Conferência do Banco de Portugal

11 - 12 de Março de 2004

Desenvolvimento Económico Português no Espaço Europeu

Portuguese Economic Development in the European Context

comunicações

proceedings

Banco de Portugal
Av. Almirante Reis, 71
1150-012 Lisboa

Economic Research Department

Printed by
Tipografia Peres, SA

Number of copies printed
800 issues

Legal Deposit no. 236982/06

ISBN 972-9479-89-5

Nota de Apresentação

Foreword

NOTA DE APRESENTAÇÃO

Esta 2ª Conferência sobre o «Desenvolvimento Económico Português no Espaço Europeu» insere-se no ciclo que o Banco de Portugal iniciou há dois anos com o objectivo de estimular trabalhos universitários de investigação aplicada aos problemas do nosso desenvolvimento económico. Não se deve estranhar este interesse do Banco nos problemas do desenvolvimento uma vez que é esse o problema central da economia portuguesa. Como se sabe, o desempenho da economia portuguesa foi bastante satisfatório ao longo de várias décadas desde os anos sessenta, situando-nos mesmo no conjunto dos 10 países do mundo com maiores taxas de crescimento. Na última década, porém, tem havido uma desaceleração sistemática da taxa de crescimento potencial. Parece que o facto de termos atingido um patamar de desenvolvimento intermédio torna agora mais difícil criar uma nova dinâmica.

O desenvolvimento económico é um fenómeno muito difícil de analisar, dada a multiplicidade de factores de que depende, e por isso mesmo a teoria económica tem vindo a alargar o campo das variáveis que considera no seu estudo. Os modelos tradicionais de Harrod-Domar ou o modelo neo-clássico de Solow situavam-se a um nível de grande abstracção «explicando» o desenvolvimento com algumas variáveis económicas e um substancial efeito residual de «progresso técnico» considerado exógeno. A endogeneização desse resíduo foi o passo seguinte que trouxe para a análise subsistemas importantes como a educação, a investigação tecnológica e o desenvolvimento de novos produtos e processos produtivos ou a qualidade das infra-estruturas em geral. Mais recentemente tem-se sublinhado a importância das «Instituições» como conjunto de normas estáveis que definem incentivos e condicionam comportamentos dos agentes económicos. Os trabalhos de Dani Rodrik , Hall e Jones, Easterly e Levine, ou de Daron Acemoglu, ilustram a importância para o desenvolvimento de diversas «Instituições», em contraponto com factores geográficos ou de disponibilidade de recursos naturais para explicar as importantes diferenças de sucesso de vários países. Que instituições? Essencialmente aquelas que asseguram as funções fundamentais de uma economia descentralizada de mercado. Instituições de criação e garantia do mercado (v.g. o papel do sistema jurídico em assegurar o primado da lei, a protecção dos direitos económicos e a execução dos contratos),

instituições de regulação do mercado, instituições de estabilização do mercado (v.g. instituições monetárias e orçamentais e instituições de supervisão prudencial do sistema financeiro) e instituições de legitimação do mercado — como o sistema político democrático ou o sistema de segurança social.

Na minha apresentação inicial da Conferência[1] procedi à análise do grau de eficácia dessas várias instituições para assegurar no nosso caso as condições propícias ao crescimento económico de longo prazo. As principais deficiências parecem situar-se na Educação, no sistema de Justiça, no sistema técnico-cientifico de Investigação & Desenvolvimento e nas regras de Finanças Públicas. Os trabalhos da Conferência que se apresentam neste volume constituem peças muito relevantes para a avaliação de algumas dessas condições institucionais que condicionam comportamentos e o funcionamento eficiente dos mercados. O debate sobre regras disciplinadoras da política orçamental, incluindo a sua dimensão regional, e a análise da eficiência de programas públicos são contributos importantes para os problemas orçamentais que o país defronta. As investigações apresentadas sobre o funcionamento do mercado de trabalho permitem, entre outros, avaliar importantes aspectos da mobilidade e flexibilidade salarial. O exame a que outros trabalhos submetem os sistemas de Regulação, Supervisão e Concorrência aplicáveis a vários sectores de actividade, constitui outro importante contributo para melhorar a eficiência de mercados em que a importância das assimetrias de informação, de externalidades com efeitos sistémicos ou de outras imperfeições, torna imprescindível a intervenção pública.

A Conferência confirmou, assim, a importância da análise das «Instituições», no sentido mais genérico do termo, para estudar o fenómeno do desenvolvimento. Os economistas não podem, porém, esquecer que não devem usar abusivamente a teoria económica como um conhecimento suficiente para escolhas fundamentais da organização social. Podemos, certamente, falar de princípios económicos e de formas de organização social que se revelaram historicamente superiores para garantir o crescimento. Por exemplo, princípios de estabilidade macroeconómica associados a regimes de baixa inflação e finanças públicas sustentáveis aparecem como importantes para o crescimento económico a longo prazo. Do mesmo modo surgem como essenciais as instituições

[1] Não incluída no presente volume, mas ver, no entanto, em www.bportugal .pt

jurídicas que asseguram o primado da lei, o Estado de Direito e a protecção dos direitos económicos de propriedade. No entanto, a variedade institucional das economias de mercado existentes no mundo, bem como as modas que sucessivamente as valorizam, revela que não existe uma forma única de assegurar aqueles princípios. Por outro lado, as escolhas subjacentes ao desenho das instituições que definem as regras do jogo económico e social estão necessariamente enformadas por valores definidores do que se entende por óptimo social e, logo, não podem ser decididas com base em análises puramente técnicas. O critério do óptimo social implica sempre mais do que um conceito de mera eficiência económica. Estas considerações servem de prevenção sobre os limites do que os economistas têm a dizer sobre os problemas do desenvolvimento a longo prazo. Trata-se, todavia, de observações que não retiram qualquer margem de importância às análises que integram os trabalhos desta Conferência, cujo contributo para a problemática do nosso desenvolvimento plenamente justifica esta iniciativa do Banco de Portugal.

Os estudos que estão na base da Conferência provieram quer de um concurso público dirigido aos Departamentos de Economia das Universidades portuguesas, quer de convites directos a alguns economistas. Seguiu-se um processo de selecção dos estudos, tarefa que foi da responsabilidade do Comité Científico, composto por Maximiano Pinheiro (Banco de Portugal e Instituto Superior de Economia e Gestão), Pedro Duarte Neves (Banco de Portugal e Universidade Católica Portuguesa), José Ferreira Machado (Universidade Nova de Lisboa), Isabel Horta Correia (Banco de Portugal e Universidade Católica Portuguesa), Pedro Portugal (Banco de Portugal e Universidade Nova de Lisboa) e Mário Centeno (Banco de Portugal e Instituto Superior de Economia e Gestão).

Agradecimentos são devidos a todos pelo trabalho realizado, tal como ao Departamento de Estudos do Banco de Portugal é devido o nosso reconhecimento pela tarefa de organização da iniciativa.

Da qualidade e conteúdo dos trabalhos seleccionados dá conta esta publicação, onde são divulgados os estudos apresentados na Conferência, que se realizou a 10 e 11 de Março de 2004, na Fundação Calouste Gulbenkian, em Lisboa.

Vítor Constâncio
Governador do Banco de Portugal

FOREWORD

This second Conference on "Portuguese Economic Development in the European Area" is part of a cycle started by Banco de Portugal two years ago, with the aim of stimulating university research work applied to the problems of Portuguese economic development. This interest of Banco de Portugal in development problems is not surprising, considering that this is the central problem of the Portuguese economy. As it is known, the performance of the Portuguese economy was quite satisfactory during several decades, from the 1960s onwards. Portugal was even among the group of 10 countries in the world with the highest growth rates. However, since the last decade the potential growth rate has been systematically decelerating. It seems that after having reached an intermediate development platform, it is now more difficult to create a new dynamic.

Economic development as a phenomenon is very difficult to analyse, given the multiple factors on which it depends. Thus, economic theory has been widening the scope of the variables considered in its study. The traditional Harrod-Domar models or the follow neo-classical model of Solow were at a level of great abstraction, "explaining" development with some economic variables and a substantial residual effect of "technical progress" considered exogenous. The endogeneisation of this residual was the next step that brought to the analysis important subsystems, such as education, technological research and the development of new products and productive processes, or the quality of infrastructures in general. More recently, the importance of "Institutions", as a set of stable rules that define incentives and condition the behaviour of economic agents, has been emphasised. The work of Dani Rodrik, Hall and Jones, Easterly and Levine, or Daron Acemoglu illustrates the importance of several "Institutions" for development, as opposed to geographical factors or the availability of natural resources, to explain the significant success differences of several countries. Which institutions? Basically those that ensure the fundamental functions of a decentralised market economy. Market creation and guarantee institutions (e.g. the role played by the legal system to ensure the primacy of the law, the protection of economic rights and the performance of contracts), market regulation institutions, market stabilisation institutions (e.g. monetary and fiscal institutions and

financial system prudential supervision institutions) and market legitimation institutions — such as the democratic political system or the social security system.

In my opening address to this Conference[1], I analysed the degree of efficacy of this group of institutions to ensure, in the case of Portugal, conditions fostering long-term economic growth. The main weaknesses seem to be education, the legal system, the technical and scientific system of R&D and public finance rules. The works of the Conference presented in this volume are relevant parts for the assessment of some of these institutional conditions that affect the behaviour and the efficient operation of the market. The discussion of rules that discipline the fiscal policy, including its regional size, and the analysis of the efficiency of public programmes are important contributions to solve the fiscal problems faced by the country. The research papers presented in this Conference on the operation of the labour market enable *inter alia* the analysis of important aspects of mobility and wage flexibility. The analysis made by other research papers of the regulation, supervision and competition systems applicable to several activity sectors, is another major contribution to improve market efficiency, in which the importance of information asymmetries, externalities with systemic effects or other imperfections makes public intervention imperative.

The Conference has thus confirmed the importance of the analysis of "Institutions", in the broad sense, for the study of the development phenomenon. However, economists cannot forget that they shall not make an abusive use of economic theory as a sufficient knowledge for the fundamental choices of social organisation. It is certainly possible to mention economic principles and social organisation forms that historically have proved to be in a better position to guarantee growth. For instance, the macroeconomic stability principles associated with low inflation regimes and sustainable public finances are important factors for long-term economic growth. Likewise, the legal institutions that guarantee the primacy of the law, the rule of law, and the protection of economic property rights are essential. However, the institutional variety of market economies existing in the world, as well as the way how they are valued over time, show that there is no single way to guarantee such principles. On the other hand, the choices underlying the design of the institutions that define the rules of the economic and social

[1] Not included in this volume, but available at www.bportugal.pt.

game are necessarily made up of values defining what is understood as social optimum and, therefore, they cannot be decided upon on the basis of purely technical analyses. The social optimum criterion always implies more than a concept of mere economic efficiency. These considerations restrain economists' assertions about long-term development problems. However, these observations do not reduce the importance of the analyses made in the research papers presented in this Conference, whose contribution to the problematics of development in Portugal fully justifies this initiative of Banco de Portugal.

The research papers presented in this Conference came from both a public call for research papers addressed to the Departments of Economy of the Portuguese universities and from the direct invitation of some economists. The scientific committee in charge of the selection process was comprised of Maximiano Pinheiro (Banco de Portugal and *Instituto Superior de Economia e Gestão*), Pedro Duarte Neves (Banco de Portugal and *Universidade Católica Portuguesa*), José Ferreira Machado (*Universidade Nova de Lisboa*), Isabel Horta Correia (Banco de Portugal and *Universidade Católica Portuguesa*), Pedro Portugal (Banco de Portugal and *Universidade Nova de Lisboa*) and Mário Centeno (Banco de Portugal and *Instituto Superior de Economia e Gestão*).

Our sincere thanks are due to all the contributors for the work they have undergone. Similarly, we place on record our recognition to the Bank's Economic Research Department for organising this Conference.

Attesting the quality of the research produced, this publication puts together the research papers presented in the second Conference "Portuguese Economic Development in the European Context", that took place on 10 and 11 March 2004, at the Calouste Gulbenkian Foundation in Lisbon.

Vítor Constâncio
Governador do Banco de Portugal

Índice

Contents

Nota de Apresentação / Foreword

Conferência do Banco de Portugal sobre "Desenvolvimento Económico Português no Espaço Europeu" / Conference held by the Banco de Portugal on "Portuguese Economic Development in the European Context"

Sessão 1 – Política Orçamental e Segurança Social / Session 1 – Budgetary Policy and Social Security

Sessão 2 – Funcionamento do Mercado de Trabalho / Session 2 – Labour Market Functioning

Sessão 3 – Avaliação de Programas Públicos / Session 3 – Evaluation of Social Programmes

Sessão 4 – Mercados e Desenvolvimento / Session 3 – Markets and Development

Conferência do Banco de Portugal "Desenvolvimento Económico Português na União Europeia" – Uma Síntese Pessoal

Banco de Portugal Conference on "Portuguese Economic Development in the European Union" – A Personal Summary

BANCO DE PORTUGAL CONFERENCE ON "PORTUGUESE ECONOMIC DEVELOPMENT IN THE EUROPEAN UNION": A PERSONAL SUMMARY

José A. Ferreira Machado[*]

1. Introduction

Following a similar initiative in 2002, Banco de Portugal organised a conference this year, held on 11 and 12 March, on the theme "Portuguese Economic Development in the European Union".

This article is a summary of the addresses made in the conference as well as the main guidelines for economic policy that can be drawn from them. This work is highly personal and represents only the opinion of the writer. There is also no claim to do justice to the technical details of the addresses themselves. This article will give more emphasis to those addresses which have a direct bearing on policy.

There were 13 addresses and they covered a wide range of themes, as diverse as "rules on public expenditure" and the regulatory framework for professions, minimum state pensions and wage mobility, the lessons to be drawn from the "40 hours law" and the articulation of sector regulators with the country's Competition Authority (available on *http://www.bportugal.pt/events/conferences/IIDEP/default.htm*). Notwithstanding the diversity of themes, it is possible to separate the addresses into two main groups. In one, the major concern relates to control of the level of public expenditure and how quality in expenditure can be assessed, with the latter point understood as the best way to reach the aims at which expenditure was targeted. In the second group the focus is on how markets function and are regulated, both for labour and for goods and services.

What is the relationship between these two themes and economic development, which was the purpose of the conference? Tiago Cavalcanti ("Business Cycle and Level Accounting: The case of Portugal") showed that the lower level of productivity in Portugal relative to the USA at the end of the 20th century (46 per cent lower) can be put down to two factors: less economic efficiency in the overall use of resources (also known as total factor productivity) and, to a lesser extent, distortions in the labour market. Economic efficiency is usually associated with the diffusion of new technologies, infrastructure quality, market competition, and the legal and institutional systems. Public expenditures provides for goods and services which have a specific value to consumers but use resources which would otherwise be available for consumption and investment. The higher the level

[*] Faculdade de Economia, Universidade Nova de Lisboa. Consultant for the Economic Research Department.

of public expenditure, or the less efficient its use, the greater the likelihood that the opportunity cost will be higher than potential benefits. In these circumstances, any increase in expenditure from levels which are already high is often compared to an adverse technological shocks. On the other hand, labour market distortions are linked to income tax and other factors such as protectionist labour laws which make it difficult to match workers with available jobs and hamper the flow of resources between companies and industries.

2. Public Expenditure: Level and Effectiveness

2.1. Expenditure Rules

In certain countries in the EU, such as Holland and Sweden, there are formal restrictions on public expenditure (known as "expenditure rules") that are imposed over and above the rules of the Stability and Growth Pact. This situation, which seems to be at the centre of attention at the moment, is somewhat paradoxical, if we look at the macroeconomic policy instruments that have been lost as a result of European monetary union. As António Pinto Barbosa pointed out in the session dedicated to Budget Policy and Social Security, this stemmed fundamentally from an "unromantic" view of governments as economic agents with limited information and their own aims, and these aims are not necessarily consistent with the "general interest" of the community. Democratic regimes have biases in favour of inflation and public expenditure and frequently use budget variables with electoral purposes in mind. The path of the Portuguese budget through the 90s is detailed in the address given by Álvaro Pina ("Fiscal Policy in Portugal: Discipline, Cyclicality and the Scope for Expenditure Rules") and it gives a very clear illustration of these two factors. Portugal recorded the biggest rise in public expenditure in the EU (3.4 pp of GDP) especially marked in primary current expenditure, which grew twice as fast in real terms as GDP. On top of this, expenditure was strikingly pro-cyclical. In the light of this evidence, one should be sceptic of arguments that oppose to "rules" (or the Pact itself) in the name of economic stabilisation. The work of Linda and Francisco Veiga ("Political-Economic Cycles in Portuguese Municipalities") is highly eloquent on this point. It demonstrates how municipal expenditure, above all in terms of investment, (and above all in the item "Other buildings and construction"), rises significantly in an election year and occasionally in the year before. Municipal councils manoeuvre economic policy to illustrate their competence at election time, thus boosting their chances of re-election.

Expenditure rules allow budgetary discipline but preserve room for automatic stabilisers to operate, since these function mainly through tax revenue. (Álvaro Pina's analysis of seven European countries shows that in Portugal tax revenue has the strongest pro-cyclical correlation.) The kind of rule being looked at here is something along the following lines: "For a horizon of T years, public expenditure (D) cannot rise above x per cent per year". This statement, however, raises a considerable number of practical problems regarding the exact definition of T, D and x.

Pina argues that the rule should have the horizon of every government mandate, with each government legitimately empowered to specify new values for the framework. António Pinto Barbosa however, in his comment on the address, argued against this, on the grounds that there should be a "constitutional" view, independent of the political calendar. As for the concept of expenditure, Pina maintains that D should be limited to the primary current expenditure of the state sub-sector. Unemployment benefits and public investment

should be excluded. It may well be that leaving out debt service payments and subsidies causes no problem, but the same cannot be said for leaving out local and regional authorities and public investment (and here the conclusions of Veiga and Veiga, op. cit. should be borne in mind). While genuine expenditure in terms of capital could well pay for itself, it is in practice very difficult to calculate the return on each slice of the "expenditure cake" and thus pick out what is genuinely capital as opposed to what is current expenditure. If therefore we leave public investment out of the expenditure frame defined by the rule, we are opening Pandora's box: the accounting creativity stemming from this would strip out the idea of discipline that the rule seeks to impose. As for the limit to expenditure growth (x), Pina proposes that growth be set in real terms (as in Holland) and not in nominal terms (as in Sweden). The point of reference should be potential GDP growth (understood as the average of the past decade, if problems involving measurement are to be avoided). This real limit would be converted annually into a nominal figure through application of the GDP deflator, laid down or validated by an independent commission. Arguments here are again divided but it seems clear that a nominal limit is more transparent and more stringent, albeit more difficult to define with all due rigour.

Independent of the specific details contained in a putative rule, what seems to be of real importance is to fix multi-year expenditure objectives. It is also of primordial importance to provide the governing and supervisory institutions with the quality and independence necessary for effective control over budgetary matters. A fundamental aspect of this discipline is also a systematic assessment of public programmes and policies.

2.2. Assessment of public service policies

The efficiency of public expenditure is an issue that can be seen from different viewpoints. We can, for example, take a raft of areas where the state intervenes and where specific policies already exist. We can then study how state intervention can be made with the least possible waste in resources. What we are looking at here is fundamentally an analysis of the frontiers of efficiency. Another view consists in assessing to what point the specific intervention yields the promised social benefits that are a priori its justification. And from this angle, we are also looking at who gains and who loses from the intervention. The two perspectives are of course complementary and are part of an overall cost/benefit appraisal of public sector programmes and policies.

There is at our disposal over 30 years of rigorous and scientifically based empirical assessment of programmes and policies. The large financial burden that many of them require justifies the widespread interest in the topic. The address of Mónica Dias ("The Empirical Evaluation of Social Programmes") focused on the methods available to evaluate public policies. The work of Gouveia and Farinha and Varejão are (methodologically varied) examples covering the assessment of specific interventions such as the minimum state pension and the "40 hours law".

Any method of evaluation requires the recreation of the (counterfactual) behaviour in an alternative economic scenario following "non participation in the programme" or if the programme does not in fact exist. In many areas, such as psychology or medicine, there are routine scientific assessments the causal effects of specific actions. Before any medicine is approved for general use, it must be subject to laboratory experiments in which patients are selected on a random basis and are given the drug, while others in a control group are given an inoffensive substance known as a placebo.

This kind of experimentation is rare in economics and other social sciences. Even if it were possible, it would not provide an assessment of programmes and policies before they are implemented. (Dias has defined this as a "structural approach" requiring the use of economic theories to build "laboratories" which allow for an evaluation of the reaction of the agents and the effects of the interventions.) But even without experimental information, it is possible to carry out assessments of programmes and policies that have in fact been implemented. The counterfactual behaviour mentioned above can be supplied by external and exogenous circumstances and this leads to what could be considered a controlled experiment (natural or quasi-experiments). We can see common examples of this in the regional variations that occur when certain programmes are implemented. Sometimes it is the passage of time ("before and after the intervention") that allows for control groups to be set up as if a controlled experiment was being undertaken. José Varejão's work, ("Reduction in the Working Week and Employment: Lessons of the 40 Hours Law") provides an illustration of a "natural experiment" since the measure in question, the law of 40 hours, can be consider an unforeseen factor; the control group will be the establishments not covered by the obligation to cut the number of hours; and the experimental group is made up of the comparable establishments which fall within the scope of the new law.

Miguel Gouveia and Carlos Farinha ("What Point Minimum Pensions?") gave an address that covers a highly relevant programme. When a pension provided according to the existing social security laws is below the minimum requirement, it is topped up to reach the minimum figure. A big number of pensioners benefit from this topping up, which is financed through the country's state budget and gives rise to considerable public expenditure. A number of governments have adopted policies leading to gradual increases in minimum pensions up to the minimum wage as part of the fight against poverty. Gouveia and Farinha show that minimum pensions are an expensive and inefficient way of reaching this objective. Only 31.25 per cent of people in a family household receiving minimum pensions are in fact below the poverty line (that is, they get an income below 60 per cent of the average income for an equivalent adult in active employment). This situation stems from the fact that pensions are not given on a "means-tested" basis. So an increase of pensions from 60 to 70 per cent of the minimum wage can be expected to lead to very low results in terms of poverty, since 64 per cent of this rise will end up being spent on people who are already above the poverty line.

The fundamental message is that the minimum wage is a poor reference point for social policy and there should be an effective means-testing procedure put in its place. The authors provided a simulation with an alternative programme, where the real situation of beneficiaries was more rigorously controlled. The result was that with half of the outlay it would be possible to achieve more significant advances against poverty. This work therefore gives us a very clear idea of something that is often mentioned but not always identified: wastage in the expenditure of public funds.

Interventions by the government do not always result in public expenditure, but they always have an effect on the way resources are allocated. "Work-share" programmes (reductions in the maximum limit that a working week can last) have been popular in Europe since the end of the 70s, partly in response to high levels of unemployment. The catchphrase has been "working less so that everyone can work". Such measures, however, have little solid basis in economic theory: there is an unequivocal rise in labour costs and this is followed by a fall in the total number of hours worked. The effects on employment are, however, ambiguous and the matter comes down to an empirical issue.

José Varejão's address ("Reduction in the Working Week and Employment: Lessons of the 40 Hours Law") looked at the impact of the law that brought in a standard 40-hour week in 1996. The variables analysed were number of hours worked, overtime, wages and employment. The study shows that the measure had the predicted effect of reducing the number of hours worked per establishment, an increase in overtime and an increase in the hourly wage rate. As for employment, there is a wide-ranging picture: in those companies where the influence of the measure is more heavily felt or where there is a big proportion of workers on the minimum wage, employment clearly falls. In aggregate terms the effect on employment is slight. As the author points out, at a time when Europe is moving towards a 35-hour week, and even if the motives are not strictly economic, it is important to bear in mind the real costs of such a shift.

Pedro Martins and Paulo Guimarães also gave addresses that have potential relevance for policy-making, in spite of the fact that neither of them refers to any specific public initiative.

Direct Foreign Investment is usually seen as having a positive impact on total productivity and thus on potential for growth in economies. Governments compete for it, often at the cost of mean huge public investments. It is therefore important to assess the real benefits provided by direct foreign investment. Pedro Martins ("Wage Differentials and Wage Spillovers of Foreign Firms: Evidence from Different Estimators") assesses this through its effects on salaries. An argument for this perspective could be the sharing of the gains from greater productivity of foreign companies.

Foreign companies pay on average 32 per cent more than domestic companies. However, controlling for features such as the educational level of the work force, the sector and size, those pay differences vanish. Does this mean that direct foreign investment has no impact? Not necessarily. In the first place, wages may not be the best indicator for assessing these benefits. Moreover, the characteristics mentioned above that "explain" the wage differential, are more frequently associated with foreign companies. The advantage of being foreign may therefore not derive from invisible factors associated with "residence", but rather to the facts that these companies are bigger and employ a more highly-educated work force.

Political decision makers are often concerned with the asymmetries in regional distribution of economic activities and are receptive to programmes that involve relocation of these activities. It is therefore of considerable interest to know which of these initiatives are more likely to be successful and which are the economic activities which are most likely to respond to such initiatives. These are the questions that are answered in the work of Paulo Guimarães and Octávio Figueiredo ("Location and Localisation of Portuguese Manufacturing Industries").

This study shows that there is only a small number of sectors which are highly clustered (above what would be expected from the mere endowment of resources in a region or from a random allocation). In some of these sectors, involving traditional industries, this clustering stems from external factors associated with historical specialisation. Some technologically advanced industries (for example, radio and television, cars or pharmaceuticals) are also found in clusters, suggesting the existence of "knowledge externalities". For these industries, public incentives for relocation will only have a modicum of success, or else the success will demand greater spending. In terms of areas of intervention, increased accessibility to the Porto-Lisbon coastal corridor is the most likely measure to succeed in spreading the economic activity throughout the country.

3. Markets

3.1. The Labour Market

The weak dynamics of labour flows both between companies and between employment and unemployment in Portugal brings with it high costs in terms of efficiency (vide Blanchard and Portugal, 2002). The existence of flexible markets is fundamental for workers and companies to find the matchings that make employment more productive. Moreover, in a context where the needs for employment oscillate tremendously, both in localisation (company, industry or region), and in skills, it is crucial to have a labour market which functions efficiently.

The adjustment costs which characterise the Portuguese labour market imply among other things longer average unemployment spells and a higher number of companies which close (as an alternative to reducing manpower). Anabela Carneiro ("The Economic Consequences of Worker Displacement in Portugal") calculates the loss in wages among individuals who spend periods out of work following closures. This provides another measure of the costs associated to unemployment. The loss of wages is substantial and has an effect up to 4 years after the episode concerned: wages are up to 11 per cent less than if the worker had not lost his or her job. A significant part of this loss derives from the loss of seniority in the company, which suggests the destruction of specific human capital. It is interesting to note that this cut in wages begins a year before the job loss. Faced with the risk of closure, workers are prepared to make concessions on wages. This fact in itself is symptomatic of the lack of fluidity in the market and the anticipation of a long period out of work. There are measures that could be taken to rein in these adjustment costs: an extension to the period for notification of redundancy and other measures that increase the possibility of finding new employment (help with job-searching and professional training).

Curiously enough, in spite of the rigidity of the institutions in the labour market, there do seem that some elements of flexibility spring here and there, providing ways of responding to the needs of employers and workers. From this angle, increased mobility becomes especially relevant.

Ana Rute Cardoso ("Wage Mobility in Portugal") compared wage mobility in Portugal and the United Kingdom, two economies with different levels of flexibility. It would be expected that the more restrictive framework in Portugal would lead to less wage mobility, i.e. fewer possibilities to change the relative position in the wage distribution. In other words, less labour flows between companies could stultify salary scales. This in fact does not happen, since Portugal and the United Kingdom are not very different in terms of mobility. There is considerable evidence that wage flexibility is high in Portugal, in spite of the lack of mobility (or perhaps because of it). For example, it is well known that average wages are sensitive to economic cycles (OECD, 1992) and that employers manage to adapt salaries that are based on collective contracts to the specific conditions of the company and the workers (Cardoso (2000) and Cardoso and Portugal (2003)). The mobility visible now reinforces this conclusion, albeit from a different angle.

Francisco Lima ("How much Mobility? Careers, Promotions and Wages") looks at other aspects of mobility, specifically workers transitions between firms and the gains thus made. A significant number of jobs are stable and long-term (40 per cent of workers over 35 have been with their company for over 10 years) but there is also a high proportion of short term links, where on average 39 per cent terminate in under a year (either on the

initiative of the worker or the employer). This situation relates mainly to young workers and is very often pejoratively referred to as "precarious". In practice, however, leaving a company tends to increase wages. This is all the more so, the younger and better educated the worker, and for transitions which occur to larger companies. From this point of view, labour mobility may take the place of careers inside an organisation, and they illustrate the gains resulting from a good match between employer and employee.

3.2. Competition

There is a consensus among economists these days on the importance of institutions for economic growth (in this context, see Tavares, 2000). Which of these are important and the reasons for their importance are more open to debate. An exception can clearly be made in the existence of independent regulators who stimulate and defend competition in markets. The two addresses presented to the conference on this theme do not approach the issue in a general and abstract way. They look at specific institutions in Portugal – the Competition Authority and the Professional Associations – and they look at which structure best stimulates competition.

Pedro Barros and Steffen Hoernig ("Sectoral Regulators and the Competition Authority: Which Relationship is best?") come to the conclusion that the overlapping of jurisdiction between entities concerned with specific sectors and the Competition Authority (which covers all sectors) may be beneficial for encouraging social well being, but there should be limits to the cooperation between them. In particular, there should not be a process of joint decision-making that means that a decision is only reached if both authorities are in agreement.

Nuno Garoupa ("Regulation of Professions in Portugal: A Case Study in Rent-Seeking") looks at how good is the regulation by two entities that are in practice the regulators of the sector: the Lawyers and the Physicians Associations. The regulation of professional activities could be based on objectives that are "in the public interest" (as a result of the asymmetry of information, for example). Many of the norms that act as regulatory devices are, however, merely of private interest, that is the wish to increase the income of established professionals. In the first category, there are, for example, the guarantees of adequate specialist training; in the second there are the restrictions on fees or the disciplinary powers that are exclusive to the Associations. The author has built an index of the quality of regulation and found that in the case of the medical profession, Portugal is the EU country where norms motivated by private interest are the most predominant. In the case of lawyers, although belonging to the group with lower regulatory quality, Portugal is not substantially different from Germany or France in this regard.

The study proposes various reforms. In terms of eliminating "restrictions to entry", I would point out the opening of the Portuguese legal market to international groups or, for the medical profession, a reduction in the influence of the Association in the process of academic training and the stimulation of competition between the medical faculties so as to give a significant boost to the number of doctors. Finally, the author proposes that the Associations should be within the jurisdiction of the Competition Authority, where restrictive practices could be detected.

4. Conclusions

In the 1st Banco de Portugal conference on economic growth in 2002, there were general recommendations regarding policies, focusing on the bases for economic growth, such as, for example, the need for investment in infrastructures, improving the education system, making the labour market more flexible and reforming the legal system and practices.

This, the 2nd conference, looked at more specific aspects of growth. In terms of the policy conclusions which can be drawn, they can be summarised as two main ideas:

1. The actions of governments must not be viewed naively. Their intervention is not always motivated by "the idea of common good", or necessarily generates net positive benefits. This is true in macro economic policy and it is enough to look at the evidence presented to the conference on the way public expenditure evolution amplifies the effects of cyclical fluctuations in the economy or on "electoral opportunism" in the investments of local authorities. This is also true for specific programmes and policies: the minimum pensions scheme was shown at the conference to be an ineffective instrument in the fight against poverty. There was also the evidence relating to the cost in working hours and consequently in production stemming from the 40 hours law approved in 1996.

Two consequences flow from this situation: on the one hand, there is the need for formal rules on public expenditure that impose intertemporal discipline on governments. Moreover, there is a need for regular assessment of programmes and policies, both ex-ante and ex-post, so as to provide checks on the promises of social benefits which give a reason for their existence and identify who gains and who loses from the intervention.

2. Flexible market functioning conducive to competition is fundamental for an efficient allocation of resources. The existence of restrictions to entry into industries or professions is a common form of ensuring economic rents to those who are already established, at the cost of social well being. These restrictions are frequently presented as being in defence of "the public interest". As with our comment on the actions of governments, we should not be naive about this. For these reasons, the activity of regulatory structures (specifically the Competition Authority), with a scope possibly extended to the professional associations, can make an important contribution to economic growth.

References

Blanchard, O and P. Portugal. "What hides behind an Unemployment rate: comparing the Portuguese and US labor markets" *American Economic Review*, 91(1), 2001.

Cardoso, A.R. "Wage differentials across firms: an application of multilevel modelling" *Journal of Applied Econometrics*, 15(4), 2000.

Cardoso, A.R. and P. Portugal. "Bargained wages, wage drift and de design of the wage setting system". IZA *Discussion Paper* 941, 2003.

OECD, *Economic Surveys*, Portugal, 1992.

Tavares, J. "Firms, Financial Markets and the Law: Institutions and Economic Growth in Portugal". Proceedings of 1st Banco de Portugal Conference on "*Desenvolvimento Económico Português no Espaço Europeu: Determinantes e Políticas*", 2002.

Sessão 1 - Política Orçamental e Segurança Social
Session 1 - Budgetary Policy and Social Security

FISCAL POLICY IN PORTUGAL: DISCIPLINE, CYCLICALITY AND THE SCOPE FOR EXPENDITURE RULES*

Álvaro M. Pina (ISEG/UTL)

Abstract

This paper first studies two aspects of Portuguese fiscal policy: its cyclical behaviour and recent trends in public accounts. Unlike in most European countries, primary spending has been strongly procyclical, and its fast growth in the nineties proved an obstacle to fiscal consolidation. Aiming at both discipline and stabilisation, the paper then proposes a multiyear expenditure rule based on stable real growth of a spending aggregate. The definition of this aggregate and the articulation of the rule with annual budgets are discussed in some detail. Finally, it is argued that creating an independent fiscal committee would enhance the rule's enforceability.

* Paper prepared for the II Conference "Desenvolvimento Económico Português no Espaço Europeu", organised by Banco de Portugal. I wish to thank the Scientific Committee of the conference for valuable comments and suggestions. The paper also benefited from conversations with António Afonso, João Ferreira do Amaral, Manuela Arcanjo, Jorge Santos and Miguel St. Aubyn. All errors are solely mine, as is the responsibility for the opinions expressed in the text.

1. Introduction

Since the late seventies, discretionary macroeconomic policy, especially of the fine-tuning type, has been regarded with suspicion, and most macroeconomists have favoured rules-based policy regimes. Arguments for rules can be grouped into two main categories, which I call, borrowing some terminology from Drazen (2000), technical and political.

On the technical side one finds a host of factors that can make discretionary interventions deliver uncertain effects, be simply ineffective or, even worse, turn out harmful. Such factors include the problem of significant lags in policy decision, implementation and/or transmission, as well as the awareness of uncertainty about the "true model" of the economy. Perhaps even more importantly, the influential paradigm of dynamic optimization under rational expectations emphasized the problem of time inconsistency, leading to the well-known inflation bias of discretionary monetary policy, but also present in some fiscal policy decisions. The same paradigm also emphasized the concept of Ricardian equivalence, with the implied ineffectiveness of deficit spending.

Political arguments for rules hinge upon the negative consequences of agents' heterogeneity and conflict of interests (Drazen, 2000) for discretionary policy actions. These can include, among others, the manipulation of policy variables to increase the chances of reelection or to pursue partisan goals; "common pool" problems leading to a spending or a deficit bias; or delays of necessary policy measures due to conflicts over burden-sharing.

Monetary policy was the first to undergo major institutional changes in the wake of the widespread preference for rules over discretion. Since the eighties many countries have significantly upgraded the level of independence of their central banks, giving them a clear mandate for price stability.

Over the past decade, fiscal policy has started to follow suit as regards the adoption of rules-based frameworks. A large literature on how institutions affect fiscal outcomes has developed, and many countries have pursued, or are pursuing, institutional reforms in the domain of fiscal policy (OECD, 2002). This "preference for rules" has been maintained even though the more extreme views about the ineffectiveness of discretionary policy have lost popularity. Long mistrusted, fiscal policy as a stabilization tool is regarded today with renewed interest. There is widespread consensus that automatic stabilisers should be allowed to operate. A cautious, even sceptical, attitude remains as regards discretionary interventions, but even the latter are acknowledged as a valuable, if imperfect, tool in situations where monetary policy is not available (Taylor, 2000; Wren-Lewis, 2000) – such as country-specific shocks in a monetary union.

In Europe, concerns with the unprecedented accumulation of public debt in the seventies and eighties were compounded by the risks posed by an undisciplined fiscal policy to the smooth operation of a monetary union, both in terms of an unbalanced policy mix and of ultimate threats to the independence of the European Central Bank (ECB). Artis and Winkler (1998) and Sapir *et al.* (2003), among others, argue that such risks are indeed the main justification for the Maastricht Treaty and the Stability and Growth Pact (SGP). These renowned examples of deficit targets belong to the wider class of numerical rules, which in turn are a subset of budgetary institutions (Alesina and Perotti, 1999). Besides adopting the SGP, European Union (EU) countries have also reformed their national fiscal

frameworks, and, among the wide variety of arrangements, expenditure rules enjoy a growing popularity (European Commission, 2003).

This paper starts by studying the performance of fiscal policy in Portugal both as regards the process of deficit reduction that led to the adoption of the euro, and from the viewpoint of cyclical stabilization. Whenever possible, comparisons with other European countries are performed. The analysis is carried out at a disaggregated level, considering the main revenue and spending items and decomposing the Portuguese general government into its subsectors.

Results stress the need for expenditure control. In the nineties, current primary outlays grew twice as fast as GDP (in real terms). Further, spending denotes a strong procyclical (and hence destabilizing) pattern. In the light of such findings, this paper then considers how an expenditure rule could be designed in Portugal. The advantages of spending targets over deficit targets are summarized. Then it is discussed how to delimit the aggregate to be targeted and how its growth rate should be set. Since deficit constraints still apply (the SGP), the remaining budget items (e.g. revenues) are also addressed; and since a multi-year horizon offers advantages, its articulation with annual budgets is considered. Finally, the paper analyses the scope for an independent authority in designing and monitoring budget policy, and tentatively assesses the proposed expenditure rule in the light of standard criteria.

Fiscal policy is complex and multidimensional – actually, multidisciplinary. This analysis of expenditure rules focuses on their macroeconomic dimension, and in particular on avoiding procyclical behaviour. In doing so, the present paper attempts to be more comprehensive and detailed than previous proposals of spending targets, such as ECORDEP (2001) or OECD (2003). Other dimensions are equally important, but they are not pursued here – for instance, the institutional implications and/or requirements of spending rules as regards the relative power of the Finance Minister in the cabinet, or the budgetary competences of Parliament. Still other aspects are only briefly touched upon, and would deserve further examination elsewhere: examples are the required timely information on the budgetary execution, the translation of an overall target into the spending allocations of individual government agencies, and the implications for their management.

The paper is structured as follows. Section 2 analyses the empirical record of Portuguese fiscal policy. Section 3 addresses the design of expenditure rules. Section 4 concludes.

2. Fiscal Policy in Portugal

2.1. Public Accounts and Fiscal Consolidation in the 1990s

The process of deficit reduction in the years preceding the introduction of the euro has been widely studied, both from a cross-country perspective (e.g. von Hagen *et al.*, 2001) and with a focus on the Portuguese economy (e.g. ECORDEP, 2001). In this section I revisit the evolution of the Portuguese public accounts between 1990 and 2000, making a comparison with the other European Union (EU) countries and drawing on the recent availability of ESA95 data for the whole period. I also decompose general government into its main subsectors, and briefly describe their relative performance.

Table 1 shows the magnitude of the main revenue and expenditure items in the year 2000, in percentage points (p.p.) of GDP, as well as their changes from 1990 to 2000 (also in p.p. of GDP)[1]. Table 2 focuses on fiscal consolidation, by considering its most widely used indicator – the evolution of the cyclically-adjusted primary balance (CAPB), decomposed into revenue and spending. Combining the information contained in the two tables, one can highlight some prominent features of the Portuguese fiscal performance.

[insert Table 1 about here]

On the revenue side, Portugal recorded the second highest increase in total revenues as a share of GDP (after Greece) – 6.8 p.p. in actual figures and 6.6 p.p. in cyclically-adjusted terms – evenly spread between the two halves of the decade (Table 2). The ensuing weight of revenues in 2000, however, was still somewhat below the EU average (42.3 per cent of GDP against an EU unweighted average of 47.8).

On the expenditure side, the Portuguese performance combined a broad convergence to the European average in levels with a deeply divergent trajectory in terms of spending dynamics. Total expenditure as a percentage of GDP, the lowest of the 15 countries considered in 1990, essentially caught up to the European average by 2000 (45.4 against 46.9 per cent), and even exceeded it in some items, such as compensation of employees. This convergence was the result of the highest increase in total spending of the EU (3.4 p.p. of GDP), in a context where most countries were actually decreasing the weight of public outlays (besides Portugal, France and Germany were the only exceptions).

The contrast in trends between Portugal and the other EU members is magnified if one looks at the composition of expenditure. There was a very strong increase in current primary spending (8.2 p.p. of GDP in cyclically-adjusted terms), accompanied by a virtual stabilisation of capital expenditure relative to GDP[2] (not out of line with the average behaviour of this item in other countries), and by huge savings in interest payments (5.4 p.p. of GDP). Portugal actually made bigger savings in these payments than more highly indebted countries, such as Italy or Belgium, due to faster convergence of nominal interest rates, especially in the first half of the decade (Portuguese rates were initially much higher) – see charts in annex 2.

Hence, the Portuguese deficit reduction in the 90s was made possible by both higher current revenues and lower interest payments, and was accompanied by an unparalleled rise in current spending, which actually absorbed most of the additional resources. Looking at the evolution of the CAPB (Table 2) actually yields the striking

[1] All EU countries have fiscal data on a ESA95 basis available from 1995 onwards, and most of them (Portugal included) already have longer series. Throughout this paper, and with the exception of Table 1, my choice has been to use only ESA95 data, which implies that sometimes the Portuguese experience is not compared to each and every EU member, but only to a subset of countries. In these cases, an alternative would have been to complete the dataset backwards by applying ESA79 growth rates to ESA95 levels. This technique is straightforward when considering series in isolation, but entails some problems when applied to series observing add-up constraints (e.g. total revenue as the sum of several items), as such constraints cease to hold. Because of this limitation, and, more fundamentally, because ESA79 and ESA95 figures are not strictly comparable, no ESA79 series were used.

[2] Notice, however, that the figures reported in Table 1 somewhat understate the growth of this item, since in 2000 expenditure was noticeably lower than in adjacent years (mainly due to a low execution of the new Community Support Framework). In 1999 capital expenditure stood at 5.9 per cent of GDP, 0.7 p.p. higher than at the beginning of the decade. Still, the sharp contrast with current primary spending remains.

conclusion that no fiscal consolidation took place: the CAPB deteriorated by 1.5 p.p. between 1990 and 2000, especially due to its performance in the second half of the decade, when Portugal was the only EU country not to consolidate[3].

[insert Table 2 about here]

Table 3 disaggregates the increases in Portuguese primary spending into the subsectors of general government (GG): Central Government (CG, in turn split into State and Autonomous Services and Funds (ASF)), Local Government (LG) and Social Security Funds[4] (SS). For those items where intersectoral transfers take place, only the amounts paid to the rest of the economy are considered, so that the subsectors add up to the consolidated figures for general government. The focus is on spending, and not on revenue, since a large proportion of the non-State sectors' revenues consist of transfers from the State[5].

[insert Table 3 about here]

The information contained in Table 3, being expressed in p.p. of GDP, reflects both the rate of growth of each item in each sector and the relative importance of each sector in the aggregate figures for the several items. In order to disentangle these two determinants, Table 4 indicates average real growth rates and the subsectors' shares. If one wishes to have a benchmark for comparison, real GDP grew at an average rate of 2.8 per cent.

[insert Table 4 about here]

One observes that, as far as intermediate consumption is concerned, the surge in the expenditures of the ASF and LG sectors account for the overall high growth of this item during the 1990s (6.3 per cent annually in real terms), with a significant impact in terms of GDP shares (1.0 and 0.5 p.p. for ASF and LG, respectively).

As regards compensation of employees, the cross-sector variation in growth rates is smaller, due to the centralised nature of many pay scales[6]. Still, the ASF sector stands out

---

3 Some studies define fiscal consolidations based on the year-on-year improvement of the CAPB, and judge their success on grounds of the ensuing reduction of the debt ratio in the following two or three years. Afonso (2002), taking this approach, identifies a successful fiscal consolidation in Portugal in 1995, which is seemingly at odds with the present study. The two analyses, however, are not contradictory: (i) I look at the CAPB evolution over several years, not at yearly changes; (ii) the 1995 CAPB change still belongs to the first half of the decade; (iii) the debt ratio actually decreased in the years after 1995 mainly due to privatisation revenues and good nominal GDP growth (relative to interest rates). This last point is further discussed in section 3.4.

4 ESA95 codes are S1311, S13111, S13112, S1313 and S1314, respectively. Local government includes regional administration. I have considered the (tiny) sector of Non-profit-making Institutions of the Central Government (S13113) together with the Autonomous Services and Funds.

5 And even as far as the remaining revenues are concerned, the autonomy of non-State subsectors is often small. For instance, the rates of local taxes must be set within strict ranges (Bronchi, 2003, p. 38).

6 Needless to say, spending on this item also depends on the volume of employment. In this area, however, statistical data is extremely scarce: I am not aware of any series of employment in the general government covering the whole decade, let alone series for the several subsectors.

as the main contributor to the expansion of this item relative to GDP (1.5 in a total of 3.2 p.p.).

Social transfers are the item responsible for the highest expansion of spending as share of GDP (3.6 p.p.). Its behaviour largely determines that of the SS sector, and vice versa, though the ASF sector also plays a non-negligible role.

Subsidies decreased in relative and even absolute (real) terms, and are nowadays of marginal importance in overall expenditure. The opposite took place with other current expenditure, a residual category that reached double-digit real growth during the 1990s[7]. The bulk of this item consists of current transfers.

Turning now to capital expenditure, and considering first gross fixed capital formation (GFCF), its growth rate (4.3 per cent) was smaller than current primary expenditure's, though still above GDP's. Sectoral trends, however, were extremely diverse, with local government alone recording a rise of 0.7 p.p. of GDP. As regards other capital expenditure (mainly investment grants and other capital transfers), the overall real growth rate was negative, though, again, with disparate sectoral evolutions. Spending by local government grew at an impressive annual rate of 17 per cent, but starting from a very small base, and thus increasing a mere 0.2 p.p. of GDP over the decade.

Overall, all the subsectors took part in the strong rise in expenditure that characterised the 1990s in Portugal. However, by jointly considering spending categories and institutional sectors, one can identify two above-average spenders: ASF, as regards intermediate consumption and personnel costs, and LG, as regards intermediate consumption and capital expenditure.

2.2. The Cyclical Behaviour of Fiscal Policy

Since the late 1990s there has been a revival in research on how fiscal policy interacts with the economic cycle. One of the strands in the literature has tried to quantify the effects of fiscal policy on output, often resorting to structural vector autoregressions (e.g. Blanchard and Perotti, 2002): Keynesian effects are generally found to prevail, though with relatively small multipliers. Another set of studies has focussed on the cyclical behaviour of fiscal policy – whether it tends to be pro or anticyclical – mainly by estimating fiscal reaction functions which include some measure of the business cycle: examples are Wyplosz (2002), Hallerberg and Strauch (2002), Lane (2003), Auerbach (2002) and Gali and Perotti (2003). Probably due to the multiplicity of national experiences, no consensus has emerged so far: for instance, the European Commission (2000) holds that procyclicality prevails, whereas Wyplosz (2002) tends to draw the opposite conclusion.

This section is concerned with the latter issue, and thus estimates some fiscal reaction functions for the Portuguese economy and for other European countries. A subset of these techniques is also applied to the subsectors of the Portuguese general government. Both the sample length and the country coverage are guided by the availability of ESA95 data (recall note 1), which starts in 1977 for Portugal. The main interest lies in determining

[7] Comparisons are affected by the circumstance that in 2000 subsidies were exceptionally lower, and other current expenditure somewhat higher (relative to GDP) than in adjacent years - a problem similar to that of note 2. Like in that note, however, the described trends would remain qualitatively unchanged if 1990 were compared with 1999 instead.

whether the overall stance of fiscal policy has been procyclical or countercyclical, considering both revenues and expenditures (and their main subcategories), and separating discretionary actions from the operation of automatic stabilisers.

As is well known, fiscal policy is said to be procyclical (countercyclical) when it contributes to increase (dampen) output fluctuations. For instance, in the event of a cyclical downturn (more accurately, of a decrease in the cyclical component of GDP), a procyclical fiscal stance translates into lower spending and/or higher taxes. Bearing in mind that a given variable is considered procyclical (countercyclical) when its cyclical component is positively (negatively) correlated with the cyclical component of output, procyclical fiscal policy implies that spending variables are procyclical and tax revenues are countercyclical. For this reason, and to avoid ambiguity, when addressing the behaviour of taxes the adjectives "procyclical" and "countercyclical" will refer to the fiscal policy stance, and not to the classification of tax variables according to the above criterion. To give an example, "countercyclical revenues" will denote a countercyclical behaviour of fiscal policy on the revenue side, collecting lower taxes in downturns.

I start by presenting the results obtained with the simplest specification – that of Lane (2003). For each budget item, an indicator of cyclicality is obtained by regressing its real growth rate on the growth rate of real GDP (*Y*). Formally, an index of cyclicality of budget item *X* is given by parameter β in the regression:

$$\Delta \log(X_t) = \alpha + \beta \Delta \log(Y_t) + \varepsilon_t \qquad (1)$$

Conversion into constant prices uses the GDP deflator, to capture spending increases both through volumes and through relative prices (Lane, 2003). The economic cycle is proxied by real GDP growth, thus avoiding the need to compute potential output (see Lane, 1998, pp. 11-12 and other references therein). However, due to the wider use of the output gap as an indicator of the cyclical position of the economy, I will also use it as a regressor in a related specification – see annex 4.

Regression (1) was run for several spending categories, as in Lane (2003), and also for the main revenue items. Table 5 reports the β coefficients and whether they are statistically significant.

In the case of revenue, positive values for β imply that fiscal policy acts in a stabilising (countercyclical) way. The same applies to a negative β for an expenditure item, implying that spending slows down when economic growth is faster. On the other hand, procyclical fiscal policy is characterised by positive β for expenditures, and negative β for revenues.

[insert Table 5 about here]

All seven countries have positive coefficients for current and total revenues, which therefore tend to move in a stabilising way. Most coefficients for specific revenue items are also positive, and the few negative estimates are not significantly different from zero. Positive values for β are partly due to automatic stabilisers, as discussed below.

Portuguese values are the highest, suggesting that revenues display a stronger countercyclical behaviour that in the other countries under consideration.

As regards expenditure, results place Portugal in sharp contrast to the six other European countries – a sharper contrast than for revenues, since now Portuguese and foreign coefficients often have the opposite sign. While abroad spending is either countercyclical or relatively neutral, in Portugal it is strongly and significantly procyclical. The items that contribute the most to this destabilising behaviour are intermediate consumption and personnel costs[8]. Conclusions hardly change when using Hodrick-Prescott cyclical components instead of growth rates in equation (1) – see annex 4.

Automatic stabilisers are inherently countercyclical, contributing to generally positive β coefficients in the case of taxes, and negative β values in the case of social transfers. To remove their influence and thus concentrate on the stance of discretionary policy, I have rerun the regressions for cyclically-sensitive items using cyclically-adjusted values.

The estimation results, given in Table 6, are broadly supportive of the previous conclusions. As one would expect, removing the effect of automatic stabilisers lowers revenue coefficients, and somewhat increases expenditure's. But Portuguese discretionary fiscal policy can still be characterised as countercyclical on the revenue side, and procyclical on the expenditure side. In both cases, there remain large differences vis-à-vis other countries.

[insert Table 6 about here]

Some papers assess cyclicality while controlling for other factors, such as a solvency-ensuring response to past deficits or debt. Examples can be found in Bohn (1998), Wyplosz (2002) or Auerbach (2002). I adapt a specification of the latter with two purposes in mind: first, to check the robustness of previous conclusions on the stance of discretionary fiscal policy; and second, to study whether the opposite stances of revenues and outlays roughly cancel out, or, instead, whether one of them prevails in overall terms.

For each country, the three equations below were estimated[9]:

$$\Delta s_t^{CA} = \alpha_1 + \beta_1 GAP_t + \gamma_1 s_{t-1} + \varepsilon_t \quad (2)$$

$$\Delta t_t^{CA} = \alpha_2 + \beta_2 GAP_t + \gamma_2 s_{t-1} + \mu_t \quad (3)$$

$$\Delta g_t^{CA} = \alpha_3 + \beta_3 GAP_t + \gamma_3 s_{t-1} + \nu_t \quad (4)$$

8 Notice that the coefficient for social transfers other than in kind (0.3) is also positive. This may seem striking, since these transfers include the anticyclical unemployment-related expenditure. However, pensions account for the bulk of this item, whereas unemployment benefits stand for a meagre fraction of the total (5.4 per cent in 2000, from Tables 1 and 11, though their accounting methodologies are not strictly comparable). Further, although positive, the coefficient is not statistically significant.

9 The first of these equations mirrors Auerbach (2002, Table 1), except that his data is quarterly, and he uses the lagged gap. The second and third simply decompose the budget balance into revenues and outlays.

Variables s^{CA}, t^{CA} and g^{CA} are the cyclically-adjusted budget balance, total revenue and total spending in p.p. of potential GDP. *GAP* denotes the output gap (also in p.p. of potential output) and *s* the actual balance in p.p. of GDP. Table 7 displays the estimation results (omitting the constant).

[insert Table 7 about here]

In the light of the above evidence, Portuguese discretionary fiscal policy has been fairly neutral: countercyclicality on the revenue side approximately compensates procyclicality in spending. One also finds a solvency-preserving feedback to past deficits (the -0.49 coefficient), working almost exclusively through revenues. Abroad, overall stances also tend to be neutral, with the exception of Denmark and Finland, where policy is found to be significantly countercyclical. The countercyclical nature of revenue in Portugal is the strongest of the seven countries considered, as is the procyclical behaviour of spending (with the exception of Italy)[10].

With different data and methods, results in this paper are therefore in line with those of Marinheiro (1998): Portuguese revenues are countercyclical, whereas spending is procyclical.

Finally, to conclude this section, regression (1) was performed for the primary expenditure items of the Portuguese general government subsectors (Table 8). Some of the ensuing coefficients must be interpreted with care, as the respective items are not significant from an economic point of view (e.g. the capital expenditure of the SS sector). One should also bear in mind that differences between sectors in aggregate items (such as current primary expenditure) partly reflect composition effects: for instance, the low procyclicality of the SS sector essentially mirrors the coefficient for overall social transfers. Still, there is a tendency for higher procyclicality in the ASF and LG sectors, mainly associated to the behaviour of intermediate consumption, wages and non-social current transfers.

[insert Table 8 about here]

2.3. The Case for Spending Control

From the analysis above one draws a double motivation for reforming Portuguese public expenditure:

- In the recent past, strong increases in spending have been the main obstacle to deficit reduction (recall section 2.1);
- Spending has shown a strong procyclical pattern, thereby hampering the effectiveness of fiscal policy as a stabilization tool.

It is well known that public expenditure has a high degree of rigidity, as most of it follows from legal obligations or contracts (ECORDEP, 2001, pp. 54-59). One should note,

[10] The small number of observations limits the scope for the analysis of subsamples. Still, drawing on Wyplosz (2002), I have studied whether the response to the output gap changed in the nineties, due to the Maastricht consolidation process. Results are available upon request and point to a stronger cyclical sensitivity of revenues in the last decade. No significant changes occur on the expenditure side.

however, that the items tending to display the highest cyclicality coefficients – intermediate consumption and gross fixed capital formation – mostly have a discretionary, non-mandatory nature, and are therefore more flexible.

The analysis in this section also shows that the need for expenditure control concerns all the subsectors of general government – especially those which are more decentralized, such as the autonomous services and funds or the local governments. Hence there is a case for stronger coordination and control within the public sector – as recent legislation has acknowledged, although in an imperfect way (see section 3.1).

3. Expenditure Targets

3.1. An Overview

The reform of fiscal rules is high on both the research and the policy agenda. Among them, expenditure rules enjoy a growing popularity. The latest report on Public Finances in EMU shows that virtually all EU15 countries have now in place some form of spending rule (European Commission, 2003, pp. 135-140). These range from recruitment restrictions in the civil service (e.g. Portugal, Greece) to fully articulate medium-term fiscal frameworks (e.g. Netherlands, Sweden, UK).

An empirical assessment of expenditure rules is made difficult not only by their diversity but also by the fact that in most countries rules were introduced in the late nineties or later (Dutch rules are among the oldest, but still dating back only to 1994). While awaiting the accumulation of empirical evidence, their growing importance rests on two main factors, which are closely related to the motivations mentioned in section 2.3. First, they can be an instrument to control and reform public expenditure, both quantitatively and qualitatively. Second, they are largely compatible with the automatic stabilisers, since these operate mainly on the revenue side. Mills and Quinet (2001) argue that "a spending rule is in essence close to a cyclically-adjusted balance target" (p. 327), but more transparent and avoiding the pitfalls of cyclical adjustment (further developed in section 3.3 below).

Unlike an expenditure target, a deficit target often implies procyclical policies, because it implies that discretionary measures will be taken to counteract the effects of automatic stabilisers, and thus ensure that the planned deficit-to-GDP ratio is attained. An extreme illustration consists in having a constant deficit target at all times (for instance, a strict balanced budget requirement), forcing discretionary policy to be systematically procyclical. A milder version of the problem occurs when the government seeks to attain the planned deficit even if, during budget execution, the cyclical position of the economy diverges from the budget's macroeconomic scenario. In this case the necessary intra-annual adjustments will be procyclical, though the overall policy stance may or may not be so, as it will depend both on those adjustments and on the planned stance.

In the Portuguese case, the stress on expenditure control has been advocated in a number of recent studies, the most prominent of which is perhaps the ECORDEP report (ECORDEP, 2001). However, recent budgetary laws emphasise deficit targets instead. The 2001 Framework Law for the State Budget, as amended by the 2002 Budgetary Stability Law[11], prescribes a non-negative primary balance for the State (article 20.1) and a non-

[11] *Lei de enquadramento orçamental* (Law no. 91/2001, on 20th August) and *Lei da estabilidade orçamental* (*Lei Orgânica* no. 2/2002, on 28th August), respectively.

negative balance for the Social Security (article 25.1) and for each and every autonomous service or fund (article 22.1), though with escape clauses. In a similar vein (viewing debt increases as roughly equal to deficits), article 84 sets annual borrowing limits by subsector, which may override those implicit in the Local Finance Law and the Finance Law of Autonomous Regions[12]. The criticisms set out in the previous paragraph hence apply[13].

This paper attempts to go further than previous proposals on how to implement expenditure ceilings in Portugal. Section 3.2 discusses the delimitation of the target, both in terms of categories of outlays and in terms of the subsectors of general government. Section 3.3 then addresses how to choose the appropriate level (or growth rate) for the ceiling, as well as the time horizon of objectives. The overall design of the budget is the object in section 3.4, which takes into account the fraction of spending left out of the ceiling, and the interactions between overall expenditure, projected revenues (under a given macroeconomic scenario), the budget balance and the path for debt. Section 3.5 tackles issues of inter and intra-annual adjustment in the face of unexpected developments, and briefly discusses problems of statistical information. Section 3.6 relates some of the previous points to the currently fashionable proposals of independent fiscal authorities. Finally, section 3.7 tentatively assesses the merits of the proposed expenditure ceilings in the light of standard criteria for the quality of fiscal rules.

A final point to make concerns the interaction between the proposed system of expenditure targets and the SGP. The latter is facing considerable tensions, due to the excessive deficit situation of several countries (Portugal, France and Germany, at the time of writing), and criticism abounds, ranging from proposals for incremental improvement (Buti, Eijffinger and Franco, 2003) to calls for more radical reform (e.g. Wyplosz, 2002). It is not the aim of this paper to contribute to such debate. Rather, SGP rules will be taken as given, and expenditure ceilings are proposed as an *additional* fiscal policy framework, compatible with the Pact's discipline – especially as regards the prohibition of deficits over 3 per cent of GDP.

3.2. Delimiting the Target

A first decision to take regards what to include in, and exclude from, the spending aggregate to be targeted. Though this aggregate should be broad, certain items – such as interest payments, public investment or unemployment-related transfers – are commonly regarded as candidates for exclusion. Subcentral levels of government are also left out of the spending rules in many European countries (European Commission, 2003). This section will examine these issues, starting with spending items and addressing subsectors in a second stage.

A common argument against the inclusion of the debt service in the spending target is that interest payments are strongly dependent on the evolution of interest rates, which in turn lies outside governments' control. If rates go up, compliance with an interest-inclusive ceiling will require cuts in other expenditure items; and if rates go down,

---

12 *Lei das Finanças Locais* (Law no. 42/98, on 6th August) and *Lei de Finanças das Regiões Autónomas* Law no. 13/98, on 24th February), respectively.

13 With some humour, and in the same vein of Barbosa (2002), who suggests that the 1998 Local Finance Law is "explosive" (see section 3.4 below), one might say that the Framework Law for the State Budget is "procyclical".

compliance can be achieved with minimal effort, even in the presence of spending overruns elsewhere (European Commission, 2003, p. 131)[14].

The issue of controllability, or lack of it, has a double dimension. Even if the government, when drafting next year's budget, could perfectly predict the path of interest rates, the problems mentioned in the paragraph above would still bite. But they can be made worse if there are forecast errors and intra-annual adjustment becomes necessary, complicating the budget's execution.

One way to address intra-annual controllability is to study the magnitude of forecast errors in interest payments. Since 1994, and under Council Regulation (EC) No. 3605/93 (later amended by Council Regulation (EC) No. 475/00), EU Member States have reported fiscal data twice a year (before 1 March and 1 September) to the European Commission[15]. The reporting includes final or provisional figures for the four previous years, and planned or forecast data for the current year. For a given variable *X*, let the *subscript* denote the year *X* refers to, and the *superscript* the year in which *X* is being reported (by 1 March). For example, letting IP represent interest payments in p.p. of GDP, IP_n^{n+1} denotes the estimate of year *n* released by 1 March of year *n+1*. One can then define the following variables:

- X^F is the current year forecast (X_n^n);
- X^{E1} is a forecast error defined as $X_n^{n+1} - X_n^n$;
- X^{E2} is a revised forecast error defined as $X_n^{n+2} - X_n^n$.

Naturally, the difference between X^{E2} and X^{E1} is caused by data revisions. Table 9 reports these variables for gross fixed capital formation and interest payments. It can be seen that interest forecast errors have been generally moderate, and practically nil since 2000 (in stark contrast to GFCF, as developed below). Part of this predictive success may, however, be due to the fact that interest rates have been low and stable by historical standards. Overall, and bearing in mind that, even if predictable, interest rates remain exogenous, there is a case to keep the debt service outside the expenditure ceilings.

[insert Table 9 about here]

What about public investment? The defence of a 'preferential treatment' for this spending item is at the heart of the time-honoured golden rule of public finances, and a natural extension would be to exclude public GFCF from the spending target. Many arguments can be found both pro and against treating public investment separately (Balassone and Franco (2000) survey some of them in the context of the debate on the 'golden rule'). Defenders of separate consideration point to the heightened vulnerability of public investment in times of fiscal retrenchment and to issues of intergenerational equity. Arguments against range from the possibility of excessive capital accumulation and low-

14 Not unlike the deficit reduction process in the run-up to EMU (see section 2.1).

15 In the remainder of this study, such reporting will be designated (in a short and loose way) as the EDP reporting (EDP standing for Excessive Deficit Procedure).

return projects to the risk of creative accounting (due to some dim borders between current and capital expenditure); it is also stressed that important areas of 'productive' public spending (e.g. education, R&D) are not classified as public investment. Without being exhaustive, some of these arguments are examined below.

I first address the relative importance of public investment in Portugal, taking the EU average as a benchmark. It is well known that a significant proportion of Community Support Framework (CSF) spending goes into infrastructural investment belonging to the GFCF of general government. Table 10 includes estimates of these CSF amounts, either considering only EU financing – column (2) of the Table – or including co-financing by the Portuguese government as well – column (3). One can therefore compute public investment exclusive of CSF support, and this is done in columns (4) and (5).

[insert Table 10 about here]

The importance of CSF-supported public GFCF varies from year to year, as rates of execution fluctuate, and hence it is more informative to consider the eight-year average reported at the bottom of the table. One sees that EU funds account for almost two-thirds of the difference between Portugal and the EU15 in the relative importance of public investment (1.1 in 1.7 p.p. of GDP, from columns (7) and (8)). And if additionally is interpreted in a strict way, comprising national co-financing, that difference shrinks to a mere 0.2 p.p. of GDP[16]. Hence, the larger weight of public GFCF in Portugal is essentially a reflection of CSF support, rather than of an autonomous investment effort by the Portuguese government surpassing those of its European counterparts[17].

One could be tempted to think of GFCF as a highly controllable budget item – not cyclically-sensitive, not controlled by outside actors, not a counterpart of entitlements. Looking at Table 9, however, yields a different picture: public investment is subject to significant forecast errors, sometimes much bigger than those of interest payments. One possible explanation lies in deliberately biased forecasts: the errors are indeed one-sided. But one can think of two other, both of which are arguments in favour of excluding public GFCF from the spending aggregate to be targeted.

The first has to do with CSF support. Fluctuations in rates of execution explain some of the forecast errors – for instance, part of the huge deviation in 2000 (Banco de Portugal, 2001, p. 120). Therefore, the reliance on external financing for some projects introduces a problem of controllability (at the yearly frequency), which is likely to be felt for some more years. The second explanation relates to the vulnerability of public investment, an item more easily reduced than wages or transfers. Part of this vulnerability translates into intra-annual cuts, which have been common in Portugal (e.g. OECD, 2003,

16 The comparison could be perfected by subtracting CSF amounts from the EU15 GFCF as well, but the latter would be only slightly affected, since the main recipients of EU funds are small economies (except Spain). On the other hand, the late nineties were a time of historically low public investment in the EU15 (European Commission, 2003, p. 77).

17 Aspects related to the "quality" or rate of return on public investment are beyond the scope of this paper. Still, it is worth mentioning that the European Commission (2003), using Granger causality tests, has recently reported evidence showing that in Portugal and Spain (unlike other EU countries) public investment has a crowding-in effect on private investment. Also beyond this study lies the issue of whether the actual level of public investment in Portugal and Europe is insufficient or not.

pp. 45-53; ECORDEP, 2001, p. 31-32) and Table 9 documents[18]. I therefore think that including GFCF in the spending ceiling would turn this item into the adjustment variable, making room for overruns in current expenditure.

A final item whose exclusion may be considered is unemployment-related transfers. In Portugal, they comprise the *subsídio de desemprego* (unemployment benefit, UB) and the *subsídio social de desemprego* (social unemployment benefit, SUB). The latter is less demanding in terms of eligibility, and correspondingly less generous in the protection provided.

From Table 11 one can see that unemployment benefits are no longer negligible in macroeconomic terms. Their combined share in GDP has been increasing at a fast pace, approaching the 1 per cent threshold. This is mainly due to an extension of coverage, as neither has unemployment displayed an upward trend, nor has the average subsidy become more generous[19]. Recent legislation[20] has made eligibility less strict, which may well result into further rises in the relative importance of these payments.

[insert Table 11 about here]

My assessment is that unemployment benefits should also be excluded from the expenditure ceiling, for two main reasons. First, because unemployment benefits act as an automatic stabiliser, and therefore contribute to countercyclical public expenditure. Second, because of the usual controllability argument, made more pressing by the growing weight of these benefits: in downturns, cuts would be needed elsewhere; in booms, room for laxity would be created. However, to reduce the scope for creative accounting, the exclusion should be precisely confined to the *subsídio de desemprego* and the *subsídio social de desemprego*, rather than defined in looser terms.

Turning now to issues of institutional coverage, it is well known that the yearly State Budget in Portugal does not set expenditure policy for the whole government sector: local and regional governments enjoy autonomy in drafting their own budgets, though subject to borrowing limits (section 3.1). Including the LG sector in the spending aggregate would therefore be politically sensitive. It would also be impractical and hard to monitor, due to the large number of local governments (Brunila and Kinnunen, 2002, p. 21) and to the paucity of timely statistical information on this subsector (see section 3.5).

To sum up, and in the light of the discussion above, the spending aggregate should be defined in terms of the outlays of the central government and of the social security, excluding interest payments, GFCF and unemployment benefits.

---

18 The table lends some support to the hypothesis that forecast errors are themselves procyclical, thereby contributing to the overall procyclicality of public GFCF (section 2): in the 1994 to 1999 period, the correlations of $GFCF^{E2}$ with the output gap and with output growth were 0.86 and 0.59, respectively (the most recent years have been excluded due to the special factors affecting 2000, partly reversed in 2001, when CSF III projects gathered pace – OECD, 2003, n.18).

19 From 1990 to 2002, the average monthly individual UB and SUB increased 75 and 58 per cent, respectively (author's computations, based on IGFSS and IIESS data), whereas nominal GDP rose 89 per cent.

20 *Decreto-Lei 84/2003*, on April 24.

3.3. Setting the Target

If an expenditure aggregate is to have a ceiling, how exactly should this be set? According to actual or to potential GDP? In nominal or in real terms? And for which time horizon? These are the questions to be addressed in this section.

To avoid procyclicality, it is essential that expenditure grows according to *potential*, rather than *actual*, GDP. Evolution according to the latter – in the simplest form, by keeping a constant share of GDP – implies, of course, the very procyclicality one wishes to escape: fast expenditure growth in booms, and restraint in slumps. The fact that potential growth should be taken as a reference does not imply, however, that expenditure should remain forever constant as a share of potential GDP. One may find desirable to change the relative size of public outlays, for reasons as diverse as the goal of reaching a budget position closer to balance, the need to make room for future pension liabilities, or a general reassessment of the role of the State in society. Such a change can be accommodated by setting a *constant, non-cyclical* differential between the growth rates of expenditure and of potential output. Taking into account the present risks of SGP violation, but also the current negative output gap, Portugal should probably set a small negative differential for the coming years – i.e., to make expenditure grow slightly slower than productive capacity.

Though relying on the concept of potential output, expenditure targets should not be set on the basis of recent estimates, or projections, of potential growth. As Mills and Quinet (2001) point out, potential output estimates are both unreliable (subject to large revisions) and often procyclical. Those authors recommend "setting a spending rule in relation to long-run potential GDP" (ibid., n. 2, p. 327).

A simple way to make this concept operational would be to take the average real growth of GDP over a long enough period. In the Netherlands, for instance, a 20-year average is used (Heeringa and Lindh, 2001). Table 12 presents 10-year averages for Portugal, suggesting that 2.5 per cent could be a prudent estimate of potential growth. An alternative, more sophisticated approach would be based on an aggregate production function and on an assessment of likely trends in factor inputs (e.g. immigration and labour force participation rates) and in productivity growth. Care would need to be taken, however, to avoid the pitfalls of the previous paragraph.

[insert Table 12 about here]

Expenditure targets should be set for the medium term, and not just for a single year. In the latter case, it is easier to circumvent targets by anticipating or postponing outlays (European Commission, 2003, p. 132). Another argument for multi-year targets lies in the greater predictability they can bring to the activities of individual government bodies, making it easier to implement objective-oriented management techniques and to assess performance (ECORDEP, 2001; OECD, 2003). Multi-year targets can be set for a government's term of office, as in the Netherlands, or in a rolling way, as in Sweden (Heeringa and Lindh, 2001): I return to this issue in section 3.5.

Multi-year spending targets can be defined in real or in nominal terms. The Netherlands uses the former possibility, Sweden the latter. The difference results from the possibility of inflation forecast errors (Brunila and Kinnunen, 2002): if inflation is higher than expected, a nominal target makes the budget's execution more difficult, by creating the need for intra-annual spending cuts; if, on the contrary, inflation is unexpectedly lower,

then extra spending becomes possible. Against this disadvantage, however, nominal targets offer the advantage of simplicity and create an incentive for wage moderation on the part of the government. Some authors also add that a nominal expenditure ceiling helps to stabilise the economy in the wake of demand-driven inflationary shocks (Mills and Quinet, 2001; Brunila and Kinnunen, 2002)[21].

Portuguese inflation, measured in Table 13 by the GDP deflator, has been among the highest in the euro area, as one might expect from the operation of the Balassa-Samuelson effect[22]. Higher average inflation tends to translate into higher variability, and thus less predictability, tilting the balance in favour of real targets. These can be set for the whole planning horizon, and then translated into nominal terms at the time of the preparation of each year's budget by applying the forecast for that year's price level (as in France, the Netherlands and other countries that use real targets – see Hallerberg *et al.*, 2001, p. 9).

[insert Table 13 about here]

One can put forward an additional argument in favour of multi-year real targets: they provide an anchor that limits the risks associated with the use of provisional data for the recent past, or estimates for the current year, as "base values". As section 3.5 documents, preliminary data are subject to significant revisions, Portugal being one of the countries hardest hit by this problem. If the base is revised, targets set on a year-by-year basis may result in undesired rates of growth. In a scheme as the one described in the previous paragraph, the risks are smaller: the GDP deflator is still subject to revisions, but the real growth component is set in advance.

The above discussion suggests setting real targets for the spending aggregate defined in section 3.2 over a horizon of 3 to 4 years. An annual real growth rate of 2 per cent (2.5 per cent of potential growth minus 0.5 per cent for consolidation purposes) could be a roughly adequate magnitude[23].

3.4. The Budget as a Whole

The previous sections have defended the adoption of multi-year spending targets, and analysed their implementation in some detail. Multiyear fiscal programming, however, should also include the remaining components of the budget and the projected path for

---

21 In my view, some of the above pros and cons should be kept in perspective. If unexpectedly high inflation results from a negative supply shock (which depresses output), then a nominal target will further destabilize the economy by reducing real spending. On the other hand, one should not forget that the government can set the "prices" of significant parts of spending (e.g. wages and pensions), at least in the short run, though it is still exposed to surprise inflation in others (e.g. intermediate consumption or health costs).

22 Due to the very short sample, country differences also reflect cyclical factors.

23 With an inflation of 3 per cent, the spending aggregate would then grow at 5 per cent in nominal terms. Seen in this light, the 4 per cent ceiling for the nominal growth of current primary expenditure proposed in ECORDEP (2001) appears too stringent: with nominal GDP growing at an average rate of 5.5 per cent (2.5+3), current primary expenditure as a share of GDP would decline at a fast pace (around 5 p.p. in 10 years). One can infer from Table 1 that in the 90s only Ireland and the Netherlands recorded comparable reductions (7.2 and 6.4 p.p., respectively), the former helped by very fast economic growth, and the latter starting from a significantly higher share. Considering, further, the pressures on spending due to ageing, my assessment is that such a 4 per cent ceiling would require very restrictive measures in many areas.

public debt. Many justifications can be offered: (i) deficits play a central role in the SGP; (ii) debt is what ultimately matters for fiscal solvency; (iii) outlays excluded from the spending target should also be monitored[24]; and so on. This section therefore addresses those components – revenues and the associated macroeconomic forecast, spending items not included in the targeted aggregate, the budget balance and public debt.

A sound budget requires a realistic macroeconomic scenario, from which forecasts for revenue and for unemployment-related transfers should follow. In Portugal, growth forecasts tend to be too rosy, as the first row of Table 14 illustrates for recent years: GDP growth has been overestimated by an average 0.7 p.p. in the 1999-2002 period, a bias only surpassed, among euro area countries, by Italy and Germany (0.8 p.p.).

[insert Table 14 about here]

Biased growth forecasts may translate into errors when predicting the budget balance. Table 14 shows that Portuguese budget forecasts have been the most biased in the euro area, with deficits being on average 0.6 p.p. of GDP worse than planned[25]. Italy and Germany come second, with a 0.4 average deterioration. The table also presents average budget forecast errors *adjusted* for the error made as regards GDP growth: if growth is lower (higher) than anticipated, the balance should worsen (improve) due to the operation of automatic stabilizers[26]. When this adjustment is made, the average errors of countries like France, Italy or Germany practically vanish, but for Portugal one still obtains an average unexpected deficit increase of 0.4 p.p. of GDP. The upshot is that, due to spending overruns and/or unrealistic revenue projections, Portuguese budget forecasts in recent years have been substantially flawed, even *beyond* what can be attributed to excessive growth optimism[27].

The implications of the analysis above are that the proposed multi-year fiscal programming should be based on a prudent macroeconomic scenario *and* that revenue forecasts should be strictly in line with such scenario, excluding "efficiency gains" due to reduced tax evasion or Laffer curve effects[28]. The quality of the projections should be assessed by an independent body, as argued in section 3.6.

---

24 As Mills and Quinet (2001, pp. 326/327) contend, reconciling the goals for debt and for the tax burden requires considering the whole expenditure.

25 It should be borne in mind that outcomes are measured on the basis of preliminary, rather than definitive, values. This implies that most of the revisions of the Portuguese deficit figure for 2001 are not included in the Table (the outcome considered is a 2.2 per cent deficit).

26 In this adjustment I have used the European Commission's estimates of the budgetary sensitivities to growth (impact on deficits, in p.p. of GDP, when GDP changes 1 percent, as reported in European Commission, 2002, p. 48). For each year, the adjusted budget forecast error is computed as the actual budget forecast error minus the product of the budgetary sensitivity by the GDP growth forecast error.

27 The period 1999-2002 witnessed a marked unanticipated slowdown of the world economy, Portugal being more affected than the European average. To some extent, it can therefore be argued that the choice of that period (for reasons of data availability) biases comparisons against Portugal. However, this bias should largely disappear when the budget balance forecast error is adjusted for the error in GDP growth, as done in the table. Further, Portuguese official growth forecasts tend to err on the side of optimism even in better times: in the period 1990-1999 real GDP growth was overestimated by an average 0.5 p.p. (figure computed on the basis of differences between "first available estimates" for year *n* taken from the budget for *n+1* and forecasts for year *n* taken from the macroeconomic scenario of the budget for year *n*).

28 The dangers of over-optimistic revenue forecasts are considered in ECORDEP (2001). It goes without saying that strong attempts to curb tax evasion should nonetheless be pursued.

What about the outlays that section 3.2 left out of the spending target? A projection for the unemployment rate should be included in the macroeconomic scenario, and be used to forecast unemployment-related expenditure, together with an assessment of other determinants, such as eligibility rates (recall section 3.2). Interest payments will depend, among other factors, on the debt stock and on ECB interest rates: it should be borne in mind that the latter are currently at historically low levels, and will therefore tend to rise. As for the relative weight of public GFCF, Portugal should maintain a positive differential vis-à-vis the EU average, ensuring full and efficient use of CSF support[29]. Again, it is to be remembered that one should not think of a spending item as a certain percentage of actual GDP, since this implies procyclicality.

As regards local government, annual borrowing limits are currently imposed (section 3.1). Relative to the Local Finance Law and the Finance Law of Autonomous Regions, those limits reinforce discipline[30]; but they are conducive to procyclical spending, as deficit limits generally are. The reason is that most local government revenues are linked to taxes, and thus follow the cycle. This is the case for fiscal revenues collected in the autonomous regions (which are given back to regional administrations, after some generous adjustments) and for municipal taxes. Transfers from the central government to municipalities in year *n* are indexed to the main nation-wide direct and indirect taxes of year *n-2* (see Bronchi, 2003, p. 37), and hence may turn out pro or anticyclical, depending on the characteristics of each particular cycle.

Balanced budget rules are common for the subcentral levels of government in many countries (most prominently the USA, but also European countries like Sweden), and their use could be considered in Portugal as well: they are, after all, a particular case of borrowing limits. More importantly, however, ways should be devised to make local government revenues cyclically neutral, as suggested in Lindh and Ohlsson (2002, p. 706) or in European Commission (2003, pp. 152-155). In any case, for the purposes of this section, a forecast of local government expenditure follows from the borrowing limit and from a forecast, or knowledge, of local revenues.

From estimated revenues and spending a certain budget balance will follow. If this is projected to exceed 3 per cent of GDP in a given year, some of the previous fiscal plans will have to be redesigned: though such a deficit could be justified in economic terms, in the context of a cyclical downturn, Portugal should not risk violating again the 3 per cent threshold. In my view, the adjustment should come from the revenue side, i.e., from higher taxes. First, because the alternative would be either to upset the spending target one wishes to implement, or to sacrifice public investment. Second, because the short run fiscal multipliers tend to be smaller (in absolute terms) for taxes than for public consumption (Lindh and Ohlsson, 2002, p. 676), and hence the procyclical impact on GDP would tend to be less serious.

Finally, deficits feed into debt, whose evolution relative to GDP ultimately determines public sector solvency. The debt-to-GDP ratio, however, depends on a host of factors. Table 15 presents the well-known decomposition of debt ratio changes into the contributions of primary balances (column 1), interest payments (column 4), nominal GDP

29 Each of these spending items would of course justify a more detailed treatment, which is beyond the scope of this paper.

30 For municipalities, both the letter of the Local Finance Law and its enforcement have been rather permissive – see Barbosa (2002) and Bronchi (2003, pp. 20-22).

growth (column 5) and stock-flow adjustments (columns 2 and 3)[31]. These adjustments correspond to changes in debt unaccounted for by the budget balance, and are further decomposed into privatization revenues used for debt repayment (column 2) and other adjustments (column 3).

[insert Table 15 about here]

Though the table would lend itself to several comments, for the purposes of this section the focus is on column 3. One sees that, more often than not, stock-flow adjustments other than privatization revenues have made debt increase[32]. The report by ECORDEP (2001, p. 39) makes a similar point, presenting as the main causes the assumption of liabilities of state-owned companies and the settlement of national health service debts. The magnitudes involved are sizeable: an average of 0.5 p.p. of GDP in the 1991-2002 period. Without non-privatization adjustments, debt at the end of 2002 would be only 51.8 percent of GDP[33].

There are, therefore, reasons for concern that these adjustments may reflect, to some extent, hidden deficits – a concern expressed, among others, by the European Commission (2003, p. 59). The implication is that in a multi-year fiscal framework the forecast for the evolution of debt should be carefully documented, with explicit consideration of stock-flow adjustments. These could be required to average a value close to zero over the planning horizon (privatizations excluded), and deviations treated along the lines suggested in the next section.

3.5. Annual Budgets, Flexibility and Adjustments

The previous sections have set out a multi-year fiscal framework, but made little or no reference to the annual budget preparation, or to corrective measures if deviations from planned trajectories emerge. The present section attempts to fill this gap.

At the time of preparing the budget for a given year, the macroeconomic forecasts contained in the multi-year planning framework would be reviewed and updated. Revised projections for unemployment and interest rates might ensue, and hence for the associated outlays as well. Revenue projections would of course also be reexamined. And the revised forecast for the GDP deflator would be multiplied by the preset real spending target, yielding a nominal expenditure ceiling.

This ceiling would not be adjusted in the wake of changes in the budgetary or macroeconomic outlook, but could respond to *some* past debt increases (relative to the multi-year path) – those due to spending overruns or stock-flow adjustments. The ensuing (projected) extra interest payments should be deducted from the spending target, thus

---

31 Columns 1 and 4 add to the budget balance. Columns 4 and 5 add to the so-called snowball effect, which is negative if nominal economic growth exceeds the implicit nominal interest rate on public debt.

32 According to the AMECO database, this trend was even more marked in the eighties. The data from that decade, however, is not strictly comparable to that presented in the table.

33 Everything else constant, i.e., with unchanged primary balances, privatization proceeds, interest rates and growth.

avoiding that such past increases are transmitted to future deficits[34]. It is suggested that this mechanism is one-sided: any debt decreases relative to the multi-year path, and concomitant interest savings, should not give rise to further spending.

As for revenues, a mechanism resembling that adopted in the Netherlands (see Heeringa and Lindh, 2001) could be implemented. The mechanism consists in comparing the updated nominal revenue projections with those following from the multi-year programming (in turn, the product of the original real forecasts - the so-called real reference levels - by the revised projection for the GDP deflator). This yields a forward-looking windfall or setback, a proportion of which gives rise to counteracting tax changes[35], while the rest is absorbed by the projected budget balance (thus reflecting automatic stabilization). One should note that those tax changes are actually procyclical: hence, it is my view that they should only be used in the event of a setback that threatens to make the deficit exceed the 3 per cent threshold.

Intra-annual adjustment mechanisms (i.e., those taking place during the execution of the budget) are also important. In Portugal they have mainly consisted in *ad-hoc* freezes in spending allocations, sometimes reversed in case of revenue windfalls (OECD, 2001, p. 50). Though useful in the absence of better alternatives, they represent a crude form of adjustment, introducing instability in the management of individual government services, and, to the extent that final outlays are made dependent on revenue collection, contributing to the procyclicality of expenditure.

It seems advisable to hold a general budget debate in Spring (as suggested in ECORDEP (2001, pp. 46-50) and already done in some other countries) and use the occasion, among other purposes, to make adjustments in spending allocations. A first source of adjustment may come from revisions to the projected GDP deflator for the current year (as in the Netherlands – Hallerberg *et al.*, 2001, p. 9), leading to across-the-board corresponding changes in allocations. A second correction concerns the feedback on past debt increases discussed above, using updated information. Finally, there should be corrections at the level of individual government services, either in the wake of past overruns or as a channel for inter-temporal reallocations.

At an aggregate level, it has been suggested above that the "punishment" for overruns be confined to the ensuing increase in interest payments. This is because annual ceilings are anchored in the multi-year programming, and therefore do not drift upwards due to past misbehaviour[36]. But at the level of individual services, overruns should be punished much more severely, both in terms of reduced future allocations and (possibly) in terms of penalties for managers. Savings in those allocations could be redistributed in some way among ceiling-compliant units, thus eliminating aggregate effects.

One should also allow for voluntary inter-temporal reallocations. Services could be allowed to transfer between years up to a certain percentage of their allocation, as it is the case in many European countries (Hallerberg et al, 2001, pp. 76/77). Aggregate effects

---

34 Consider a spending overrun of a given nominal magnitude X, which makes public debt increase by a similar amount. The interest on X will accrue to the total debt service. However, if such interest is subtracted from the expenditure ceiling, then total (i.e., interest inclusive) spending will not be affected by X. With time, nominal GDP growth will make the debt ratio converge to the baseline path (i.e. the path in the absence of the overrun X).

35 This proportion is different for windfalls and setbacks, and also depends on the projected deficit ratio.

36 They should not be revised downwards: two wrongs don't make one right!

could be limited, or eliminated, by only allowing those reallocations which net out in yearly terms.

Section 3.3 mentioned two possibilities for multi-year programming: rolling targets, or term-of-office targets. This paper endorses the latter. While rolling targets may ensure a higher degree of continuity of the framework, they also make it more complex: if, as suggested above, the multi-annual programming serves as an anchor to annual budgets, then it seems advisable that it is not revised itself each year. Term-of-office targets also make it simpler to assess the performance of each government, while fully respecting its democratic right to change policy priorities: namely, a newly-elected executive should be able to set a new steady real rate of growth for the spending aggregate, according to its own vision of what the role and size of government should be. It would be nonetheless desirable that the main political parties agree to set some bounds to the choice of such rate.

A final issue, which underlies many others, regards the reliability and timeliness of statistical information on budgetary outcomes. There is a widespread perception that Portugal fares poorly on this count. Table 16 presents one indicator in this area: the magnitude of data revisions to budget balances as a percentage of GDP. The average absolute revision for each EU country, given in the last row of the table, shows Portugal in an uncomfortable position, second only to Luxembourg[37].

[insert Table 16 about here]

This problem has two main dimensions. First, there is a lack of timely information on budgetary outcomes for some of the subsectors of general government, the problem being most acute in the case of local government. Recent Portuguese budgetary laws (see section 3.1) are well aware of this limitation. They reinforce the reporting obligations of non-State subsectors (articles 63 to 65 and 88 of the 2001 Framework Law for the State Budget, as amended by the 2002 Budgetary Stability Law), and introduce the possibility (though in a discretionary, non-automatic way) that transfers from the State are suspended if the required information is not provided (article 89). It is still too early to assess the effectiveness of these measures.

Second, the government may be tempted to do "creative accounting" so as to present budgetary data in a politically more convenient way. Some mechanisms to restrain this temptation are already in place. Most prominently, the Eurostat publicly announces whether or not it certifies the figures for past years contained in the EDP reporting of each country. Though valuable, this form of auditing has nonetheless displayed limitations, not being able to avoid very significant data revisions[38]. Further, national projections for the current year are simply not published by the Commission services, thus being exempt from the previous auditing mechanism. Hence, there is a case for reinforcing independent control over official estimates and projections – as the next section suggests.

---

37 The major slippage of 2001 heavily contributes to this result, but does not fully explain it.

38 To give a few examples, the reported figures for the 2001 Italian deficit were 1.4, 2.2 and 2.6 per cent of GDP in March 2002, August 2002 and March 2003, respectively. All were successively certified by Eurostat, though with some explanation for the changes. Eurostat did not validate the 2001 Portuguese deficit as reported in March 2002 (2.2 per cent of GDP), but the value of 2.75 per cent presented in the Commission's Spring Forecasts, published one month later, in April, was still far off the later 4.1 per cent estimate (in the August 2002 reporting).

3.6. A Role for Independent Fiscal Authorities?

Some recent contributions have proposed delegating some of the fiscal policy decision powers to independent authorities – 'National Fiscal Councils' in the terminology of Eichengreen *et al.* (1999), 'Fiscal Policy Committees' (FPCs) according to Wyplosz (2002).

These proposals draw on the experience of delegating monetary policy to independent central banks, and claim to offer an improvement on the credibility *versus* flexibility trade-off associated to fiscal rules. Unlike these, an FPC will retain enough flexibility to adjust the fiscal stance to short term developments, while preserving a credible commitment to long-run debt sustainability.

According to both Eichengreen *et al.* (1999) and Wyplosz (2002), the mandate of an FPC consists precisely in ensuring debt sustainability[39]. This translates into the legal power of setting the annual budget balance (or the debt change limit) before the budget is prepared by the government and parliament; and, once this has happened, of checking the spending and revenue projections, retaining a veto right if those projections fail to comply with the imposed balance. The FPC would therefore have the ability to adjust the budget balance to the state of the economy, under the constraint that higher deficits at a certain point in time would need to be (have been) compensated by a more restrictive stance in future (past) years.

The authors of these proposals are at pains to stress that democratic decision-making and accountability would not be removed from fiscal policy. First, because only the macroeconomic dimension of the policy (as encapsulated in the balance figure) would be delegated to an FPC. The structural and redistributive dimensions, associated to the overall size and composition of revenues and outlays, would still be entrusted to elected politicians. Second, because the FPC would be instrument-independent, but not goal-independent; and its members, appointed for long and non-renewable terms of office, would be subject to *ex-ante* and *ex-post* parliamentary oversight.

A milder version of these proposals consists in having an independent body with advisory and consultation roles – a 'Court of Wise Persons' (CWP), as Wyplosz (2001) puts it. The CWP could resemble an FPC in its mandate, composition and independence, but without legal clout. Its analyses and recommendations would be made public, with an adequate degree of formality and publicity. Like an FPC, the CWP would have its own technical support staff. Institutions with some of the characteristics of a CWP already exist in a number of countries – e.g. the UK's Audit Commission or the Belgian High Council for Finances. The danger of what essentially is a "naming and shaming" approach is that indiscipline-prone governments will be able to turn disagreements with the wisemen committee into a minor embarrassment (Wyplosz, 2001, p. 28), carrying virtually no political costs.

This paper does not endorse the proposal of an FPC. First, it does not seem realistic that a country like Portugal, with no tradition whatsoever of independent bodies in the area of fiscal policy, immediately jumps to an extreme form of such bodies – an FPC. Second, and more fundamentally, because I think that some of the arguments behind the proposal are flawed, as argued below.

---

[39] A concept to be made operational at a constitutional stage (i.e., when the FPC is set up) or by the political power (e.g., a debt target for the length of the legislature) – see Wyplosz (2002).

The separation between the macroeconomic and the structural/redistributive roles of fiscal policy is highly debatable. A given deficit target, especially a tight one, will have an impact on the composition of the budget: some outlays are more vulnerable than others (e.g. recall the discussion about public investment in section 3.2), not all tax rates are equally hard to change, and so forth. Reciprocally, changes in the size and composition of budgets will have macroeconomic consequences even if the deficit does not change – an insight dating back to Haavelmo's balanced budget multiplier, and reasserted in recent studies (e.g. Lindh and Ohlsson, 2002).

If the FPC sets a deficit target for the following year, there is a danger that procyclical measures might ensue, as argued above in section 3.1. A solution put forward by FPC proponents is to allow for in-year target revisions, but these could entail credibility problems. Strong institutions are not created overnight, and hence one cannot discard either the danger of time inconsistency in the path for deficits – for instance, expanding now and leaving offsetting surpluses for future years[40].

While an FPC does not seem a good option, setting up a "wisemen committee", endowed with a well-qualified support staff[41], would be a welcome development. The details of its composition will not be dealt with here, though the usual requirements of independence and long terms of office would apply. To maximize the impact of the committee's opinions and recommendations, they would be made public and formally communicated to the government, the parliament and to international organizations (e.g. the European Commission, the IMF and the OECD).

What would such a committee do? I list below some possible tasks, which follow from the analysis of previous sections. The general principle is that the committee would have full responsibility for macroeconomic forecasts, and an auditing and advisory role in other budgetary matters. Hence, the committee would:

- prepare the macroeconomic scenario for the multi-year programme, including an assessment of the rate of long run economic growth;
- be consulted beforehand, and deliver an opinion on the rate of real growth of the spending aggregate chosen by the executive;
- be consulted beforehand, and deliver an opinion on the remaining aspects of the multi-year programming;
- prepare an updated macroeconomic scenario (i.e., forecast) for the preparation of each annual budget;
- be consulted beforehand, and deliver an opinion on the annual budget, with particular emphasis on the quality of revenue forecasts;
- prepare an updated estimate of the current year's GDP deflator before the Spring budgetary debate;

---

40 A similar kind of time inconsistency can be found in the successive Stability Programmes of some countries, including Portugal. A suggestive illustration can be found in OECD (2003, p. 55).

41 As regards technical support, one could introduce an element of competition between research centres, as suggested in recommendation no. 4 of ECORDEP (2001).

- deliver an opinion on the quality of the budgetary execution projections for the current year and of the provisional accounts for recent past years prepared by the government, including the EDP reporting[42];
- prepare an annual report on fiscal policy, assessing overall compliance with the expenditure rule framework; the report should also document the evolution of public debt, and set the projection of accrued interest payments to be deducted from the spending target.

It is worth mentioning that a recent study, the so-called Sapir Report (Sapir *et al.*, 2003), also proposes establishing independent national fiscal bodies, which are termed Fiscal Auditing Boards (FABs). There is some overlap between the proposed roles of these boards and the list of tasks above, especially as regards the mission of auditing the quality of fiscal data on the recent past and of current estimates and projections. Other tasks are somewhat different (the FABs would also assess the budgetary impact of alternative policy decisions and examine the implications of different economic assumptions), since the missions proposed in the present paper follow from a specific framework of expenditure targeting.

From the assignment of tasks suggested above, it follows that the "wisemen committee" would play the central role in the enforcement of the expenditure rule. The most widely quoted reference in this area is Inman (1996), who characterizes enforcement of balanced budget rules (BBR) in U.S. states along three dimensions. First, whether access to the enforcer in case of BBR violation is open or closed (open access meaning that any taxpayer may bring a case against the fiscal decision-maker). Second, whether the enforcer is independent (from the fiscal decision-maker) or partisan. Finally, whether penalties for infringement are small or large. According to Inman, a strong BBR has open access, an independent enforcer and large penalties, whereas a weak BBR has the opposite characteristics.

By Inman's criteria, the suggested "wisemen committee" would be an independent enforcer, but penalties may be regarded as small, since the committee's opinions, though widely publicized, are mostly non-binding. No open access is envisaged, either: the committee would not deliver opinions on demand from third parties. Therefore, enforcement would be mostly of a weak type. I argue, however, that adopting stronger forms of enforcement could prove a poisoned chalice. For instance, giving the committee the power to impose changes to the budget (e.g. spending cuts) would make it an easy scapegoat for unpopular measures; and opening access could well place the committee in the uncomfortable position of having to decide on a complaint brought against the government by an opposition party. In both cases the committee's standing in the eyes of the public would be undermined, casting doubts on its independence and ultimately destroying its ability to effectively "name and shame".

---

42 Currently the Ministry of Finance and the National Institute of Statistics (INE) divide responsibility for the figures contained in the EDP reporting: the Ministry reports years *n* (the current one), *n-1* and sometimes *n-2*, and INE reports years for which finalised accounts already exist (*n-3*, *n-4*, sometimes *n-2*). It should be clear that the "wisemen committee" opinion would focus only on the figures under the responsibility of the Ministry of Finance.

3.7. An Assessment of the Proposed Rule

This section assesses the proposed framework for fiscal policy in the light of a set of criteria first put forward by Kopits and Symansky (1998) and recently applied by Buti *et al.* (2003) to the SGP. According to these authors, fiscal rules should ideally be well-defined, transparent, simple, flexible, adequate to their final goal, enforceable, consistent and underpinned by public finance reforms. These eight attributes are analysed in turn, though their assessment entails some subjectivity and has a tentative character. For conciseness, the term "expenditure rule" will often be used to designate to whole fiscal policy framework suggested in the present paper.

The proposed expenditure rule is **well-defined**, both in terms of the variables to be targeted and as regards the institutional coverage (subsectors of the general government): this has been the object of section 3.2. Later sections have also clarified that targets are to be set in real terms for a government's term of office, and addressed their articulation with annual budgets.

The rule also fares reasonably well in terms of **transparency**, though on this count the "wisemen committee" plays an essential role. Variables are to be defined according to the ESA95, and items which do not map into a specific ESA95 category, such as unemployment benefits, are to be carefully circumscribed. To a large extent, the "wisemen committee" is able to prevent the use of biased forecasts, since it sets the underlying macroeconomic scenario and the estimate for the GDP deflator. It also assesses the quality of provisional figures on the budget's execution, though in this respect the problems of limited information (next to last paragraph of section 3.5) will remain a constraint in the near future. A careful monitoring of the evolution of public debt also contributes to transparency.

Admittedly, the suggested rule is far from **simple**. There are unavoidable trade-offs between simplicity and other criteria (especially flexibility and adequacy to final goal), and the proposed framework is tilted towards the latter. It follows that the "wisemen committee" will need to pay special attention to communicate its views to the "educated layman", and not just to the restricted circle of professional economists. Though far from simple, it should be noted that the rule contains no escape clauses, which would make it more complex and less transparent.

The absence of escape clauses for special circumstances does not prevent the rule from having a high degree of "built-in" **flexibility**. The spending target excludes items which are cyclically-sensitive or otherwise hard to control. Its definition in real rather than nominal terms also contributes to flexibility, as do the suggested intra-annual adjustment mechanisms.

Adequacy to final goals is one of the most important qualities of the expenditure rule. By summarising the two main flaws in Portuguese public expenditure, Section 2.3 implicitly set out the major goals of the rule: (i) to avoid procyclicality in public outlays and (ii) to enhance fiscal discipline, easing compliance to the SGP and ensuring public solvency. The definition of the spending target, as well as the suggested treatment of revenues, contribute to achieving both goals, by making room for the operation of automatic stabilisers, while restraining discretionary activism. A close monitoring of debt and the suggested feedback on past expenditure overruns (section 3.5) are also conducive to discipline and solvency.

Insufficient **enforceability** is a typical problem of fiscal rules, as Buti *et al.* (2003) acknowledge in the case of the SGP. The issue has been addressed in section 3.6, which attributes to a "wisemen committee" a central role in enforcement. Some additional remarks are made here. First, it is essential that the two largest Portuguese political parties agree on a common fiscal framework, to increase the chances that it outlives the government that initiates it. Such an agreement will probably require that each government retains some freedom over the steady rate of real expenditure growth to be chosen for its term of office, though it would be desirable to avoid huge variability from one government to the next. Second, many elements of the fiscal framework can and should be written into law, but some others cannot and/or should not: the value of the rate of real expenditure growth is a case in point. Hence some scope for judgement and political decisions will remain, which in turn reinforces the value of an agreement between the main parties. Finally, as argued in sections 3.5 and 3.6, the law should set penalties for mismanagement at the level of individual government units, but the final responsibility for fiscal decision-making should remain entrusted to the elected government, rather than to an independent committee.

Finally, the proposed expenditure rule fares well in terms of **consistency** and **linkage to public finance reforms**. There are no apparent internal contradictions, and a virtuous circle may emerge between the rule and some structural reforms in the public administration. Namely, the rule requires multi-year programming, but at the same time it creates favourable conditions for government services to plan their activities over the medium run, by ensuring a more regular stream of financial allocations. The rule should also stimulate an improvement in the quality and timeliness of fiscal statistics. Needless to say, the proposed framework is not an all-encompassing answer to public finance problems, and some important areas, such as tax reform, are left out of the analysis.

To sum up, the suggested rule seems to satisfy most criteria, though with some potential weaknesses due to its complexity and problems of enforcement. These are areas where the "wisemen committee" will have a crucial role, and hence one may conclude that the creation of such an independent body will greatly enhance the chances of success of the proposed fiscal framework. However, one should not lose sight of the fact that the above assessment exercise is entirely prospective, and that even for other countries the empirical evidence on expenditure rules is still very thin (section 3.1). If such a rule comes to be implemented in Portugal, practical experience may well show that some of this paper's proposals are not the best solution. Again, an independent committee will be in a privileged position to critically assess the rule itself, and propose adjustments.

4. Concluding Remarks

This section summarizes the main conclusions of the study, and adds a final note on fiscal stabilization policy.

The Portuguese deficit reduction of the nineties differed substantially from those of other European countries: it was accompanied by steep increases in current primary outlays, which prevented any fiscal consolidation as conventionally defined. Capital expenditure (in p.p. of GDP) increased only marginally. Financing came from higher taxes and lower interest payments – an item for which savings (in p.p. of GDP) exceeded even Belgium's and Italy's. Among the general government subsectors, the Autonomous

Services and Funds recorded the highest growth rate of current primary spending, followed by Local Government.

Public expenditure in Portugal has also behaved in a procyclical way. On the basis of a sample starting in 1977, personnel costs, intermediate consumption and gross fixed capital formation have contributed the most to a degree of procyclicality which is high by European standards. Public revenue, on the contrary, has reacted to the cycle in a stabilizing way, even if one removes the influence of automatic stabilizers. The cyclically-adjusted budget balance has been characterized by a relatively neutral stance, reflecting roughly offsetting influences from revenue and spending.

There are, therefore, important benefits to be derived from a better control of public expenditure, both in terms of policy discipline and in terms of an increased stabilization capacity. Drawing on the experience of other countries, the paper has studied how an expenditure rule could be designed in Portugal. As clarified in the introduction, the focus has been on the macroeconomic framework, and aspects related to institutional implications, as well as implementation details at a decentralized level, have received much less attention. It also goes without saying that the true test of the quality of the suggested rule will be its performance in practice, if it ever comes to be implemented.

The central element of the strategy consists in setting multi-year real spending targets according to a stable growth rate guided by (though not necessarily equal to) the rate of long run economic growth. When drafting an annual budget, the respective real target is converted to nominal terms using a projection of the GDP deflator. The spending target excludes, for different reasons, interest payments, unemployment-related transfers, public investment and the outlays of local government. Taking the SGP as given, revenues are to be adjusted whenever there is the prospect of a deficit over 3 per cent of GDP. Expenditure adjustments should be much more limited, and only intended to prevent that past overruns and unexpected off-budget debt increases have an impact on future deficits.

The suggested framework is relatively demanding in terms of macroeconomic forecasts, an area where the track record of Portuguese fiscal authorities is poor. The paper hence proposes that the responsibility for producing those forecasts is given to an independent agency. This body would also have to be consulted at various stages of the budget process, delivering (non-binding) opinions and recommendations on issues such as the government's choice of the multi-year real growth rate for the spending aggregate, the quality of the revenue projections (conditional on the macroeconomic forecasts supplied by the agency) or the timeliness and accuracy of reports on the current and last years' budget execution.

The attentive reader may at this stage raise the following objection: the proposed fiscal framework amounts to a broadly neutral structural stance, especially on the expenditure side; this may represent an improvement relative to the procyclicality of the past, but still falls short of a more active use of fiscal policy as a countercyclical tool, as considered in Seidman (2001) or Lindh and Ohlsson (2002). True, and I offer two justifications. First, timely countercyclical actions are generally on the revenue side, and therefore the suggested expenditure rule framework could easily be adapted to accommodate this possibility. Second, there is a case for gradualism. Discretionary interventions are harder to get right, and Portuguese fiscal policy has a fragile "infrastructure" in many respects (e.g. forecasts, spending control, timeliness and accuracy of information). It may be a better strategy to start by developing the ability to follow a simpler, rules-based regime.

Annex 1

Table A.1

Correspondence between ESA95 and AMECO codes

Budget items	**ESA95 codes**	**AMECO codes**
1. Taxes on production and imports (indirect taxes)	D2r	UTVG
2. Current taxes on income and wealth (direct taxes)	D5r	UTYG
3. Social contributions	D61r	UTSG
4. Other current revenue	(1)	(6)
5. Total current revenue (1+2+3+4)		**URTG-UKTTG**
6. Intermediate consumption	P2	not available
7. Compensation of employees	D1p	UWCG
8. Social transfers	(2)	not available
Of which: Social transfers other than in kind	D62p	UYTGH
9. Interest	D41/D41(EDP)	UYIG (5)
10. Subsidies	D3p	UYVG
11. Other current expenditure	(3)	(7)
12. Total current expenditure (6+7+8+9+10+11)		**UUTG-UIGG0-UKOG**
13. Gross saving (5-12)	**B8g**	**USGG**
14. Capital transfers received (capital revenue)	D9r	UKTTG
15. Total revenue (5+14)		**URTG**
16. Gross fixed capital formation	P51	UIGG0
17. Other capital expenditure	(4)	UKOG
18. Total expenditure (12+16+17)		**UUTG (5)**
19. Net lending(+)/net borrowing(-) (budget balance) (15-18)	**B9/B9(EDP)**	**UBLG (5)**

(1): P11+P12+P131+D7r+D4r+D39r

(2): D62p+D6311p+D63121p+D63131p

(3): D29p+D42p+D43p+D44p+D45p+D5p+D7p

(4): P52+P53+K2+D9p

(5): Until 1998, D41; from 1999, D41 (EDP); variables UUTG and UBLG are similarly affected

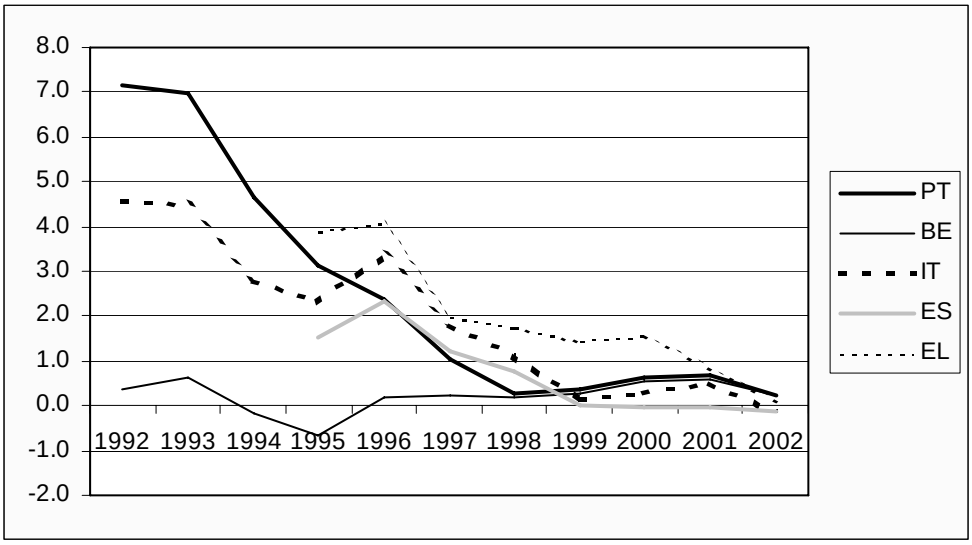
(6): obtained by difference

(7): obtained by difference; it includes social transfers in kind

Annex 2

Chart A.1

Implicit interest rate on public debt, spread *vis-à-vis* Germany (p.p.)

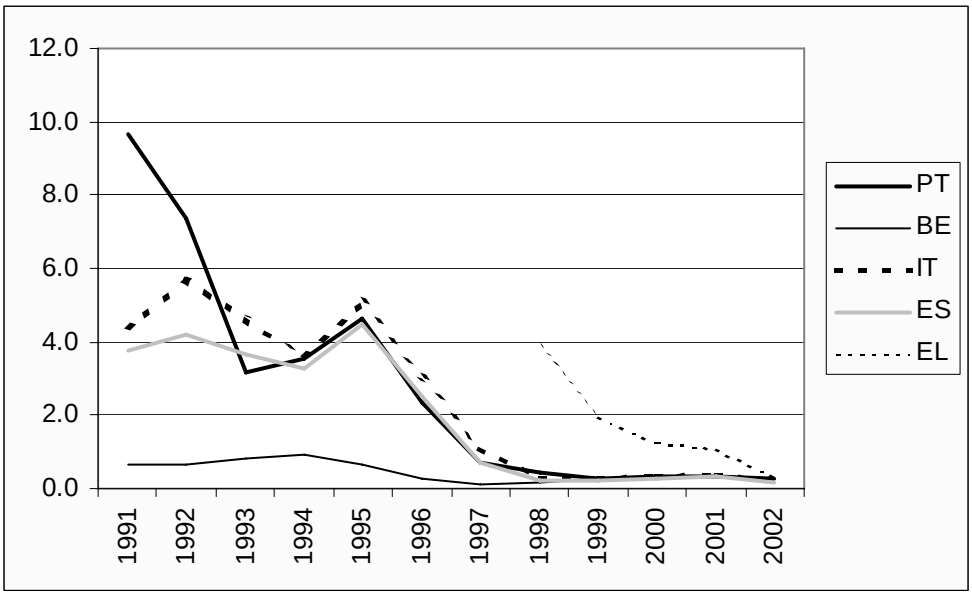


Source: AMECO.

Notes: the implicit interest rate on public debt for year *n* is defined as interest payments in year *n* divided by the stock of public debt at the end of year *n-1*.

Chart A.2

Nominal long-term interest rates, spread *vis-à-vis* Germany (p.p.)



Source: AMECO.

Notes: Nominal long-term interest rates are presented as annual averages and correspond, from 1995 onwards, to the central government 10-year benchmark bond.

Annex 3

The method followed by the European Commission to compute cyclically-adjusted budget balances is presented in European Commission (2002). Potential output is estimated on the basis of a Cobb-Douglas aggregate production function. The budgetary sensitivity to the output gap rests on the tax and expenditure elasticities computed by the OECD (Van den Noord, 2000). The Commission services use a weighted average of the four OECD tax elasticities (personal income taxes, corporate taxes, social security contributions and indirect taxes), and report cyclically-adjusted figures for overall revenues, rather than for individual items. As for expenditure, aggregate figures are likewise reported, though obtained under the assumption that only unemployment-related spending responds to the cycle in an automatic way.

To derive cyclically-adjusted values for individual budget categories, I have started from the AMECO values for total revenue and total expenditure (codes URTGAP and UUTGAP), and for the production function-based output gap (code AVGDGP). The difference between effective and adjusted expenditure has been attributed to social transfers other than in kind. Adjusted values for indirect taxes and social contributions have been computed using the Commission's output gap and the respective OECD elasticities. Formally,

$$T^{CA} = T.\left(\frac{Y^*}{Y}\right)^{\alpha}, \qquad \text{(A.1)}$$

where T is effective revenue, T^{CA} is cyclically-adjusted revenue, Y and Y^* denote actual and potential GDP, and α is the elasticity. The remainder of the difference between effective and adjusted revenue has been imputed to direct taxes.

Annex 4

As a robustness check, equation (1) was rerun using the cyclical components of the fiscal item *X* and of output *Y*, both obtained by the Hodrick-Prescott (HP) filter, instead of their growth rates. Cyclical components were defined as the percentage difference between actual and filtered figures; hence, the regressor becomes the widely used HP output gap. This approach also deals with the possible objection that in equation (1), as set out in the main text, country comparisons may be blurred by the fact that the average growth of fiscal items in the sample period differed across nations – though this should be captured by parameter α, not β.

Table A.2 presents the results, which are broadly supportive of Table 5. Portuguese coefficients tend to be close in both tables, both numerically and in terms of statistical significance. The comparisons made in the main text between Portugal and the other countries remain broadly valid, though contrasts are now somewhat milder: the Portuguese coefficient for total revenues becomes the second highest (after Denmark), and the procyclical behaviour of total primary expenditure in Portugal is now matched by Italy.

[insert Table A.2 about here]

Table A.3, also based on HP filtering, is the counterpart of Table 6 in the main text, and conveys essentially the same message. The main differences between growth rates and HP cyclical components are, again, those mentioned in the previous paragraph (Denmark with the most stabilising revenue behaviour and Italy matching Portugal in procyclical total primary spending).

[insert Table A.3 about here]

For completeness, Table A.4 reports HP-based regressions for the expenditure of the general government subsectors, thus checking the robustness of Table 8. There is some variation in the coefficient estimates, but it mainly affects relatively small items (e.g. capital expenditure of central government), rather than the larger aggregates. Nonetheless, the previous conclusions are generally confirmed: higher (and statistically significant) procyclicality in the ASF and LG sectors, mainly due to the behaviour of intermediate consumption, wages, non-social current transfers and (for LG) FBCF.

[insert Table A.4 about here]

Annex 5

The gap between the forecast of a variable and its reported realisation results not only from genuine forecast errors but also from methodological/definitional changes. This is a well-known problem in studies on forecast assessment (e.g. Artis and Marcellino, 2001), widely acknowledged but (due to data constraints) seldom addressed. Without offering a complete solution, I have gone some way in purging forecast errors from statistical changes.

First, no comparisons were made between ESA79 forecasts and ESA95 realisations. The Portuguese Stability and Growth Programme of February 2000 presents public sector accounts according to both ESA79 and ESA95. Because this Programme is almost exactly contemporaneous to the March 2000 EDP reporting (the first on a ESA95 basis), ESA79 figures were used when appropriate[43].

Second, a similar adjustment was made regarding a methodological change in interest payments introduced in the March 1997 reporting. At the time, the *Ministério das Finanças* prepared a note providing values according to the old and new definitions, and the former were used to assess previous forecasts.

[43] For instance, IP^{E1} for 1999 was computed using the ESA79 figure for 1999 interest payments from the February 2000 Programme.

Table 1

Budget items in 2000 and their changes in the nineties (p.p. of GDP)

	Belgium		Denmark		Germany		Greece		Spain		France		Ireland		Italy	
	2000	Δ 00/90	2000	Δ 00/90	2000	Δ 00/91	2000	Δ 00/90	2000	Δ 00/90	2000	Δ 00/90	2000	Δ 00/90	2000	Δ 00/90
Current revenue	49.0	2.8	56.7	1.4	46.6	2.8	44.3	n.a.	38.4	n.a.	50.9	2.3	35.1	-3.9	45.8	3.4
Indirect taxes	13.1	1.3	17.2	0.5	12.0	1.0	15.2	n.a.	11.7	n.a.	15.5	0.6	13.2	-1.2	15.0	4.3
Direct taxes	17.3	1.7	29.6	1.3	12.5	1.2	10.8	n.a.	10.5	n.a.	12.3	4.0	13.6	0.4	14.7	0.5
Social contributions	16.1	0.0	3.3	1.0	18.6	1.4	14.0	n.a.	13.3	n.a.	18.2	-2.3	5.6	-1.5	12.7	-1.7
Other current revenue	2.5	-0.2	6.7	-1.3	3.4	-0.8	4.4	n.a.	2.9	n.a.	4.9	-0.1	2.7	-1.6	3.3	0.3
Capital revenue	0.5	0.2	0.5	-0.1	0.4	0.1	2.7	n.a.	0.6	n.a.	0.4	0.4	1.3	-0.1	0.4	0.2
Total revenue	49.5	3.0	57.3	1.3	47.0	2.8	47.0	13.5	39.0	0.2	51.3	2.7	36.4	-4.1	46.2	3.6
Current expenditure	46.3	-4.6	52.5	-2.6	45.2	2.7	41.7	n.a.	35.2	n.a.	48.5	2.4	27.2	-13.0	44.3	-4.6
Intermediate consumption	3.1	0.2	7.9	0.7	3.9	-0.5	5.1	n.a.	4.1	n.a.	5.3	-0.2	5.2	-0.5	5.0	0.0
Compensation of employees	11.4	0.3	17.0	-0.8	8.2	-0.8	11.7	n.a.	10.5	n.a.	13.5	1.0	7.8	-2.6	10.6	-2.0
Social transfers other than in kind	15.3	-0.9	17.3	-0.6	18.8	3.1	16.6	n.a.	12.3	n.a.	17.8	1.0	8.0	-3.8	16.8	1.3
Interest	6.8	-5.2	4.2	-3.1	3.4	0.5	7.0	n.a.	3.3	n.a.	3.1	0.2	2.0	-5.8	6.5	-4.1
Subsidies	1.5	-0.2	2.2	-0.1	1.7	-0.4	0.2	n.a.	1.2	n.a.	1.2	-0.6	0.8	-0.4	1.2	-0.7
Other current expenditure	8.3	1.2	3.9	1.1	9.3	0.9	1.1	n.a.	3.8	n.a.	7.5	1.0	3.3	0.1	4.3	0.8
Capital expenditure	3.1	0.6	2.2	0.4	3.2	-1.5	7.2	n.a.	4.7	n.a.	4.2	-0.4	4.8	1.6	3.7	-1.7
Gross fixed capital formation	1.8	0.1	1.7	0.1	1.8	-0.9	4.1	n.a.	3.1	n.a.	3.2	-0.2	3.7	1.6	2.4	-0.9
Other capital expenditure	1.3	0.5	0.6	0.3	1.4	-0.6	3.1	n.a.	1.5	n.a.	0.9	-0.2	1.0	0.0	1.3	-0.8
Total expenditure	49.4	-4.0	54.7	-2.3	48.4	1.3	48.9	-0.4	39.9	-2.7	52.7	2.0	31.9	-11.3	48.0	-6.3
Budget balance	0.1	7.0	2.6	3.6	-1.4	1.6	-1.9	13.9	-0.9	2.9	-1.4	0.7	4.5	7.3	-1.8	9.9

Sources: AMECO (Annual Macro Economic Database, European Commission, Spring 2003 release) and Eurostat. In the remainder, the former will be referred to simply as AMECO.
Notes: All figures are ESA95 data, except the changes from 1990 to 2000 (Δ 00/90) in Greece, Spain and Sweden. For these three countries ESA95 fiscal series start later than 1990, and hence values for 1990 had to be inferred by deducting ESA79 growth rates from the earliest ESA95 figures available. This shortcut was applied to total revenue and total expenditure (with the budget balance obtained by difference), but not to their components, for the reasons given in note 1. For (unified) Germany ESA95 data starts in 1991, and hence changes in budget items are from this year to 2000 (Δ 00/91). Annex 1 gives ESA95 codes for the several items considered, and establishes a correspondence to AMECO variables. UMTS proceeds, received in 2000 in the case of Germany, Spain, Italy, the Netherlands, Austria, Portugal and the United Kingdom, have been excluded.

Table 1 (cont.)

Budget items in 2000 and their changes in the nineties (p.p. of GDP)

	Luxemburg		Netherlands		Austria		Portugal		Finland		Sweden		U. Kingdom	
	2000	**Δ 00/90**	**2000**	**Δ 00/90**	**2000**	**Δ 00/90**	**2000**	**Δ 00/90**	**2000**	**Δ 00/90**	**2000**	**Δ 00/90**	**2000**	**Δ 00/90**
Current revenue	45.5	-2.5	47.0	-2.2	50.6	0.1	40.8	6.8	55.6	1.1	60.7	n.a.	40.4	0.0
Indirect taxes	14.4	2.6	12.1	1.8	14.6	-1.0	14.4	1.5	13.6	-1.7	16.3	n.a.	13.8	1.6
Direct taxes	15.8	-1.1	12.1	-3.0	13.3	1.6	10.4	2.7	21.4	3.9	21.2	n.a.	16.6	-0.2
Social contributions	11.4	-1.1	17.1	0.6	16.9	1.3	11.8	2.3	12.2	-0.6	14.9	n.a.	7.6	0.1
Other current revenue	4.0	-2.9	5.7	-1.6	5.8	-1.9	4.2	0.2	8.3	-0.4	8.3	n.a.	2.4	-1.6
Capital revenue	0.2	0.0	0.5	0.2	0.2	0.1	1.4	0.0	0.3	0.2	0.2	n.a.	0.3	0.0
Total revenue	45.7	-2.5	47.4	-2.1	50.8	0.1	42.3	6.8	55.9	1.3	60.9	-7.2	40.6	0.0
Current expenditure	34.4	-3.0	42.4	-8.4	48.6	0.6	40.1	3.1	46.1	1.1	54.5	n.a.	37.5	-0.3
Intermediate consumption	3.1	0.2	6.3	0.2	5.3	-0.2	4.4	1.2	8.1	-0.3	10.2	n.a.	11.5	3.2
Compensation of employees	7.9	-2.1	10.0	-1.0	11.1	-0.8	15.0	3.2	13.2	-2.0	15.7	n.a.	7.2	-4.3
Social transfers other than in kind	13.9	-0.3	11.8	-6.5	18.5	0.7	12.4	3.0	16.5	1.6	17.5	n.a.	13.3	1.3
Interest	0.3	-0.1	3.9	-2.0	3.6	-0.4	3.2	-5.4	2.9	1.4	4.0	n.a.	2.8	-1.0
Subsidies	1.6	-0.9	1.5	-0.8	2.4	-0.7	1.1	-0.7	1.5	-1.4	1.6	n.a.	0.5	-0.4
Other current expenditure	7.5	0.3	8.9	1.6	7.8	2.0	4.1	1.8	4.0	1.7	5.5	n.a.	2.2	0.8
Capital expenditure	5.2	-0.7	3.5	-0.5	4.0	-1.0	5.4	0.2	2.8	-1.4	2.9	n.a.	1.6	-2.9
Gross fixed capital formation	4.1	-0.5	3.2	0.1	1.5	-1.5	3.8	0.5	2.6	-1.1	2.9	n.a.	1.1	-1.5
Other capital expenditure	1.1	-0.3	0.3	-0.7	2.5	0.6	1.5	-0.3	0.2	-0.3	0.0	n.a.	0.4	-1.4
Total expenditure	39.6	-3.7	45.9	-8.9	52.6	-0.4	45.4	3.4	48.9	-0.4	57.4	-5.6	39.1	-3.2
Budget balance	6.1	1.2	1.5	6.8	-1.9	0.6	-3.2	3.4	6.9	1.7	3.4	-1.6	1.6	3.1

Table 2

Changes in cyclically-adjusted budget items (p.p. of potential GDP)

	Revenue			**primary expenditure**			**primary balance (CAPB)**		
	1990-1995	1995-2000	1990-2000	1990-1995	1995-2000	1990-2000	1990-1995	1995-2000	1990-2000
BE	1.9	1.0	3.0	0.5	0.6	1.2	1.4	0.4	1.8
DK	2.1	-0.7	1.4	5.1	-2.7	2.5	-3.0	1.9	-1.1
DE	n.a.	0.9	n.a.	n.a.	-0.5	n.a.	n.a.	1.3	n.a.
EL	n.a.	7.8	n.a.	n.a.	4.8	n.a.	n.a.	3.0	n.a.
ES	n.a.	0.8	n.a.	n.a.	-1.3	n.a.	n.a.	2.1	n.a.
FR	0.5	2.2	2.7	1.9	0.0	1.8	-1.3	2.2	0.9
IE	-1.7	-2.2	-3.9	-1.6	-3.3	-4.9	-0.1	1.1	1.0
IT	3.2	0.5	3.6	-2.7	0.2	-2.5	5.8	0.3	6.1
NL	-2.5	0.5	-2.0	-5.7	-0.8	-6.5	3.2	1.3	4.5
AT	0.7	-0.6	0.0	2.3	-2.5	-0.2	-1.7	1.9	0.2
PT	3.4	3.2	6.6	2.6	5.5	8.2	0.8	-2.3	-1.5
FI	0.5	0.8	1.2	2.4	-4.6	-2.2	-1.9	5.4	3.4
SE	n.a.	0.8	n.a.	n.a.	-5.4	n.a.	n.a.	6.2	n.a.
UK	-1.8	1.8	0.0	1.9	-4.3	-2.5	-3.7	6.1	2.4

Source: AMECO.

Notes: Annex 3 gives details on the cyclical adjustment of fiscal variables. The countries are Belgium (BE), Denmark (DK), Germany (DE), Greece (EL), Spain (ES), France (FR), Ireland (IE), Italy (IT), the Netherlands (NL), Austria (AT), Portugal (PT), Finland (FI), Sweden (SE) and the United Kingdom (UK). No cyclically-adjusted figures are available for Luxemburg.

Table 3

Change in budget items from 1990 to 2000 by subsectors (p.p. of GDP)

	GG	CG			LG	SS
		State	ASF	Total CG		
Current primary expenditure	8.6	1.7	3.2	4.8	0.9	2.9
Intermediate consumption	1.2	-0.2	1.0	0.8	0.5	0.0
Compensation of employees	3.2	1.4	1.5	2.9	0.3	0.0
Social transfers	3.6	0.1	0.6	0.6	0.0	2.9
Subsidies	-0.7	-0.2	-0.1	-0.4	0.0	-0.3
Other current expenditure	1.2	0.6	0.3	0.8	0.1	0.3
Capital expenditure	0.2	n.a.	n.a.	-0.7	0.9	0.0
Gross fixed capital formation	0.5	-0.3	0.1	-0.2	0.7	0.0
Other capital expenditure	-0.3	n.a.	n.a.	-0.5	0.2	0.0

Source: Instituto Nacional de Estatística – Portugal.

Notes: UMTS proceeds have been excluded. Some components of "other capital expenditure" are only available for CG as a whole (i. e, without disaggregating into State and ASF), giving rise to lack of data (n.a.) in some entries of this table. The same happens with Table 4.

Table 4

Average real growth rates from 1990 to 2000 and subsectors' shares

		GG	CG			LG	SS
			State	ASF	Total CG		
Current primary expenditure	**growth rate**	**5.5**	**4.4**	**7.1**	**5.5**	**5.7**	**5.5**
	1990 share	**84.7**	**35.2**	**21.9**	**57.1**	**9.4**	**33.5**
	2000 share	**87.3**	**31.6**	**25.4**	**57.0**	**9.5**	**33.5**
Intermediate consumption	growth rate	6.3	1.3	9.3	5.9	8.1	0.4
	1990 share	9.4	37.1	36.0	73.1	22.8	4.1
	2000 share	10.4	23.0	47.7	70.7	27.0	2.3
Compensation of employees	growth rate	5.3	4.8	6.9	5.5	4.4	1.3
	1990 share	35.3	57.4	26.5	83.9	13.6	2.5
	2000 share	35.5	54.9	30.9	85.8	12.5	1.7
Social transfers	growth rate	5.9	3.7	7.0	5.8	7.1	5.9
	1990 share	31.2	7.3	11.0	18.3	0.9	80.8
	2000 share	33.2	6.0	12.2	18.2	1.0	80.9
Subsidies	growth rate	-2.0	-0.7	-0.9	-0.8	1.0	-9.3
	1990 share	5.2	49.3	24.0	73.3	5.2	21.5
	2000 share	2.5	56.4	26.7	83.1	7.0	9.9
Other current expenditure	growth rate	10.2	12.4	8.5	10.7	8.0	10.2
	1990 share	3.6	33.1	31.1	64.2	12.6	23.2
	2000 share	5.7	40.1	26.6	66.7	10.2	23.1
Capital expenditure	**growth rate**	**2.5**	**n.a.**	**n.a.**	**-0.7**	**7.4**	**6.6**
	1990 share	**15.3**	**n.a.**	**n.a.**	**68.0**	**30.5**	**1.5**
	2000 share	**12.7**	**n.a.**	**n.a.**	**52.8**	**45.2**	**2.0**
Gross fixed capital formation	growth rate	4.3	-1.6	3.7	1.7	6.7	9.7
	1990 share	10.0	23.5	30.7	54.2	44.9	0.9
	2000 share	9.1	13.2	28.9	42.1	56.4	1.5
Other capital expenditure	growth rate	-1.6	n.a.	n.a.	-3.9	17.0	4.1
	1990 share	5.4	n.a.	n.a.	93.7	3.8	2.5
	2000 share	3.6	n.a.	n.a.	79.9	16.6	3.4
Total primary expenditure	**growth rate**	**5.1**	**n.a.**	**n.a.**	**4.6**	**6.4**	**5.5**
	1990 share	**100.0**	**n.a.**	**n.a.**	**58.7**	**12.6**	**28.6**
	2000 share	**100.0**	**n.a.**	**n.a.**	**56.4**	**14.1**	**29.5**

Source: Instituto Nacional de Estatística – Portugal.

Notes: UMTS proceeds have been excluded. Shaded entries are average real growth rates, where nominal budget items have been deflated by the GDP deflator. The shares in the GG olumn are defined relative to total primary expenditure. The shares in the remaining columns are relative to the GG total of the respective item.

Table 5

Cyclicality coefficients for budget items in several European countries

	PT	AT	BE	DK	FR	IT	FI
Revenue							
Current revenue	1.3 ***	0.2	0.2	1.0 ***	0.7 **	0.2	0.7 ***
Indirect taxes	1.1 **	0.7 **	0.4	1.8 ***	0.4	0.7	1.1 ***
Direct taxes	1.8 **	-0.7	0.1	1.1 ***	1.7 *	-0.4	1.4 ***
Social contributions	1.1 **	-0.1	0.0	-1.3	0.6	0.3	0.1
Other current revenue	1.7	1.2	0.8	-1.0	-0.5	0.2	-0.5
Capital revenue	-2.6	-3.2	0.9	-1.0	-15.0	-7.5	-0.2
Total revenue	1.3 ***	0.2	0.2	0.9 ***	0.5 *	0.2	0.7 ***
Expenditure							
Current expenditure	0.6 *	0.1	-0.1	-0.3	-0.5	-0.2	-0.6 ***
Intermediate consumption	1.8 **	1.4 *	-0.3	-0.4	-1.0	0.6	-0.1
Compensation of employees	1.3 ***	0.5	0.3	-0.2	-0.2	0.6	0.0
Social transfers other than in kind	0.3	-0.3	-0.3	-0.5	-0.3	-0.6	-1.4 ***
Interest	-0.5	-0.7	-0.3	-0.7	-2.2	-0.8	-2.1 **
Subsidies	-1.3	-0.5	-0.3	-1.6	-1.3	-1.5	-0.4
Other current expenditure	0.6	-0.2	0.0	0.9	0.0	0.8	0.0
Capital expenditure	1.7	0.4	-2.5	-1.2	3.4 ***	3.6 *	0.5
Gross fixed capital formation	1.8	1.0	0.9	1.4	2.6 ***	2.9	1.0
Other capital expenditure	1.6	-0.1	-10.1	-12.7 **	6.5 *	4.3	-2.2
Total expenditure	0.8 **	0.1	-0.3	-0.3	-0.2	0.2	-0.5 ***
Current primary expenditure	0.8 **	0.1	-0.1	-0.3	-0.4	0.0	-0.5 ***
Total primary expenditure	0.9 ***	0.1	-0.4	-0.3	-0.1	0.3	-0.5 **

Sources: AMECO and Eurostat.

Notes: ESA95 data, with UMTS proceeds excluded. The sample is 1977-2002 for all countries, except France (1978-2002) and Italy (1980-2002). For the final years of the sample, data is provisional. Estimation is by OLS. However, if residuals are serially correlated, OLS estimates of standard errors do not offer a reliable basis for inference. Hence, estimation was also performed by GMM, in such a way as to replicate OLS estimates of β, but with heteroskedasticity and autocorrelation consistent covariances. Asterisks *, ** and *** denote significance at 10, 5 and 1% levels, and are based, for each variable and country, on the highest p-value of the different estimations (OLS and GMM, the latter with several possibilities for the number of autocorrelation terms).

Table 6

Cyclicality coefficients for cyclically-adjusted budget items

	PT	AT	BE	DK	FR	IT	FI
Revenue							
Current revenue	0.6 *	-0.3	-0.6 ***	0.2	0.1	-0.7	0.0
Indirect taxes	0.5	0.3	-0.4	0.5	-0.2	-0.6	0.4 *
Direct taxes	0.7	-1.6 *	-1.0 *	0.4	0.7	-1.4	0.4
Social contributions	0.3	-0.5	-0.8 ***	-1.9	0.1	-0.2	-0.8 **
Total revenue	0.6 *	-0.3	-0.6 ***	0.2	0.0	-0.7	0.0
Expenditure							
Current expenditure	0.7 *	0.1	0.1	0.1	-0.3	-0.1	-0.3 *
Social transfers other than in kind	0.6	-0.3	0.3	0.7	0.1	-0.4	-0.5 *
Total expenditure	0.8 **	0.1	-0.1	0.0	0.0	0.2	-0.3
Current primary expenditure	0.9 ***	0.1	0.2	0.1	-0.2	0.1	-0.2
Total primary expenditure	1.0 ***	0.1	-0.1	0.1	0.1	0.4	-0.2

Sources: AMECO and Eurostat.

Notes: see notes under Table 5. Cyclically-adjusted figures for overall revenues and expenditures are available in AMECO; for particular revenue and expenditure items, OECD elasticities (Van den Noord, 2000) were used. Annex 3 gives details.

Table 7 (cont.)

Sensitivity to cyclical and budgetary conditions

		GAP		s_{-1}	
PT	Δs^{CA}	0.04		-0.49	**
	p-value	0.75		0.02	
	Δt^{CA}	0.26	***	-0.42	***
	p-value	0.00		0.00	
	Δg^{CA}	0.22	**	0.07	
	p-value	0.04		0.68	
AT	Δs^{CA}	0.04		-0.25	
	p-value	0.80		0.21	
	Δt^{CA}	0.04		0.01	
	p-value	0.78		0.95	
	Δg^{CA}	0.00		0.26	
	p-value	0.97		0.15	
BE	Δs^{CA}	-0.35		0.00	
	p-value	0.24		0.98	
	Δt^{CA}	-0.28	**	0.04	
	p-value	0.03		0.42	
	Δg^{CA}	0.07		0.03	
	p-value	0.80		0.73	
DK	Δs^{CA}	0.45	**	-0.14	
	p-value	0.05		0.21	
	Δt^{CA}	0.13		-0.16	*
	p-value	0.47		0.08	
	Δg^{CA}	-0.32		-0.03	
	p-value	0.14		0.81	
FR	Δs^{CA}	-0.01		-0.22	*
	p-value	0.96		0.09	
	Δt^{CA}	-0.06		0.00	
	p-value	0.55		0.99	
	Δg^{CA}	-0.06		0.22	**
	p-value	0.52		0.03	
IT	Δs^{CA}	-0.20		0.01	
	p-value	0.45		0.95	
	Δt^{CA}	0.18		-0.17	**
	p-value	0.41		0.02	
	Δg^{CA}	0.38	*	-0.18	**

Table 7 (cont.)

Sensitivity to cyclical and budgetary conditions

		GAP		s_{-1}	
	p-value	0.08		0.02	
FI	Δs^{CA}	0.33	***	-0.18	**
	p-value	0.00		0.03	
	Δt^{CA}	0.10		0.00	
	p-value	0.36		1.00	
	Δg^{CA}	-0.22	**	0.18	**
	p-value	0.04		0.04	

Sources: AMECO.

Notes: ESA95 data, with UMTS proceeds excluded (except in the case of variable s_{t-1}). The sample is 1977-2002 for all countries, except France (1978-2002) and Italy (1980-2002). For the final years of the sample, data is provisional. Estimation is by OLS. Asterisks *, ** and *** denote significance at 10, 5 and 1% levels.

Table 8

Cyclicality coefficients by subsectors

	GG	**CG**				**LG**	**SS**
		State		ASF		Total CG		
Current primary expenditure	0.8 **	0.9 *		1.7 ***		1.0 **	1.7 ***	0.3
Intermediate consumption	1.8 **	1.2		3.7 ***		1.8 *	1.4 **	0.9
Compensation of employees	1.3 ***	1.3 **		1.4 ***		1.3 ***	1.6 ***	2.0 **
Social transfers	0.3	0.2		0.1		0.4	4.5 **	0.2
Subsidies	-1.3	-0.3		2.7		-0.6	0.2	-4.4
Other current expenditure	1.1	-1.7		6.6 *		0.4	3.5 **	2.7
Capital expenditure	1.7	n.a.		n.a.		1.7	1.9	4.6 **
Gross fixed capital formation	1.9 **	3.3		0.4		1.8	2.2 *	4.4
Other capital expenditure	1.7	n.a.		n.a.		2.0	0.8	5.2
Total primary expenditure	0.9 ***	n.a.		n.a.		1.0 **	1.7 **	0.3

Source: Instituto Nacional de Estatística and Ministério das Finanças – Portugal.

Notes: UMTS proceeds have been excluded. The estimation samples are 1978-2002 for GG, CG, LG and SS, and 1988-2002 for the State and ASF. Estimation methods and meaning of asterisks as in Table 5.

Table 9

Forecasts and forecast errors of GFCF and interest payments (p.p. of GDP, Portugal)

	$GFCF^{F}$	$GFCF^{E1}$	$GFCF^{E2}$	IP^{F}	IP^{E1}	IP^{E2}	*i*
1994	4.7	-0.3	-0.8	5.9	-0.1	0.0	12.1
1995	4.7	-0.5	-0.8	5.5	0.3	0.1	10.9
1996	4.5	-0.3	-0.4	5.2	-0.4	-0.4	8.9
1997	4.4	-0.1	-0.1	4.6	-0.3	-0.3	7.3
1998	4.4	-0.2	-0.1	3.7	-0.2	-0.3	6.4
1999	4.3	0.3	-0.2	3.4	-0.2	-0.3	6.2
2000	4.9	-1.1	-1.2	3.2	-0.1	-0.1	6.3
2001	4.3	-0.3	-0.2	3.1	0.0	0.0	6.2
2002	4.1	-0.5	n.a.	3.1	-0.1	n.a.	5.7

Sources: author's calculations based on data from the March EDP reporting (several years), kindly supplied by Ministério das Finanças. Implicit interest rate on public debt (i) from AMECO.

Notes: the notation is explained in the main text, and methodological aspects are detailed in annex 5.

Table 10

Public GFCF in Portugal and in the EU average (p.p. of GDP)

q	GFCF in PT	CSF: EU	CSF: all	GFCF in PT net of CSF		GFCF in EU15	Differences between PT and EU15		
	(1)	(2)	(3)	(4) = (1)-(2)	(5) = (1)-(3)	(6)	(7) = (1)-(6)	(8) = (4)-(6)	(9) = (5)-(6)
1995	3.7	1.2	1.8	2.5	1.9	2.6	1.1	-0.1	-0.7
1996	4.2	1.4	2.1	2.8	2.1	2.5	1.7	0.3	-0.4
1997	4.4	1.3	1.8	3.1	2.5	2.2	2.1	0.8	0.3
1998	3.9	0.9	1.3	3.0	2.7	2.3	1.7	0.8	0.4
1999	4.1	0.8	1.1	3.4	3.1	2.3	1.9	1.1	0.8
2000	3.8	0.7	0.9	3.2	2.9	2.3	1.6	0.9	0.6
2001	4.1	1.0	1.4	3.1	2.7	2.3	1.8	0.8	0.4
2002	3.6	0.8	1.3	2.8	2.4	2.2	1.4	0.6	0.2
1995-2002	4.0	1.0	1.5	3.0	2.5	2.3	1.7	0.6	0.2

Sources: AMECO and DPP (*Departamento de Prospectiva e Planeamento, Ministério das Finanças*) for estimates of CSF amounts (see text for details).

Table 11

Unemployment benefits (p.p. of GDP) and unemployment rate in Portugal

	UB	SUB	UB+SUB	U rate
1990	0.13	0.06	0.20	4.8
1991	0.20	0.08	0.28	4.2
1992	0.28	0.11	0.40	4.3
1993	0.45	0.15	0.60	5.6
1994	0.50	0.18	0.67	6.9
1995	0.49	0.19	0.68	7.3
1996	0.46	0.20	0.66	7.3
1997	0.43	0.19	0.62	6.8
1998	0.41	0.18	0.59	5.1
1999	0.43	0.19	0.62	4.5
2000	0.49	0.18	0.67	4.1
2001	0.51	0.19	0.70	4.1
2002	0.61	0.22	0.83	5.1

Sources: IGFSS and IIESS (Ministério da Segurança Social e do Trabalho) for UB and SUB, AMECO for GDP and the unemployment rate (variable code ZUTN).

Table 12

Average rate of real GDP growth in Portugal over previous 10 years

1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
3.2	3.5	3.7	3.6	3.4	3.1	2.8	2.8	2.5	2.5

Source: AMECO.

Table 13

Annual % changes in the GDP deflator (current euro area members)

	PT	AT	BE	FR	IT	FI	NL	IE	LU	DE	ES	EL
1999	3.1	0.7	1.4	0.5	1.6	-0.3	1.6	4.1	3.1	0.5	2.7	3.0
2000	3.2	1.4	1.3	0.7	2.1	2.9	4.2	4.3	2.8	-0.3	3.5	3.4
2001	4.7	1.6	2.0	1.4	2.7	3.6	5.3	5.3	2.3	1.4	4.2	3.4
2002	4.6	1.3	2.3	1.7	2.7	1.3	3.2	5.6	1.7	1.6	4.4	3.7
average	3.9	1.3	1.7	1.1	2.3	1.8	3.5	4.8	2.5	0.8	3.7	3.4
st. dev.	0.8	0.4	0.4	0.5	0.5	1.5	1.4	0.6	0.5	0.8	0.6	0.2

Source: AMECO.

Table 14

Forecast errors in Stability or Convergence Programmes (p.p. of GDP, 1999-2002)

		PT	AT	BE	FR	IT	FI	NL	IE	LU	DE	ES	EL
real GDP	avg error	-0.7	-0.5	-0.4	-0.5	-0.8	-0.6	-0.4	0.9	-0.3	-0.8	-0.2	-0.2
Growth	avg abs error	0.7	0.8	1.0	0.7	1.1	1.3	1.6	1.9	2.8	0.9	0.3	0.4
	avg error	-0.6	0.0	0.2	-0.2	-0.4	0.5	0.3	-0.2	0.0	-0.4	0.2	-0.3
Budget	avg abs error	0.6	0.4	0.3	0.5	0.5	0.6	0.9	1.9	1.6	0.8	0.3	0.8
Balance	Avg error, adj.	-0.4	0.2	0.4	0.0	-0.1	0.9	0.6	-0.5	0.2	0.0	0.2	-0.2
	avg abs error, adj.	0.5	0.6	0.4	0.2	0.4	1.4	0.8	1.7	1.1	0.4	0.3	0.7

Sources: European Commission (2000, 2001, 2002, 2003) and information on national Stability or Convergence Programmes available at a European Commission's website

(http://europa.eu.int/comm/economy_finance/index_en.htm).

Notes: the table reports summary statistics for forecast errors, defined as "first available estimates" minus forecasts. Forecasts are values for year *n* contained in the Programme prepared at the end of year *n-1* or at the beginning of year *n*. "First available estimates" are the values for year *n* given in the following Programme. The sample has four observations: *n* = 1999, 2000, 2001, 2002. The forecast error for *n* = 1999 is affected by the change from ESA79 to ESA95.

Table 15

Portuguese debt-to-GDP ratio – level and decomposition of changes (p.p. of GDP)

	debt	Δdebt	primary balance	privatiz. revenues	other stock--flow adj.	interest effect	Growth Effect	snowball effect
		(1)+(2)+...+(5)	(1)	(2)	(3)	(4)	(5)	(4)+(5)
1990	58.3							
1991	60.7	2.5	-1.2	-0.6	3.1	8.8	-7.6	1.2
1992	54.4	-6.3	-3.8	-1.5	-2.7	8.5	-6.8	1.7
1993	59.1	4.7	0.3	-0.4	-0.3	7.7	-2.7	5.1
1994	62.1	3.0	1.1	-0.2	0.0	6.6	-4.5	2.1
1995	64.3	2.2	-0.8	-0.8	2.0	6.3	-4.5	1.7
1996	62.9	-1.4	-0.6	-1.7	-0.4	5.4	-4.0	1.3
1997	59.1	-3.8	-0.7	-3.4	0.6	4.2	-4.6	-0.4
1998	55.0	-4.1	-0.3	-1.4	-1.2	3.5	-4.7	-1.2
1999	54.3	-0.7	-0.4	-0.6	0.7	3.2	-3.6	-0.4
2000	53.3	-1.0	-0.4	-0.6	0.3	3.2	-3.5	-0.3
2001	55.6	2.3	1.1	-0.2	1.5	3.1	-3.2	-0.1
2002	58.1	2.6	-0.3	-0.3	2.8	3.0	-2.7	0.3

Sources: AMECO and IGCP (Instituto de Gestão do Crédito Público).

Table 16

Revisions to the budget balance/GDP ratios (p.p. of GDP)

	BE	DK	DE	EL	ES	FR	IE	IT	LU	NL	AT	PT	FI	SE	UK
1994	0.1	0.5	0.1	0.4	0.5	0.2	0.2	-0.1	0.3	0.4	-0.5	0.2	-0.7	-0.4	0.1
1995	0.4	-0.4	0.0	0.0	-0.8	0.1	-0.1	0.1	0.6	-0.3	0.9	-0.6	0.6	0.2	0.4
1996	0.1	1.0	0.5	-0.1	-0.2	0.0	0.5	0.1	-0.2	-0.1	-0.1	0.8	-0.7	0.1	-0.3
1997	0.2	-0.3	0.0	0.2	0.1	0.0	0.2	0.0	1.2	0.5	0.6	-0.1	-0.3	0.1	0.0
1998	0.3	0.4	0.4	-0.1	-0.8	0.2	-0.2	-0.1	1.1	0.2	-0.4	0.2	0.3	-0.1	-0.3
1999	0.2	0.1	-0.2	-0.2	0.0	0.2	0.1	0.1	2.3	0.4	-0.1	0.0	-0.5	0.0	0.1
2000	0.1	0.0	-0.1	0.1	0.0	0.0	0.0	-0.2	0.5	0.2	-0.3	-0.2	0.3	-0.4	-0.3
2001	0.3	0.3	0.0	-1.5	-0.1	-0.1	-0.6	-1.2	1.4	-0.1	0.2	-1.9	0.3	-0.2	-0.2
avg abs rev	0.21	0.38	0.18	0.33	0.31	0.11	0.25	0.24	0.95	0.26	0.38	0.50	0.45	0.18	0.22

Sources: European Commission (EDP reporting).
Notes: revisions are defined as the difference between the figure for year n reported in March of *n+2* and the figure for year n reported in March of *n+1*. Year n is given in the first column. Values for 1998 are affected by the move from ESA79 to ESA95.

Table A.2

Cyclicality coefficients for budget items in several European countries

	PT	AT	BE	DK	FR	IT	FI
Revenue							
Current revenue	1.1 ***	0.6 ***	0.1	1.5 ***	0.5 **	0.2	0.7 ***
Indirect taxes	0.8 ***	0.7 **	0.6 **	2.4 ***	0.0	0.8	1.4 ***
Direct taxes	1.3 **	0.4	-0.4	1.7 ***	1.7	-0.1	1.5 ***
Social contributions	1.0 ***	0.3	0.0	0.2	0.3	-0.2	-0.1
Other current revenue	1.3 *	1.6 **	1.2	-1.0 **	-0.1	0.5	-0.7 **
Capital revenue	0.9	-2.2	1.8	2.3	-47.7	-16.2	0.2
Total revenue	1.1 ***	0.6 ***	0.1	1.5 ***	0.4 **	0.0	0.7 ***
Expenditure							
Current expenditure	0.8 ***	0.2	-0.1	-0.5 *	-0.6 ***	-0.1	-0.7 ***
Intermediate consumption	1.4 ***	0.5	-0.2	-0.3	-0.3	1.1 **	-0.1
Compensation of employees	1.5 ***	0.4	0.0	-0.4	-0.5 **	1.3 **	0.1
Social transfers other than in kind	0.3	-0.2	-0.5	-1.3 ***	-0.5 **	-0.5 *	-1.6 ***
Interest	-0.1	0.1	0.1	0.9	-2.9 ***	-2.8 *	-2.8 ***
Subsidies	-0.3	0.5	-0.2	-3.3 ***	-1.6 *	-1.0	-0.8 *
Other current expenditure	0.8	0.3	0.2	1.3 **	-0.2	3.6 ***	0.5
Capital expenditure	1.1 *	0.2	-0.9	-2.2 **	2.7 ***	3.9 ***	0.8 **
Gross fixed capital formation	2.0 ***	0.4	0.5	1.9 *	2.7 ***	3.8 ***	1.3 ***
Other capital expenditure	-1.0	-0.7	-3.6	-27.1 ***	2.7	4.2	-1.6
Total expenditure	0.8 ***	0.2	-0.1	-0.5 **	-0.4 *	0.3	-0.6 ***
Current primary expenditure	0.9 ***	0.2	-0.2	-0.7 **	-0.5 **	0.6 *	-0.6 ***
Total primary expenditure	0.9 ***	0.2	-0.2	-0.8 **	-0.2	0.9 ***	-0.5 ***

Sources: AMECO and Eurostat.

Notes: Those under Table 5 apply. Further, HP filtering used $\lambda = 100$, as is often the case in yearly series. To alleviate the end-point bias problem, filtering took place until 2004, using the European Commission Spring 2003 forecasts. Notice that, for the purposes of this paper, a "start-point bias" exists as well. It has been dealt with in a limited way, because in the case of Portugal (and also for some other countries) there are no observations available before the start of our estimation sample; still, the series for filtering starts one year before the series for estimation, matching the initial observation lost due to first differencing in table 5.

Table A.3

Cyclicality coefficients for cyclically-adjusted budget items

	PT	AT	BE	DK	FR	IT	FI
Revenue							
Current revenue	0.4 *	0.2	-0.7 ***	0.8 ***	-0.1	-0.6 *	0.1
Indirect taxes	0.3	0.3	-0.1	1.2 ***	-0.6 **	-0.4	0.7 ***
Direct taxes	0.2	-0.5	-1.5 ***	1.1 ***	0.7	-1.0 *	0.7 **
Social contributions	0.2	-0.1	-0.8 ***	-0.3	-0.1	-0.7	-0.9 ***
Total revenue	0.4 *	0.2	-0.7 ***	0.8 ***	-0.1	-0.8 *	0.1
Expenditure							
Current expenditure	0.9 ***	0.2	0.1	-0.1	-0.5 **	0.0	-0.4 ***
Social transfers other than in kind	0.6	-0.2	0.1	-0.2	-0.1	-0.3	-0.8 ***
Total expenditure	0.9 ***	0.2	0.1	-0.2	-0.2	0.3	-0.4 ***
Current primary expenditure	1.0 ***	0.2	0.1	-0.3	-0.3	0.7 **	-0.3 ***
Total primary expenditure	1.0 ***	0.2	0.0	-0.4	-0.1	1.0 ***	-0.2 *

Sources: AMECO and Eurostat.

Notes: those under Tables 6 and A.2 also apply here.

Table A.4

Cyclicality coefficients by subsectors

	GG	CG			LG	SS
		State	ASF	Total CG		
Current primary expenditure	0.9 ***	0.6	1.3 ***	1.0 ***	2.0 ***	0.4
Intermediate consumption	1.4 ***	1.3	1.8 ***	1.3 **	1.8 ***	1.1
Compensation of employees	1.5 ***	0.7 *	1.0 ***	1.4 ***	2.0 ***	1.7 ***
Social transfers	0.4	0.2	0.4	0.9 *	4.1 ***	0.2
Subsidies	-0.5	0.3	3.2	-0.8	1.4	-0.9
Other current expenditure	1.3	-0.8	4.7 **	0.6	3.2 ***	2.6 **
Capital expenditure	1.0	n.a.	n.a.	0.3	2.3 ***	4.4 ***
Gross fixed capital formation	1.8 ***	-0.7	-0.9	1.2	2.5 ***	5.1
Other capital expenditure	-0.9	n.a.	n.a.	-1.3	0.7	4.2 **
Total primary expenditure	0.9 ***	n.a.	n.a.	0.9 ***	2.1 ***	0.5

Source: *Instituto Nacional de Estatística* and *Ministério das Finanças* – Portugal.

Notes: those under Table 8 also apply here. HP filtering performed as in Table A.2, though without any end-point bias correction (due to the absence of expenditure projections for the subsectors).

References

Afonso, A. (2002), "Efeitos das Consolidações Orçamentais na UE-15: o Caso de Portugal", *Direcção-Geral de Estudos e Previsão*, Ministry of Finance, Working Paper no. 25, January.

Alesina, A. and R. Perotti (1999), "Budget Deficits and Budget Institutions", in J. Poterba and J. von Hagen (eds.), *Fiscal Institutions and Fiscal Performance*, NBER, University of Chicago Press, 13-36.

Artis, M. and M. Marcellino (2001), "Fiscal Forecasting - the Track Record of the IMF, OECD and EC", *Econometrics Journal*, 4(1), 20-36.

Artis, M. and B. Winkler (1998), "The Stability Pact: Safeguarding the Credibility of the European Central Bank", *National Institute Economic Review*, 163, 87-98.

Auerbach, A. (2002), "Is There a Role for Discretionary Fiscal Policy?", NBER Working Paper 9306.

Balassone, F. and D. Franco (2000), "Public Investment, the Stability Pact and the 'Golden Rule'", *Fiscal Studies*, 21(2), 207-229.

Banco de Portugal (2001), *Relatório do Conselho de Administração 2000*, Lisboa.

Barbosa, A. S. P. (2002), "Nota sobre uma lei 'explosiva'", *Boletim Económico*, Banco de Portugal, Dezembro, 27-29.

Blanchard, O. and R. Perotti (2002), An Empirical Characterisation of the Dynamic Effects of Changes in Government Spending and Taxes on Output", *Quarterly Journal of Economics*, November, 1329-1368.

Bohn, H. (1998), "The Behavior of U.S. Public Debt and Deficits", *Quarterly Journal of Economics*, August, 949-963.

Bronchi, C. (2003), "The Effectiveness of Public Expenditure in Portugal", OECD Economics Department Working Paper no. 349, February.

Brunila, A. and H. Kinnunen (2002), "Spending limits and fiscal discipline in euro area countries", *Bank of Finland Bulletin*, 1, 18-22.

Buti, M., S. Eijffinger and D. Franco (2003), "Revisiting the Stability and Growth Pact: Grand Design or Internal Adjustment?", CEPR Discussion Paper 3692.

Drazen, A. (2000), *Political Economy in Macroeconomics*, Princeton University Press.

ECORDEP (2001), *Relatório da Estrutura de Coordenação para a Reforma da Despesa Pública*, Ministério das Finanças, September.

Eichengreen, B., R. Hausmann and J. von Hagen (1999), "Reforming Budgetary Institutions in Latin America: The Case for a National Fiscal Council", *Open Economies Review*, 10, 415-442.

European Commission (2000), *European Economy. Public Finances in EMU 2000*.

European Commission (2001), *European Economy. Public Finances in EMU 2001*.

European Commission (2002), *European Economy. Public Finances in EMU 2002*.

European Commission (2003), *European Economy. Public Finances in EMU 2003*.

Gali, J. and R. Perotti (2003), "Fiscal Policy and Monetary Integration in Europe", mimeo, March.

Hallerberg, M. and R. Strauch (2002), "On the Cyclicality of Public Finances in Europe", *Empirica*, 29, 183-207.

Hallerberg, M., R. Strauch and J. von Hagen (2001), *The Use and Effectiveness of Budgetary Rules and Norms in EU Member States*, Report prepared for the Dutch

Ministry of Finance by the Institute of European Integration Studies (Final Version: June 11, 2001).

Heeringa, W. and Y. Lindh (2001), "Dutch Versus Swedish Budgetary Rules: a Comparison" in Banca d'Italia, *Fiscal Rules* (papers presented at the Bank of Italy workshop held in Perugia, 1-3 February 2001), 487-513.

Inman, R. (1996), "Do Balanced Budget Rules Work? U.S. Experience and Possible Lessons for the EMU", NBER Working Paper no. 5838.

Kopits, G. and S. Symansky (1998), "Fiscal Policy Rules", IMF Occasional Paper no. 162.

Lane, P. (1998), "On the Cyclicality of Irish Fiscal Policy", *The Economic and Social Review*, 29(1), 1-16.

Lane, P. (2003), "The Cyclical Behaviour of Fiscal Policy: Evidence from the OECD", *Journal of Public Economics*, 87(12), 2661-2675.

Lindh, Y. and H. Ohlsson (2002), "Fiscal Policy in a Currency Union – Instruments and Implementation" in Banca d'Italia, *The Impact of Fiscal Policy* (papers presented at the Bank of Italy workshop held in Perugia, 21-23 March 2002), 665-713.

Marinheiro, C. (1998), "O Pacto de Estabilidade e a Politica Orçamental Portuguesa: a aplicação de um modelo VAR", *Notas Económicas*, 10, 181-99.

Mills, P. and A. Quinet (2001), "The Case for Spending Rules" in Banca d'Italia, *Fiscal Rules* (papers presented at the Bank of Italy workshop held in Perugia, 1-3 February 2001), 319-330.

OECD (2001), *OECD Economic Surveys – Portugal*, April.

OECD (2002), *OECD Economic Outlook* no. 72.

OECD (2003), *OECD Economic Surveys – Portugal*, February.

Sapir, A. *et al.* (2003), *An Agenda for a Growing Europe: Making the EU Economic System Deliver*, Report of an Independent High-Level Study Group established on the initiative of the President of the European Commission, July.

Seidman, L. (2001), "Reviving Fiscal Policy", *Challenge*, 44(3), 17-42.

Taylor, J. (2000), "Reassessing Discretionary Fiscal Policy", *Journal of Economic Perspectives*, 14(3), 21-36.

Van den Noord, P., (2000) "The Size and Role of Automatic Stabilisers in the 1990s and Beyond", OECD Economics Department Working Paper No. 230.

von Hagen, J., A. Hughes Hallett and R. Strauch (2001), "Budgetary Consolidation in EMU", ECFIN Economic Papers 148.

Wren-Lewis, S. (2000), "The Limits to Discretionary Fiscal Stabilization Policy", *Oxford Review of Economic Policy*, 16(4), 92-105.

Wyplosz, C. (2001), "Fiscal Policy: Institutions vs. Rules", report prepared for the Swedish Government's Committee on Stabilization Policy in the EMU, December.

Wyplosz, C. (2002), "Fiscal Discipline in EMU: Rules or Institutions?", paper prepared for the April 16, 2002 meeting of the Group of Economic Analysis of the European Commission.

PARA QUE SERVEM AS PENSÕES MÍNIMAS?

Miguel Gouveia
FCEE, Universidade Católica Portuguesa

Carlos Farinha Rodrigues
ISEG, Universidade Técnica de Lisboa e
Departamento de Síntese Económica de Conjuntura do INE

Resumo

Quando uma pensão resultante das regras usuais da Segurança Social é baixa, essa pensão é reforçada com "complementos sociais" para que o total atinja o nível da pensão mínima. Uma larga proporção dos pensionistas portugueses recebe pensões mínimas gerando despesas públicas muito significativas. Os Governos portugueses adoptaram a política de aumentar faseadamente as pensões mínimas até ao nível do salário mínimo

Neste trabalho usamos o Inquérito aos Orçamentos Familiares de 2000 para estudar as características redistributivas das pensões mínimas. A análise mostra que as pensões mínimas são instrumentos caros e ineficazes de solidariedade. Só 31.25% das pessoas que vivem em agregados familiares recebendo pensões mínimas são pobres

Usamos também os dados para simular um aumento das pensões de 60% para 70% do salário mínimo líquido em 2000. Os resultados mostram que os indicadores de pobreza se reduziriam muito pouco apesar da tal medida acarretar um acréscimo elevado da despesa pública, principalmente porque 67.9% deste acréscimo incide em pessoas não pobres. Em contrapartida um programa simulado com base numa condições de recursos muito grosseira consegue obter resultados melhores em termos de pobreza e equidade e gastar menos de metade dos recursos.

1. Introdução.

1.1. As Pensões Mínimas e as Prioridades da Política Social

O Governo actual, na linha de uma política que já vinha sendo seguida pelos Governos anteriores, adoptou uma política de crescimento real acelerado das pensões mínimas do regime geral da Segurança Social tendo como objectivo atingir a breve prazo um montante equivalente ao salário mínimo nacional (líquido das contribuições para a Segurança Social).

A Lei n.º 32/2002 de 20 de Dezembro, que aprova as bases da Segurança Social, contém a definição da política no seu artigo 38º que citamos integralmente:

Artigo 38.º

Princípio de convergência das pensões mínimas

1. Os mínimos legais das pensões de invalidez e de velhice são fixados, tendo em conta as carreiras contributivas, com referência e até ao limite do valor da remuneração mínima mensal garantida à generalidade dos trabalhadores, deduzida da quotização correspondente à taxa contributiva normal do regime dos trabalhadores por conta de outrem.
2. As pensões que não atinjam o valor mínimo previsto no número anterior correspondentes às suas carreiras contributivas são acrescidas do complemento social previsto na alínea c) do n.º 1 do artigo 57.º, de montante a fixar na lei.
3. Sem prejuízo do disposto no número seguinte a fixação dos mínimos legais das pensões de invalidez e de velhice convergirá para o valor da remuneração mínima mensal garantida à generalidade dos trabalhadores, deduzida da quotização correspondente à taxa contributiva normal do regime dos trabalhadores por conta de outrem, e será estabelecida com base no sistema de escalões relacionados com as carreiras contributivas:

    a) Até 14 anos de carreira contributiva inclusive, será igual a 65% da remuneração mínima mensal garantida à generalidade dos trabalhadores, deduzida da quotização a que se refere o n.º 1 do presente artigo;
    b) Entre 15 e 20 anos de carreira contributiva inclusive, será igual a 72,5% da remuneração mínima mensal garantida à generalidade dos trabalhadores, deduzida da quotização a que se refere o n.º 1 do presente artigo;
    c) Entre 21 e 30 anos de carreira contributiva inclusive, será igual a 80% da remuneração mínima mensal garantida à generalidade dos trabalhadores, deduzida da quotização a que se refere o n.º 1 do presente artigo;
    d) Mais de 30 anos de carreira contributiva, será igual à da remuneração mínima mensal garantida à generalidade dos

trabalhadores, deduzida da quotização a que se refere o n.º 1 do presente artigo.

4. O escalonamento de convergência das carreiras contributivas previsto no número anterior, será concretizado, de forma gradual e progressiva, no prazo máximo de quatro anos contado após a data da entrada em vigor do Orçamento de Estado para 2003.
5. A verificação de condições económicas, orçamentais ou outras excepcionalmente adversas, poderão justificar uma dilação máxima de um ano na aplicação do disposto nos números anteriores.

Diário da República nº 294, I Série A, de 20 de Dezembro de 2002

Em resumo, as pensões de velhice em Portugal dependem do número de anos de contribuição e dos salários declarados que constituíram a base dessas contribuições. No entanto, se a aplicação das regras gerais resultarem num montante (pensão estatutária) inferior a um limite definido (que é função do número de anos de contribuições), à pensão estatutária soma-se um complemento social por forma ao total igualar a pensão mínima correspondente.

A política de fazer crescer rapidamente as pensões mínimas por forma a convergir com o salário mínimo parece ser bastante consensual entre os políticos portugueses. Quer nos debates públicos em geral quer no Parlamento a discussão política parece centrar-se mais sobre a velocidade com que a convergência para o salário mínimo deveria ocorrer do que sobre a desejabilidade deste objectivo final. Na prática a oposição clama sempre que o Governo não está a fazer crescer as pensões mínimas o suficiente! Pelo lado dos Governos, a única razão que trava um crescimento mais acelerado são as dificuldades em compatibilizar essa trajectória com as restrições orçamentais do sector público.

A ausência de uma visão crítica desta política poderia levar a pensar que, com excepção do óbvio custo orçamental, ela seria inquestionavelmente a política mais correcta e eficaz do ponto de vista de redistribuição do rendimento, combate à pobreza e provisão de segurança económica para grupos vulneráveis da população. Contudo, há alguns aspectos menos positivos da política seguida que não devem ser descurados.

1.2. Os Problemas Potenciais

A principal função de um sistema de pensões é assegurar a substituição dos rendimentos do trabalho quando se dá incapacidade de os gerar, seja por idade, incapacidade ou morte. Trata-se de um mecanismo de seguro social e não de um mecanismo cujo objectivo primordial seja a redistribuição do rendimento. Por isso, analisar o sistema de pensões apenas do ponto de vista do seu impacto na desigualdade ou na pobreza seria adoptar uma visão extremamente reducionista.

No entanto, ao introduzir no seu regime geral de pensões as pensões mínimas e os complementos de pensão, a Segurança Social portuguesa criou um subprograma de redistribuição do rendimento que pode e deve ser estudado com especificidade dada a sua relevância orçamental e política. Note-se que não é sequer defensável que as pensões mínimas sejam vistas apenas como uma não-linearidade de um mecanismo de seguro social porque a legislação da Segurança Social estipula que os complementos de pensão devem ser pagos pelo sistema fiscal (Orçamento do Estado) e não pelas contribuições para a

Segurança Social. A única conclusão possível é que se trata então de um mecanismo puro de redistribuição do rendimento a partir dos recursos públicos do país e que se deve dirigir aos quem têm mais necessidades. Em última análise isso significa que a desejabilidade de tal programa deverá sempre aferir-se pelos seus efeitos nos níveis de pobreza. E é ao chegar a este ponto da conceptualização que começamos a encontrar alguns problemas.

De um ponto de vista da análise económica, não é nada óbvio que a política definida em relação às pensões mínimas seja particularmente adequada, quer numa óptica de consolidação das finanças públicas, quer numa óptica de eficiência quer, acima de tudo numa óptica de equidade e redistribuição correcta do rendimento e redução da pobreza.

Numa **óptica de sustentabilidade das finanças públicas**, **aumentos das pensões mínimas são bastantes onerosos** em termos do aumento das despesas com pensões da Segurança Social e das transferências do Orçamento de Estado para a Segurança Social que daí decorrem dada a obrigatoriedade de o Orçamento do Estado pagar à Segurança Social os complementos de pensão. Isso significa que qualquer aumento da pensão mínima tem um impacto recorrente (sustentado) de primeira ordem nas despesas públicas. Para se ter uma noção do impacto orçamental das pensões mínimas basta referir que em 2000 44.2% dos pensionistas de velhice do regime geral recebiam pensões inferiores ou iguais a 170€ mensais (INE (2003)).

Numa óptica de eficiência, o aumento das pensões mínimas amplia o conjunto de trabalhadores para quem deixa de existir qualquer relação marginal entre as contribuições para a Segurança Social e os benefícios gerados por essas contribuições. Isso significa que as contribuições têm todos os custos de eficiência associados com impostos (carga excedente, evasão, desincentivo à procura e oferta de trabalho, etc.) em vez do custo de eficiência muito menor associado com a aquisição de benefícios obrigatórios (por exemplo poupança forçada). Isso significa que as pensões mínimas não devem ser vistas como um programa anti-pobreza clássico de provisão universal (cf. Besley (1990)) mas como um programa com claros desincentivos, à semelhança dos programas típicos com condição de recursos ("means testing").

Numa **óptica de equidade, o impacto das pensões mínimas medido em termos de redução de índices de pobreza é bastante mais pequeno do que geralmente se supõe. A razão para isso está na ausência de *targetting***: não há nenhum mecanismo de "*means testing*" que garanta que os recipientes das pensões mínimas sejam alvos prioritários, ou seja que estejam abaixo da linha de pobreza. De facto, há duas ordens de razões que permitem antecipar que em muitos casos os recursos gastos com os complementos sociais não terão impacto sobre os níveis de pobreza.

Por um lado, há muitos pensionistas que correctamente anteciparam as suas pensões reduzidas e têm outros meios de sustentação. Por exemplo, muitos dos trabalhadores independentes (comerciantes, profissionais liberais, etc.) consideram a Segurança Social um mau negócio e usam a discricionariedade que as suas circunstâncias e a lei lhes facultam para contribuir o mínimo possível para a Segurança Social. Isso significa que as pensões estatutárias serão muito reduzidas e nalguns casos inferiores às pensões mínimas. No entanto, pessoas nestas circunstâncias têm outras fontes de rendimento durante a reforma, por exemplo continuam a ser proprietários de estabelecimentos comerciais, retêm participações em empresas, dispõem de património suficiente para garantir um nível de vida adequado, etc.

Por outro lado, alguns dos pensionistas recipientes de pensões mínimas poderão ter de facto rendimentos explícitos muito reduzidos, mas têm um nível de vida superior aos

rendimentos e patrimónios explícitos dado viverem integrados em agregados familiares situados claramente acima da linha de pobreza.

Colocam-se igualmente **questões de equidade horizontal**. Só têm direito às pensões mínimas do regime geral pessoas que tenham sido trabalhadores beneficiários (activos). Para outros residentes idosos ou incapacitados em Portugal só há um direito à Pensão Social, do regime não contributivo, com valores mais reduzidos. Ora, na medida em que as contribuições feitas enquanto trabalhadores activos geram um direito à pensão estatutária, a atribuição de um complemento acima desse montante faria sentido tendo em conta uma presumível situação de pobreza. Nesse caso, **porquê discriminar contra outras pessoas que estão em situações semelhantes mas que não fizeram parte do regime geral?** Dado que o complemento de pensão é atribuído em função de uma presunção de necessidade, porque não tratar da mesma forma todos os igualmente necessitados? Se o problema que os complementos de pensão visam resolver é um problema de solidariedade, então não parece haver justificação para discriminar entre pessoas igualmente necessitadas dessa solidariedade. Como já referimos, o problema da iniquidade horizontal é tanto mais grave quanto o facto de a legislação da Segurança Social, estabelecer que o financiamento dos complementos de pensão se deve fazer a partir do Orçamento do Estado e não das contribuições para a Segurança Social.

Uma última fonte de problemas potenciais com a política de crescimento das pensões mínimas prende-se com a **utilização do salário mínimo como indicador de objectivos de natureza redistributiva**. De um ponto de vista técnico esta utilização carece de sentido. A identificação de quem tem um nível de vida demasiado baixo deve ser feita com base em linhas de pobreza do agregado em que cada pessoa vive, tendo em conta por um lado todas as fontes de rendimento do agregado e por outro lado um indicador das suas necessidades, o qual deverá ter em conta a sua composição demográfica. A utilização de um preço mínimo do mercado de trabalho para tais funções afigura-se-nos completamente inadequada.

Acresce que a utilização do salário mínimo como instrumento na definição das transferências sociais reduz, simultaneamente, os graus de liberdade na condução da política de Segurança Social e da política laboral. Determinadas situações poderão aconselhar um crescimento mais lento do salário mínimo numa óptica de política laboral mas um crescimento mais rápido numa óptica de política redistributiva, ou vice-versa, uma contradição inultrapassável com o sistema proposto. Por outro lado, o próprio sistema político poderá originar decisões menos acertadas: a concorrência entre partidos ou a acção de grupos de pressão poderão conduzir a modificações do salário mínimo apenas para obter objectivos na área das pensões, podendo isso prejudicar seriamente o funcionamento do mercado de trabalho para trabalhadores menos qualificados ou no início de carreira. A definição de pensões mínimas com base em linhas de pobreza oficiais (como já acontece oficialmente noutros países e implicitamente em Portugal no caso do Rendimento Mínimo Garantido) permitiria desacoplar políticas laborais de políticas de pensões.

1.3. Objectivos e Metodologia

Este trabalho visa estimar, com o rigor possível, o impacto das pensões mínimas na distribuição do rendimento e dos índices de pobreza. Os dados a utilizar para a análise serão retirados do Inquérito aos Orçamentos Familiares de 2000.

Note-se que na medida em que as pessoas vivem inseridas em agregados familiares há uma partilha conjunta do orçamento. Dada a inexistência de informação sobre a distribuição do consumo intra-agregado, temos que a única forma de estudar estatisticamente a situação de pobreza é aceitá-la como uma característica do agregado e não como uma característica onde possa existir heterogeneidade individual dentro de cada agregado. As fontes de rendimento de cada indivíduo do agregado são somadas e é o total que conta. Isso significa, por exemplo, que quando um pensionista recebe uma pensão mínima que faz o rendimento total do seu agregado passar para lá da linha de pobreza então essa pensão mínima teve um impacto não só no nível de pobreza do pensionista como no de todos os outros membros do agregado em que o pensionista está inserido.

A metodologia consistiu em definir um esquema de imputação que, para cada pessoa na base de dados e para cada agregado familiar, define se são ou não recipientes de pensões mínimas e em caso positivo quais os montantes anuais envolvidos. Uma vez resolvido o problema anterior, será calculada uma bateria de indicadores de despesas e redistributivos (despesas com pensões mínimas, índices de desigualdade, índices de pobreza, % do montante gasto em pensões mínimas que foram para agregados pobres, etc.).

Seguidamente criaremos um cenário contrafactual simulando um aumento das pensões mínimas e identificando o impacto em termos de despesas e efeitos na distribuição do rendimento e níveis de pobreza. Em todos estes casos assumimos irrealisticamente que não há reacções comportamentais por parte dos trabalhadores e pensionistas às políticas públicas. Isso significa que os efeitos por nós identificados podem ser vistos apenas como sendo uma primeira aproximação, sendo necessário explicitar e modelizar o comportamento dos agentes envolvidos para obter estimativas mais rigorosas dos efeitos das políticas públicas.

2. A Distribuição do Rendimento, a Desigualdade e a Pobreza

A fonte dos dados usados neste trabalho é o Inquérito aos Orçamentos Familiares de 2000. Ao nível dos agregados domésticos privados (ADP) usamos dados para os rendimentos totais e por fonte de rendimento, juntamente com algumas características do agregado, incluindo a sua dimensão, quer em número de pessoas quer em termos da escala de equivalência. Uma característica importante de cada ADP é o seu ponderador, o qual reflecte a taxa de amostragem usada pelo INE na construção dos dados. Estes ponderadores indicam grosso modo quantos agregados da população portuguesa são representados por cada observação (ou seja são o inverso da taxa de amostragem). Todos as estatísticas e cálculos apresentados neste trabalho usam estes ponderadores pelo que os resultados não são comparáveis com quaisquer outros resultados já publicados onde esses ponderadores tenham sido ignorados. Todos os montantes foram convertidos para Euros.

A distribuição do rendimento que melhor permite aferir das condições de vida das populações é a distribuição individual do rendimento por adulto equivalente. Para tal, torna-se necessário proceder à transformação da distribuição inicial do rendimento de forma a considerar simultaneamente a construção do rendimento equivalente e a consideração de cada indivíduo como unidade de análise.

O rendimento equivalente é obtido, dividindo o rendimento total de cada agregado pela sua dimensão em termos de "adultos equivalentes", utilizando a escala de equivalência modificada da OCDE. Esta escala atribui um peso de 1.0 ao primeiro adulto num ADP, 0.5 aos restantes adultos e 0.3 a cada criança. O rendimento equivalente permite ter em conta

as diferenças na dimensão e composição dos diferentes agregados na análise da distribuição.

De forma a passarmos da distribuição do rendimento por agregado para a correspondente distribuição individual, e assim obtermos uma mais apropriada medida do bem-estar de cada pessoa na sociedade, o rendimento equivalente do agregado é atribuído a cada indivíduo que o constitui. Os resultados apresentados utilizando a distribuição individual do rendimento por adulto equivalente são, assim, definidos em termos do número de pessoas e não de agregados.

O Quadro 1 apresenta os indicadores para a distribuição individual dos rendimentos equivalentes anuais. O rendimento individual equivalente médio é de 8851€. As distribuições individuais foram construídas combinando os ponderadores de cada agregado familiar com o número de membros do agregado.

Quadro 1

A distribuição individual do rendimento equivalente

Percentis			
5%	2833	Observações	10 020
25%	4863	Pessoas representadas	10 106 339
50%	7164	Média	8851
75%	10335	Desvio Padrão	6711
95%	21498	Variância	45038180

Fonte: cálculos dos autores usando os dados do Inquérito aos Orçamentos das Famílias de 2000 com inclusão dos ponderadores amostrais. Valores do rendimento em euros.

A partir dos dados é possível calcular indicadores que permitem quantificar os níveis de desigualdade da distribuição do rendimento bem como os níveis de pobreza.

Na abordagem da desigualdade, serão utilizadas diversas medidas de forma a ilustrar não só a assimetria existente em diferentes partes da distribuição, como os fundamentos dos diferentes juízos normativos na avaliação da desigualdade.

O índice de Gini é, provavelmente, a medida de desigualdade mais utilizada. Se representarmos por y_i o rendimento de cada indivíduo, e ordenarmos estes de forma crescente com o rendimento, o Índice de Gini pode ser calculado como:

$$G = 1 + \left(\frac{1}{n}\right) - \left(\frac{2}{(n^2\mu)}\right)\sum_{i=1}^{n}(n - i + 1)Y_i \qquad (1)$$

onde μ representa o rendimento médio da distribuição. Este índice é particularmente sensível a transferências efectuadas no meio da distribuição. O decréscimo na desigualdade resultante de uma transferência regressiva será tanto maior quanto mais perto da moda da distribuição os agentes envolvidos se encontrarem.

O índice de Atkinson é obtido através da expressão seguinte:

$$A = 1 - \left(\frac{1}{\mu}\right)\left[\left(\frac{1}{n}\right)\sum_{i=1}^{n}\left(y_i^{1-\varepsilon}\right)\right]^{\frac{1}{1-\varepsilon}} \quad (2)$$

onde ε é um parâmetro de aversão à desigualdade. A sensibilidade do Índice de Atkinson a diferentes partes da distribuição depende do valor atribuído ao parâmetro de aversão à desigualdade ε. Quanto maior ε, maior o peso atribuído aos rendimentos de menor magnitude.

Na apreciação das famílias e indivíduos em situação de pobreza colocam-se duas questões de natureza metodológica: a primeira é a identificação da linha de pobreza, isto é, do limiar de rendimento abaixo do qual uma família pode ser considerada pobre. A segunda questão prende-se com as medidas de pobreza. Embora a quantificação da proporção dos pobres existentes numa dada sociedade seja um indicador importante do seu nível de bem-estar, é igualmente relevante dispor de informação que nos permita avaliar das condições de vida da população pobre.

No que concerne à determinação da linha de pobreza, utilizaremos neste trabalho o valor correspondente a 60% do rendimento mediano como limiar de pobreza. Esta é a abordagem corrente seguida nos diferentes países da União Europeia e recomendada pelo Eurostat.

As medidas de pobreza adoptadas neste trabalho, são as medidas de Foster-Greer-Thorbecke (FGT), dadas pela expressão:

$$P_\alpha = \left(\frac{1}{n}\right)\sum_{i=1}^{q}\left(\frac{1-y_i}{z}\right)^{\alpha} \quad (3)$$

onde z é a linha de pobreza, y_i é o rendimento da unidade de observação i, n é o número de unidades de observação na amostra, q é o número de pobres e α é um parâmetro de aversão à pobreza. O valor do parâmetro α determina o tipo de índice estimado.

Se α=0 apenas o número de pobres conta, pelo que $P_\alpha = q/p$. O índice consiste no rácio entre o número de pobres e a população total, ou seja, temos a **prevalência** da pobreza ou, usando uma terminologia mais consagrada, a taxa da pobreza.

Quando α=1 o indicador de pobreza corresponde à soma, para todos os pobres, das diferenças entre os seus rendimentos e a linha de pobreza, expressa como percentagem da linha de pobreza. Trata-se de uma medida da **intensidade** da pobreza, já que maiores níveis individuais de pobreza, na forma de rendimentos mais baixos, conduzem a maiores valores para o índice agregado.

É também tradicional na literatura calcular a medida de pobreza admitindo que, a gravidade da situação de privação cresce mais do que proporcionalmente com a distância a que o rendimento está da linha de pobreza. Trata-se de uma tentativa de medir a **severidade** da pobreza. No caso particular de α=2, admite-se que a gravidade cresce com o quadrado da distância proporcional do rendimento à linha de pobreza.

O Quadro 2 apresenta os valores para os índices de Atkinson com parâmetro de aversão à desigualdade de 1 e de 2 e para o índice de Gini.

Quadro 2

Indicadores de desigualdade

Distribuição Individual do Rendimento Equivalente

Índice	Valor
Atkinson2	0.3131
	(0.0066)
Atkinson1	0.1802
	(0.0047)
Gini	0.3469
	(0.0052)

Nota: Erros Padrão das estimativas entre parêntesis

O Quadro 3 apresenta os valores estimados para os três indicadores de pobreza da família de Foster-Greer-Thorbecke (taxa de pobreza, défice de recursos relativo dos pobres (intensidade) e severidade da pobreza). Os valores baseiam-se numa linha de pobreza relativa igualmente apresentada. Como já referido, esta linha de pobreza segue a convenção usada pelo Eurostat para modificar o critério primeiramente sugerido por Fuchs (1967) construímos a linha de pobreza como sendo 60% do rendimento equivalente mediano.

Quadro 3

Indicadores de pobreza

Distribuição Individual do Rendimento Equivalente

Índice	Valor
Linha de Pobreza	4298.7 €
	(46.27)
Taxa de Pobreza: F0	0.1855
	(0.0056)
Intensidade da Pobreza: F1	0.0461
	(0.0019)
Severidade da Pobreza: F2	0.0168
	(0.0010)

Nota: Erros Padrão das estimativas entre parêntesis

De acordo com os padrões adoptados as estimativas apontam para que 18.6% da população portuguesa seja pobre, ou seja tenha um rendimento equivalente abaixo de 4299 €. As características da distribuição dos rendimentos dentro da sub-população pobre são importantes e a literatura recente sugere a utilização de medidas destas características de através dos índices de intensidade e severidade da pobreza (cf. entre muitos outros Jenkins e Lambert (1997)).

3. Desigualdade, Pobreza e Fontes de Rendimento

O conjunto dos índices estatísticas e gráficos apresentados permite quantificar o cenário de base sobre o qual faremos a simulação de Políticas Sociais e a imputação dos seus efeitos aos vários instrumentos e programas. Contudo, os tipos de políticas sociais que nos propusemos estudar requerem uma análise mais detalhada. Em particular, convém ter uma ideia precisa da decomposição dos rendimentos dos agregados e dos indivíduos pelas suas diferentes fontes. Naturalmente as fontes que iremos estudar em detalhe são as mais directamente ligadas às políticas sociais, em particular as pensões de velhice e o rendimento mínimo garantido.

O Quadro 4 apresenta alguns dados estatísticos sobre a composição do rendimento equivalente dos portugueses.[1] As fontes indicadas são a desagregação permitida pelos dados disponíveis do IOF 2000. Os rendimentos do trabalho por conta de outrem são a principal fonte de rendimento, a que se seguem as pensões de velhice e em terceiro lugar os rendimentos não monetários (autoconsumo, autolocação, etc.).

Shorrocks (1982) demonstrou que é possível obter uma decomposição única da desigualdade pelas várias fontes de rendimento. A decomposição de Shorrocks usa o facto de a variância de uma soma ser igual à soma das covariâncias dos totais com as parcelas e atribui a cada parcela, ou fonte de rendimento, uma contribuição para a desigualdade igual ao rácio da covariância respectiva sobre a variância dos totais.

O Quadro 4 apresenta detalhes sobre o papel distributivo de cada fonte de rendimento. Por exemplo, os rendimentos do trabalho por conta de outrem são 52% dos rendimentos dos agregados mas são responsáveis por 60% da desigualdade. Por outro lado, as pensões de velhice são responsáveis por 15% dos rendimentos mas 9 % da desigualdade. Isso significa que as pensões contribuem para aumentar a desigualdade, se bem que de forma mais atenuada que outras fontes de rendimento. Um caso interessante é o do rendimento mínimo garantido, o qual apresenta uma covariância negativa com o rendimento total. Isso significa que é a única fonte de rendimento que explicitamente reduz a desigualdade, embora o seu efeito seja pequeno (-0,2 %).

[1] Tal como no caso do rendimento total, o rendimento para cada fonte de todos os indivíduos num agregado foi somado e esta soma dividida pela escala de equivalência para se obterem os rendimentos equivalentes de cada pessoa em cada agregado familiar.

Quadro 4

Contribuição de cada fonte de rendimento para a desigualdade

	Percentagem do rendimento global	Covariâncias	Contribuição para a Desigualdade
Rendimentos Trabalho CO	51.5	60.0%	Acima da média
Rendimentos Trabalho CP	10.1	10.8%	Acima da média
Outros Rendimentos Privados	4.9	9.1%	Acima da média
Pensões Velhice	15.0	8.9%	Abaixo da média
Outras Pensões	2.2	0.3%	Abaixo da média
Benefícios Sociais	2.6	0.5%	Abaixo da média
Rendimento Mínimo Garantido	0.2	-0.2%	Reduz a desigualdade
Rendimentos Não Monetários	13.5	10.6%	Abaixo da média

Nota: Covariâncias expressas como percentagem da variância do rendimento equivalente.

Ao contrário do que se passa para a desigualdade, para a pobreza não há uma metodologia consagrada para medir a contribuição de cada fonte de rendimento. Por essa razão prosseguimos a análise seguindo uma metodologia própria. Admitindo que queremos fixar uma linha de pobreza que sirva de referencial único para múltiplos cenários, não interessa apenas a dispersão relativa dos rendimentos mas igualmente o seu montante absoluto. Adoptamos então um princípio de identificar o contributo de cada fonte criando distribuições hipotéticas com a mesma média mas composições diferentes (inspirando-nos na ideia que a literatura designa como “mean preserving spread”).

Assim, para identificar o contributo para a pobreza de cada fonte de rendimento criamos uma distribuição do rendimento artificial em que essa fonte de rendimento foi anulada, mas as restantes fontes de rendimento são aumentadas de forma proporcionalmente homogénea de forma a manter a média da distribuição. A diferença entre o indicador de pobreza no cenário de base e no cenário artificial mede a contribuição da fonte de rendimento para a pobreza. Uma diferença negativa significa que a fonte reduz a pobreza. Os resultados encontram-se no Quadro 5.

Quadro 5

Contribuição de cada fonte de rendimento para a pobreza

	F0	F1	F2
Rendimento Global	**0.1855**	**0.0461**	**0.0168**
Sem Rendimentos Trabalho CO	0.3239	0.1626	0.1095
Sem Rendimentos Trabalho CP	0.2253	0.0739	0.0371
Sem Outros Rends. Privados	0.1810	0.0466	0.0185
Sem Pensões Velhice	0.2664	0.1293	0.0880
Sem Outras Pensões	0.2019	0.0568	0.0244
Sem Benefícios Sociais	0.1958	0.0531	0.0219
Sem Rendim. Mínimo Garantido	0.1867	0.0499	0.0203
Sem Rendims. Não Monetários	0.2112	0.0579	0.0231

Todas as fontes de rendimento contribuem para diminuir os índices de pobreza, com a única excepção dos "outros rendimentos privados". A introdução dos "outros rendimentos privados" aumenta ligeiramente a taxa de pobreza mas diminui os outros dois índices de pobreza.

Uma interpretação possível destes resultados é que cada fonte de rendimento tem um conjunto de agregados que dele depende especialmente. A eliminação de cada fonte (mesmo que compensada ao nível agregado) aumenta a pobreza mediante o empobrecimento dos grupos particularmente associados a tal fonte de rendimentos.

Quadro 6

Pessoas com rendimentos predominantemente de uma só fonte

Fonte de rendimento	% de todas as pessoas	% das pessoas com rend. dessa fonte
Rendimentos Trabalho CO	23.1	32.4
Rendimentos Trabalho CP	2.1	8.0
Outros Rendimentos Privados	0.3	0.7
Pensões Velhice	7.0	18.3
Outras Pensões	0.4	2.7
Benefícios Sociais	0.3	0.5
Rendimento Mínimo Garantido	0.1	2.4
Rendimentos Não Monetários	0.0	0.0

Nota: Considera-se uma fonte de rendimento como predominante se incluir mais de 80% do rendimento global.

O Quadro 6 é parcialmente compatível com essa interpretação. No total, 33 % das pessoas baseia mais de 80% dos seus rendimentos numa só fonte. Mesmo assim, o caso das pensões de velhice é surpreendente. Só 7% da população depende em mais de 80% dos rendimentos das pensões de velhice. Dentro do conjunto de pessoas que pertencem a

agregados recebendo pensões de velhice só 18% depende destas para mais de 80% do rendimento. Isso significa que os agregados recebendo pensões de velhice têm fontes de rendimento muito mais diversificadas do que geralmente se crê na sociedade portuguesa. Na próxima secção estudamos em mais detalhe a relação entre as pensões mínimas de velhice e os rendimentos totais dos agregados que as recebem.

4. As Características Distributivas das Pensões Mínimas

O Quadro 7 contém os valores dos principais parâmetros de política social relevantes para o problema em estudo.

Quadro 7

Pensões e salários mínimos

	Pensão Social	Pensão Mínima	Salário Mínimo
1999	117.7	162.6	305.8
2000	124.7	169.6	318.2
2001	134.0	179.6	334.2

Pensão social em 2001 para <70 anos. Pensões mínimas para 14 anos de contribuição ou menos.

Os dados do IOF incluem apenas uma variável para pensões de velhice ou reforma pelo que não temos nos dados uma separação entre pensões do regime contributivo e do não do contributivo. Em particular não temos uma identificação explícita de quem recebe pensões mínimas do regime geral ou de quem recebe pensões sociais. No entanto é possível ultrapassar este problema quando se tem em conta os dados. Para isso trabalhamos com os dados individuais para todas as pessoas com rendimentos incluídas no IOF 2000.

O Quadro 8 apresenta as principais estatísticas da distribuição das pensões de velhice pelos agregados familiares portugueses. A maioria dos agregados não recebe pensões, pelo que se trata de uma distribuição com muitos valores nulos.

Quadro 8

Distribuição das pensões de velhice por agregado

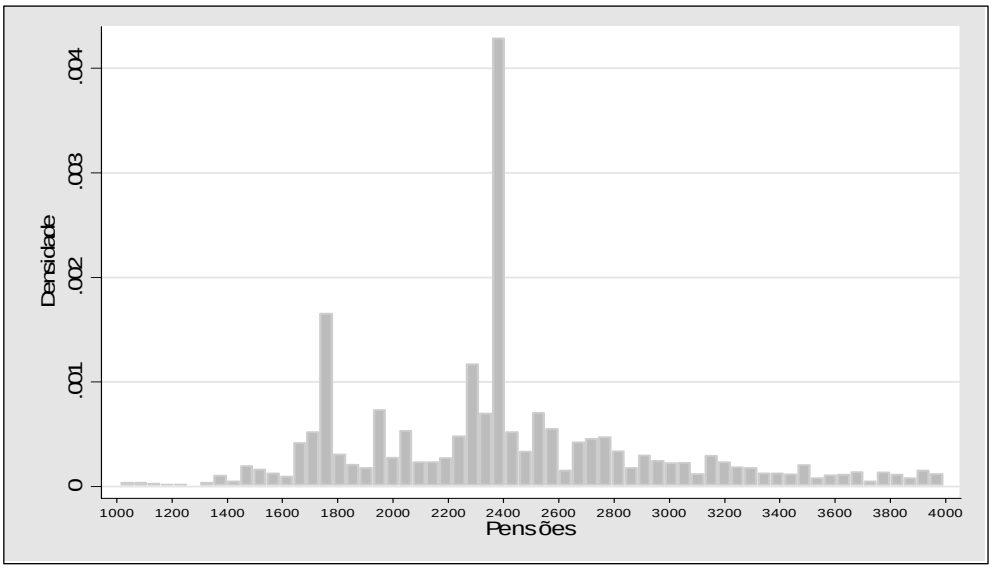
Percentis			
5%	0	Observações	10020
25%	0	Agregados representados	3 599 272
50%	0	Média	2719.2
75%	3491.6	Desvio Padrão	5116.6
95%	12278.5	Variância	26179596

A Figura 1 apresenta uma parte da distribuição dos montantes anuais das pensões de velhice (excluem-se observações com montantes nulos e algumas observações com montantes relativamente elevados a fim de permitir a representação mais detalhada). Na

Figura 1 destacam-se claramente dois picos. O primeiro pico inclui observações muito perto do montante 1745.79 €, o valor anual da pensão social para 2000 (recebida 14 vezes por ano). O segundo pico, o mais elevado, inclui observações muito perto do valor 2374.28 €, o valor anual da pensão mínima mais baixa para 2000.

Figura 1

Distribuição individual das pensões de velhice



Os dados reflectem claramente a prevalência elevada de pessoas recebendo os montantes associados aos dois tipos de pensões. Não é óbvio como interpretar as pensões intermédias entre os dois valores, mas esse conjunto pode corresponder a um número elevado de situações.[2]

Para efeitos da análise que se segue faremos uma hipótese que se nos afigura como razoável: consideraremos que todas as pensões entre 2300 € e 2600 € anuais são pensões mínimas. Estes valores excluem pensões sociais e incluem as pensões mínimas mais elevadas atribuídas a quem tenha maiores histórias contributivas. Isso permite-nos identificar as pessoas que recebem as pensões mínimas e, após alguma manipulação dos dados, identificar os agregados onde há pelo menos uma pessoa a receber uma pensão mínima.

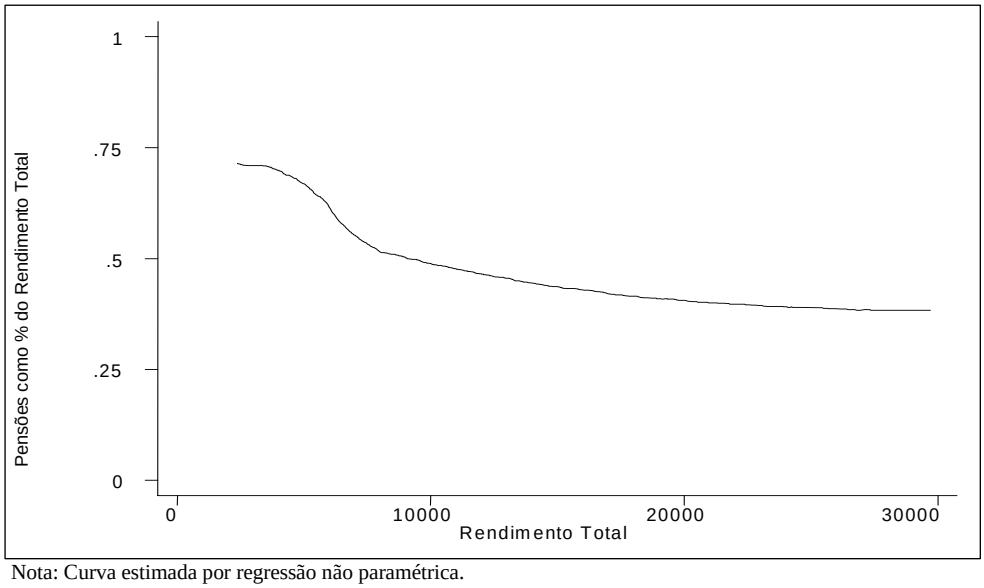
Qual é a natureza da relação entre as pensões de velhice e os rendimentos totais dos agregados que as recebem? Para o caso dos agregados recipientes de pensões mínimas,

[2] Para além do caso óbvio de informação inexacta propositadamente ou por "recall bias" temos a considerar casos de pessoas que não receberam a mesma pensão durante todo o ano ou cuja pensão seja determinada em condições especiais (RESSA, etc.).

a Figura 2 ilustra a relação entre as pensões de velhice e o rendimento total existente através do gráfico da estimativa não paramétrica do "share" das pensões de velhice no rendimento total condicional a cada nível de rendimento. Quanto maior o rendimento menor a quota das pensões. É possível ver que a partir de rendimentos anuais superiores a dez mil euros, a quota é em média inferior a 50%.

Figura 2

Pensões e rendimentos totais dos agregados com pensões mínimas



Nota: Curva estimada por regressão não paramétrica.

Uma análise adicional dos dados revela que para 41,6% dos agregados recebendo pensões mínimas, as pensões de velhice (não só as mínimas como possivelmente outras pensões de velhice) constituem menos de metade dos rendimentos totais do agregado. Para reforçar este último dado apresentamos no Quadro 9 a composição média dos rendimentos totais para os quartis da distribuição do rendimento total em todos os agregados que recebem pensões mínimas.

Quadro 9

Composição dos rendimentos dos agregados com pensões mínimas

	Rendimento Total Médio	Pensões de Velhice Médias	Composição em %
1º Quartil	3755.7	2632.5	70.1
2º Quartil	6041.6	3934.3	65.1
3ª Quartil	9317.9	4631.0	49.7
4º Quartil	19893.8	6351.0	31.9

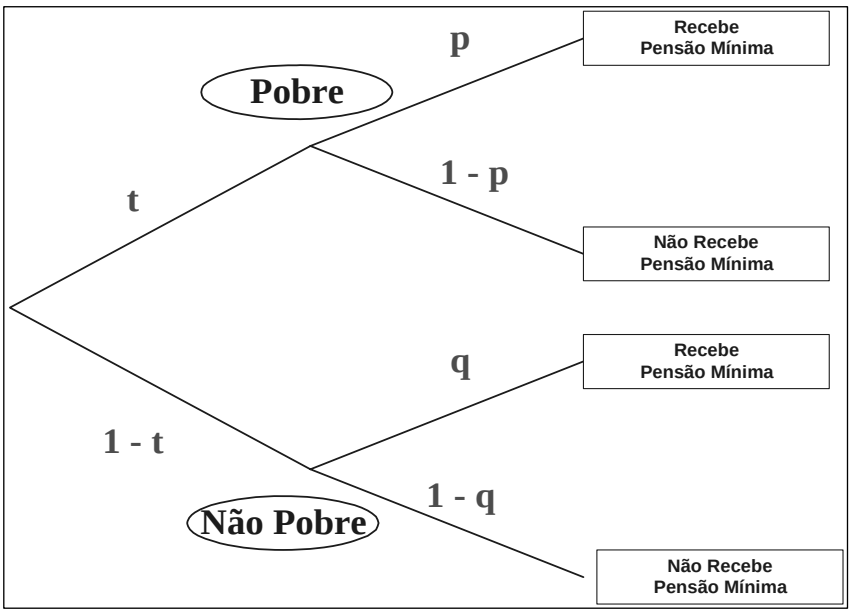
Em média mesmo nos 25% de agregados com rendimentos totais mais baixos cerca de 30% do rendimento é extra pensões de velhice.

O aspecto essencial a reter é que os agregados que recebem pensões mínimas recebem igualmente rendimentos significativos de outras fontes e que este facto se verifica para os diferentes níveis de rendimento, embora de forma não homogénea. Este resultado contraria a percepção comum e muitas vezes implicitamente usada como fundamentação para decisões em Política Social de as pensões serem quase exclusivamente a única fonte de rendimento dos agregados que as recebem. Essa situação está longe da verdade para as pensões de velhice em geral e para as pensões mínimas em particular.

A informação assim produzida é da maior relevância para podermos caracterizar que tipos de agregados beneficiam das pensões mínimas, em particular para cruzarmos o estatuto de recipiente de pensão mínima com ser ou não pobre. Um diagrama elementar de probabilidades condicionais poderá ser útil para organizar conceitos e é apresentado na Figura 3.

Figura 3

Probabilidade de ser pobre e receber uma pensão mínima



O diagrama permite precisar as ideias ainda que a um nível abstracto. "t" é a taxa de pobreza (proporção da população com rendimento equivalente abaixo da linha de pobreza), "p" é proporção dos pessoas pobres que beneficiam de pensões mínimas e "q" é a proporção dos pessoas não pobres que beneficiam das mesmas pensões.

Vejamos agora as estimativas para estas probabilidades baseadas nos dados do IOF 2000, apresentadas no Quadro 10. O Quadro mostra que 18.55% ("t") das pessoas são pobres e que 11,04% pertencem a agregados que recebem pensões mínimas. Dentro do conjunto das pessoas pobres 18.6% ("p") beneficiam de pensões mínimas e destes dados resulta que 9.3% ("q") das pessoas não pobres beneficiam de pensões mínimas.

Quadro 10

Pobreza e pensões mínimas

População em Geral	
Observações no IOF 2000	10020
Pessoas Representadas	10 106 339
Proporção Pobre	0.1855
Proporção com Pensão Mínima	0.1104
Rendimento Equivalente Médio	8851.3
Sub-População Pobre	
Observações no IOF 2000	2860
Pessoas Representadas	1 874 836
Proporção Com Pensão Mínima	0.1860
Rendimento Equivalente	3231.0
Sub-População com Pensão Mínima	
Observações no IOF 2000	1558
Pessoas Representadas	1 115 978
Proporção Pobre	0.3125
Rendimento Equivalente Médio	6193.9

Os dados mais importantes para o nosso propósito são as probabilidades condicionais a beneficiar de uma pensão mínima. Com efeito só 31.25% das pessoas que beneficiam de pensões mínimas são pobres. Isso significa que é difícil justificar uma política de aumento das pensões mínimas com base em razões de solidariedade social já que mais de 68% das pessoas beneficiadas directamente por tais políticas não são pobres.

Os valores acima reflectem a situação existente, ou seja já incluindo as pensões mínimas, pelo que é possível argumentar que se os agregados que recebem as pensões mínimas recebessem pensões mais pequenas então a) uma percentagem muito maior estaria em situação de pobreza e b) a percentagem de pobres que recebe pensão mínima seria maior. Para estudar esta questão simulámos a distribuição do rendimento dos agregados assumindo que todos os pensionistas anteriormente identificados como recebendo pensões mínimas estariam agora a receber apenas a pensão social. Os resultados podem ser vistos no Quadro 11.

Quadro 11

Experiência hipotética: reduzir as pensões mínimas para o nível da pensão social

População Geral	
Observações no IOF 2000	10020
Pessoas Representadas	10 106 339
Pobreza Simulada na População Geral	0.1949
Sub-População com Pensão Mínima	
Observações no IOF 2000	1558
Pessoas Representadas	1 115 978
Pobreza Simulada na sub-População com Pensão Mínima	0.3913
Sub-População Pobre Simulada	
Observações no IOF 2000	3028
Pessoas Representadas	1 969 394
Proporção com Pensão Mínima na sub-População pobre simulada	0.2217

Os resultados mostram que a redução das pensões simulada teria um efeito pequeno sobre a taxa de pobreza, que passaria de 18.55% para 19.49%, e que mesmo assim só 39% das pessoas em agregados com estas pensões mínimas seriam pobres. Por outro lado é interessante notar que a hipotética redução da pensão mínima simulada levaria a uma redução das despesas com pensões de velhice de 390 milhões de Euros. Esses recursos hipotéticos poderiam levar a poupanças avultadas para o Orçamento de Estado ou alternativamente permitiriam aumentar os montantes gastos noutras áreas da política social como o Rendimento Mínimo Garantido, as pensões sociais, os programas da Acção Social, etc.

5. Efeitos Redistributivos da Convergência das Pensões Mínimas.

A política de convergência das pensões mínimas com o salário mínimo pode ser agora avaliada em termos das suas consequências redistributivas. Em 2000 o valor das pensões mínimas mais baixas era de 60% do salário mínimo (líquido das contribuições do trabalhador para a Segurança Social). Nesta secção simulamos as implicações de aumentar as pensões mínimas (que identificámos com pensões entre os 2300 e os 2600 € anuais) para 70% do salário mínimo líquido, i.e. **2775.63** € anuais. Naturalmente pensões entre 2600€ e **2775.63** € são também aumentadas.

A inferência feita a partir da amostra dos IOF 2000 diz-nos que a despesa em pensões acarretada pelo aumento simulado da pensão mínima seria de 226.7 milhões de Euros. Põe-se a questão de saber como financiar esta despesa. Uma hipótese seria definir um padrão de distribuição desta carga (por exemplo reduzindo proporcionalmente os rendimentos líquidos do trabalho por conta própria e por conta de outrem). No entanto, a escolha seria sempre muito subjectiva pelo que ignoraremos as questões de equilíbrio geral e orçamental do financiamento da despesa.

O efeito do aumento das pensões mínimas na desigualdade da distribuição individual do rendimento equivalente pode ser visto no Quadro 12. Há um redução da desigualdade entre 0.6 e 1.2%, dependendo do indicador usado. Quanto aos indicadores de pobreza, a redução é entre 3.8 e 4.8%.

Quadro 12

Aumento das pensões mínimas: efeitos nos indicadores de desigualdade e pobreza

Índice	Cenário de Base	Aumento das Pensões Mínimas	Variação em %
Atkinson2	0.3131	0.3095	-1.15
	(0.0066)	(0.0067)	
Atkinson1	0.1802	0.1780	-1.20
	(0.0047)	(0.0047)	
Gini	0.3469	0.3447	-0.63
	(0.0052)	(0.0052)	
Taxa de Pobreza: F0	0.1855	0.1785	-3.77
	(0.0056)	(0.0059)	
Intensidade da Pobreza: F1	0.0461	0.0439	-4.67
	(0.0019)	(0.0018)	
Severidade da Pobreza: F2	0.0168	0.0160	-4.75
	(0.0010)	(0.0009)	

Nota: Erros Padrão das estimativas entre parêntesis

De certa maneira podemos dizer que as mudanças apresentadas no Quadro 12 se conseguem à custa de 226.7 milhões de euros, ou seja de sobrecarregar outras partes da economia não explicitamente consideradas na análise. Podemos formular duas questões para avaliar a eficácia da estratégia seguida. A primeira é: **que outros resultados seria possível obter gastando os mesmos 226.7 milhões de euros**? A segunda é: **será possível obter a mesma redução dos níveis de pobreza gastando menos dinheiro?**

Não vamos apresentar neste trabalho uma resposta directa à primeira questão, mas vamos apresentar informação suficiente para responder à segunda. Uma parte das pensões mínimas acrescidas vai para agregados que eram pobres antes de tal medida e que continuam a ser pobres depois. Uma outra parte vai para agregados que já não eram pobres antes e que o são ainda menos depois do aumento das pensões. Uma terceira parte dos recursos gastos no aumento das pensões mínimas consiste em transferências feitas para agregados que eram pobres antes do aumento mas que depois ficam com rendimentos acima da linha da pobreza. No caso deste último grupo de agregados pode-se decompor a despesa feita em duas fatias. A primeira faz subir o rendimento desses agregados até este igualar a linha de pobreza. A segunda fatia faz os rendimentos subir acima da linha de pobreza.

O Quadro 13 decompõe a despesa total gerada pelo aumento simulado das pensões de acordo com os efeitos descritos no parágrafo anterior. Os resultados constantes do Quadro 13 mostram uma ineficácia elevada da política de aumento das pensões mínimas no

que diz respeito ao alívio de situações de pobreza: mais de dois terços dos recursos gastos são desperdiçados na medida em que não dão qualquer contribuição para a redução dos níveis de pobreza.

Quadro 13

Decomposição da despesa gerada pelo aumento das pensões mínimas

Despesa	Milhões €	%
Montante total gasto	226.7	100.0
1. Despesas em agregados sempre abaixo da linha de pobreza	62.4	27.5
2. Despesas em agregados que deixam de ser pobres, parte até rendimentos igualarem linha de pobreza	10.5	4.6
1+2. Despesa total contribuindo para redução da pobreza	***72.8***	***32.1***
3. Despesas em agregados que deixam de ser pobres, parte gerando rendimentos acima da linha de pobreza	9,4	4.2
4. Montante gasto em agregados sempre acima da linha de pobreza.	144.4	63.7
3+4. Despesa total não contribuindo para a redução da pobreza	***153.8***	***67.9***

Seguindo a metodologia preconizada por Beckerman (1979, 1981) este aumento das pensões mínimas apresenta uma Eficiência Vertical de cerca de 36% enquanto que a sua Eficiência na Redução da Pobreza se limita a 32.1%[3]. Estes valores revelam-se extremamente baixos quando comparados com os obtidos através de outros programas sociais, nomeadamente os associados à existência de uma condição de recursos. Por exemplo, na avaliação da eficiência do RMIG em Portugal, Rodrigues (2001) encontrou valores de 92% e de 89% respectivamente.

A conclusão anterior sobre a ineficácia da política das pensões mínimas pode ser criticada porque só um programa com um "targetting" perfeito tem desperdícios nulos, i.e. um programa onde só pessoas de facto pobres beneficiam dos recursos gastos[4]. A crítica seria a de que nenhum programa no mundo real poderia ser perfeito pelo que haverá sempre falhas. A questão pode então pôr-se da seguinte forma: será que é possível ter programas operacionais que tenham menos desperdício de recursos que as política das pensões mínimas?

A resposta é sim. Com base no IOF de 2000 é possível simular um programa arbitrário e simplista, mas que goza de uma eficácia muito superior à da política de

[3] De acordo com Beckerman a Eficiência Vertical do Programa representa a Proporção das Transferências recebidas pelas famílias inicialmente pobres enquanto que a Eficiência na Redução da Pobreza mede a redução do défice de recursos dos agregados em situações de pobreza.

[4] Note-se que nesta argumentação ignoramos quer os custos administrativos quer os custos da ineficiência económica atribuíveis aos incentivos gerados pelos programas de transferências.

aumento das pensões mínimas[5]. O programa simulado, a que chamaremos *Rmig Modificado*, consiste em pura e simplesmente dar mais dinheiro aos agregados que já recebem o Rendimento Mínimo Garantido através de um aumento proporcional de 90%[6] desta prestação para todos estes agregados. Os resultados da simulação do programa Rmig Modificado podem ver-se no Quadro 14.

Os resultados obtidos não são exactamente iguais aos do aumento das pensões mínimas sendo ligeiramente melhores numas dimensões da desigualdade e da pobreza e ligeiramente piores noutras. No cômputo geral não é muito forçado muito dizer que o programa *Rmig Modificado* é equivalente ou mesmo superior ao aumento de pensões no que diz respeito à equidade da distribuição final do rendimento e isto apesar de custar menos de metade do aumento das pensões mínimas.

Quadro 14

Despesa e efeitos nos indicadores de desigualdade e pobreza do *RMIG modificado*

Indicador	Aumento da Pensão Mínima	*Rmig Modificado*	Diferença (%)
Despesa	226.7 M €	110.76 M €	-51.13
Atkinson2	0.3095	0.3073	-0.7
Atkinson1	0.1780	0.1775	-0.3
Gini	0.3447	0.3447	0.0
Taxa de Pobreza: F0	0.1785	0.1804	1.0
Intensidade da Pobreza: F1	0.0439	0.0429	-2.3
Severidade da Pobreza: F2	0.0160	0.0153	-4.3

A razão para a superioridade do programa *Rmig Modificado* está obviamente no seu targeting. Apesar de imperfeito, os resultados do targeting para este programa, correspondentes ao Quadro 13 para o caso das pensões mínimas, são que 68.2% dos recursos contribuem para a redução da pobreza. Apesar deste programa ainda apresentar níveis de desperdício de 31.8%, o seu nível de eficácia é mais do dobro da eficácia do aumento das pensões mínimas!

6. Conclusões

Os resultados obtidos confirmam que as pensões mínimas são instrumentos caros e ineficazes de solidariedade. Só 31.25% das pessoas que vivem em agregados familiares

5 Pode interpretar-se o argumento que se segue como uma espécie de raciocínio por absurdo já que um programa arbitrário, sem bons critérios, é mesmo assim muito melhor que a política de aumento das pensões mínimas. Dar mais dinheiro a quem já recebe o rendimento mínimo é muito pior do ponto de vista de *targetting* do que aumentar o nível do rendimento garantido pelo programa. Por exemplo, o aumento do rendimento garantido implicaria a entrada de novas famílias para o programa, o que não acontece por definição no programa arbitrário que usamos como exemplo.

6 Valor encontrado por tentativa e erro.

recebendo pensões mínimas são pobres. Isso pode ser explicado porque estes agregados têm outros rendimentos. Para 41,6% dos agregados recebendo pensões mínimas as pensões de velhice constituem menos de metade dos rendimentos totais do agregado. Por comparação mais de 90% das pessoas em agregados recebendo o Rendimento Mínimo Garantido (RMIG) são pobres (veja-se Rodrigues (2001)).

Por outro lado um aumento do esforço de convergência das pensões para o salário mínimo de 60% para 70% produziria resultados diminutos em termos de equidade. Os resultados obtidos mostram que os indicadores de pobreza se reduziriam menos de 5% apesar da despesa elevada que tal política acarreta. Tais resultados devem-se ao facto de 67.9% das despesas geradas pelo acréscimo das pensões mínimas serem feitas em pessoas não pobres. Em contrapartida, um programa simulado com base numa condições de recursos muito grosseira consegue apesar disso obter melhores resultados em termos de pobreza e equidade e gastando menos de metade dos recursos.

O acréscimo das pensões mínimas com objectivos redistributivos, nos moldes em que existem actualmente, afiguram-se-nos como uma utilização ineficiente de avultados recursos públicos. A Segurança Social, na sua componente de solidariedade, deverá adoptar uma nova lógica e abandonar os objectivos contraproducentes de convergência para o salário mínimo, inserindo condições de recursos nos programas com objectivos redistributivos.

Referências

Atkinson, A. B. (1998), Poverty in Europe, Blackwells Publishers, Oxford.

Atkinson, A. B. (1987), "On the Measurement of Poverty", *Econometrica*, vol. 55, nº 4, pp 749-764.

Atkinson, A. B. (1983), The Economics of Inequality (2nd ed.).

Atkinson, A. B. (1970), "On the Measurement of Inequality", *Journal of Economic Theory*, 244-63.

Atkinson, A.B. and Bourguignon, F. (eds.) (1999), *Handbook of Income Distribution*. Amsterdam, Netherlands.

Beckerman, W. (1981), "The Impact of Income Maintenance Programmes on Poverty in Canada", *World Employment Programme Research Working Paper* 2-23/98, International Labour Office, Geneva.

Beckerman, W. (1979), "The Impact of Income Maintenance Programmes on Poverty in Four Developed Countries", International Labour Office, Geneva.

Besley, T. (1990), "Means Testing versus Universal Provision in Poverty Alleviation Programmes," *Economica*, 57:119-129.

Cowell, F, (1999), Measurement of Inequality, in Atkinson, A.B. and Bourguignon, F. (eds.), *Handbook of Income Distribution*. Amsterdam, Netherlands.

Cowell, F.(1994), Measuring Inequality, LSE Handbooks in Economics, London (2nd ed.).

Foster, J., Greer, J. and Thorbecke, E. (1984), "A Class of Decomposable Poverty Measures", *Econometrica*; 52(3), 761-66.

Fuchs, V. (1967), "Redefining Poverty and Redistributing Income", *The Public Interest*, 8, 88-95.

Gouveia, M. and Rodrigues, C.F. (2002), "The Impact of a Guaranteed Minimum Income Program in Portugal", *Public Finance and Management*, II(2), 2002.

Gouveia, M. and Tavares, J. (1995), "The Distribution of Household Income and Expenditure in Portugal: 1980 and 1990", *Review of Income and Wealth*, 41 (1), 1-17.

Instituto Nacional de Estatística. (2003), *Estatísticas da Protecção Social* 2000, INE, Lisboa.

Instituto Nacional de Estatística. (2002), *Inquérito aos Orçamentos Familiares* 2000 – Principais Resultados. INE, Lisboa: Série Estudos.

Jenkins, S. e Lambert, P. (1997), "Poverty Orderings, Poverty Gaps, and Poverty lines", *Oxford Economic Papers*, vol. 49, pp 93-107.

Rodrigues, Carla (2003), "Inequality and Poverty in retirement age groups: an analysis for Portugal" DGEP/Ministério das Finanças – *Working Paper* nº 32

Rodrigues, C.F. (2001), "Effectiveness and efficiency of the Guaranteed Minimum Income Programme in Portugal", *ISEG/DE Working Paper nº WP8/2001/DE/CISEP*.

Rodrigues, C.F. (1999), "An Assessment of Income Distribution and Poverty in Portugal Using Different Sources of Income", *Eurostat - Studies and Research, Proceedings of the Seventh Seminar on Income Distribution and different sources of Income, Eurostat, Luxembourg.*

Rodrigues, C.F. (1993), The Measurement and Decomposition of Inequality in Portugal [1980/81 - 1989/90]", Microsimulation Unit *Discussion Paper* MU9302, Cambridge, Department of Applied Economics.

Shorrocks, A.F. (1983), "The impact of Income Components on the Distribution of Family Incomes", *Quarterly Journal of Economics*, nº 98, pp 311-326.

CICLOS POLÍTICO-ECONÓMICOS NOS MUNICÍPIOS PORTUGUESES*

Linda Gonçalves Veiga
Núcleo de Investigação em Políticas Económicas
Universidade do Minho
P-4710-057 Braga – Portugal
Tel.: +351-253-604568
Fax: +351-253-676375
E-mail: linda@eeg.uminho.pt

Francisco José Veiga
Núcleo de Investigação em Políticas Económicas
Universidade do Minho
P-4710-057 Braga – Portugal
Tel.: +351-253-604534
Fax: +351-253-676375
E-mail: fjveiga@eeg.uminho.pt

Resumo:

O presente artigo testa a existência de ciclos político-económicos no âmbito da política orçamental dos municípios de Portugal Continental, durante o período 1979-2000. Os resultados empíricos revelam claramente o comportamento eleitoralista dos autarcas que, em anos de eleições, aumentam os défices e as despesas municipais, com destaque para rubricas altamente visíveis para o eleitorado. Na medida em que o eleitoralismo dos autarcas gera ineficiências na alocação dos recursos e pode, ocasionalmente, dificultar o cumprimento das regras estipuladas no Pacto de Estabilidade e Crescimento, consideramos ser benéfica a imposição de regras que limitem a gestão discricionária das finanças locais.

* Estamos gratos à Fundação para a Ciência e a Tecnologia pelo financiamento concedido ao projecto"Eleitores, governantes e a economia" (POCTI/2001/ECO/37457) e aos assistentes de investigação César Coelho, Cláudia Ribeiro e Liliana Veiga pelo seu eficiente trabalho.

1. Introdução

Durante muito tempo, as análises da política orçamental foram essencialmente normativas, sendo os governantes encarados como ditadores benevolentes que visavam apenas o bem-estar social. Com o aparecimento da chamada Escola das "escolhas públicas",[1] as abordagens tornaram-se fundamentalmente positivas e passou a ser atribuída maior atenção à possibilidade dos interesses pessoais dos responsáveis pela política económica influenciarem as medidas adoptadas. Recentemente, o estudo da influência de factores políticos na condução da política orçamental tem sido muito profícuo, gerando um elevado número de publicações nesta área.[2]

O presente artigo insere-se nesta linha de investigação ao testar a existência de ciclos político-económicos no âmbito da política orçamental dos municípios Portugueses durante o período compreendido entre 1979 e 2000. A motivação para o trabalho resultou, fundamentalmente, de quatro ordens de razões:

1º. Trata-se de uma linha de investigação muito activa em termos internacionais, sobretudo no que diz respeito à análise do comportamento dos governos centrais, estando o poder local ainda relativamente pouco estudado;

2º. O número de trabalhos científicos a debruçar-se sobre a realidade portuguesa é extremamente escasso;

3º. Na medida em que os ciclos eleitorais geram ineficiências na alocação dos recursos económicos, a comprovação da sua existência sugere a adopção de medidas que diminuam o poder discricionário dos governantes;

4º. No actual contexto da União Europeia, os países membros estão sujeitos a regras na condução da política orçamental, pelo que o estudo da necessidade de disciplinar os autarcas nesta matéria adquire ainda maior relevância.

A investigação realizada revela a existência de um comportamento oportunista[3] dos autarcas que se traduz na tentativa de mostrarem maior competência nas vésperas das eleições, através de um aumento das despesas, particularmente em categorias altamente visíveis pelo eleitorado, como as despesas de investimento em viadutos, arruamentos e obras complementares e viação rural. Os testes econométricos demonstram ainda que a dimensão do ciclo oportunista não depende do facto do partido do autarca dispor de maioria na Assembleia Municipal nem de ele vir a ser candidato nas eleições seguintes. No entanto, a magnitude do ciclo parece depender da ideologia, nomeadamente autarcas de esquerda são em média mais eleitoralistas que os de direita. É ainda possível comprovar que os défices tendem a aumentar durante os períodos pré-eleitorais.

Tendo em conta que a tendência para elevar a despesa e o endividamento pouco antes das eleições autárquicas poderá, ocasionalmente, dificultar o respeito do Pacto de

---

1 Mueller (1997) efectua uma digressão sobre esta escola de pensamento.

2 Para uma uma extensa revisão bibliográfica sobre a economia política das finanças públicas veja-se Persson e Tabellini (2002).

3 A expressão "oportunista" está consagrada na literatura de "*Political Economics*". Neste contexto ela significa eleitoralismo, ou seja, a adopção de medidas de política económica tendo em vista o êxito eleitoral.

Estabilidade e Crescimento e que a instabilidade nas despesas de investimento acarreta ineficiências na alocação dos recursos, entendemos ser benéfica a imposição de regras à gestão das finanças públicas locais que dificultem a gestão eleitoralista das mesmas pelos autarcas.

O artigo está organizado da seguinte forma. Após a introdução é apresentada a revisão bibliográfica sobre ciclos político-económicos. É depois feita uma digressão sobre os municípios portugueses e são descritas as fontes de dados estatísticos. Segue-se a apresentação da metodologia adoptada e dos resultados empíricos. De seguida, é discutido o impacto macroeconómico dos ciclos político-económicos locais. Por fim, são expostas as principais conclusões e as implicações de política económica.

2. Ciclos Político-Económicos e o Poder Local

A literatura sobre os ciclos político-económicos[4] procura analisar de que forma a ideologia dos governantes e o seu desejo de serem reeleitos influenciam a evolução da economia. Nordhaus (1975) apresentou um dos primeiros modelos sobre esta temática, com o modelo dos políticos oportunistas, que estimulam a economia na véspera dos escrutínios de forma a aumentarem a sua probabilidade de reeleição e, após as eleições, eliminam a inflação que resulta deste comportamento através de uma recessão . Este modelo admite expectativas adaptativas, o que permite aos governantes escolherem a sua combinação preferida de inflação e desemprego ao longo da curva de Phillips e enganarem sistematicamente o eleitorado antes das eleições.

A revolução das expectativas racionais exigiu a reformulação do modelo de Nordhaus, tendo surgido diversos trabalhos que procuram reconciliar a hipótese de políticos oportunistas com eleitores racionais. Estes trabalhos assumem que embora todos os governos sejam oportunistas, eles diferem no seu nível de competência que conhecem antes do eleitorado. Os eleitores apenas conseguem determinar a competência dos governos depois de observarem os resultados económicos das políticas por estes implementadas. Desta forma, na véspera das eleições, governantes oportunistas tiram partido dessa assimetria de informação procurando sinalizar um elevado grau de competência ao eleitorado, dando origem ao ciclo político-económico.

Entre os modelos racionais de oportunismo destacam-se os trabalhos de Rogoff e Sibert (1988) e Rogoff (1990) segundo os quais, sendo a política orçamental uma temática complexa para a maioria dos eleitores, é de esperar que ela seja utilizada pelos governantes para sinalizarem a sua competência. Em Rogoff e Sibert (1988) é assumido que o governo deve providenciar um montante fixo e observável de bens públicos (G), sendo a sua competência avaliada pelo montante de impostos que tem de cobrar para fornecer G. A restrição orçamental é definida como:

$$G = \varepsilon_t + \tau_t + \Delta_t \tag{1}$$

onde ε_t representa a competência do governo, τ_t o nível de impostos não distorcionários e Δ_t o montante de impostos distorcionários (receitas de senhoriagem). No modelo, a utilidade de cada eleitor depende do seu rendimento (y – constante), dos impostos cobrados, da perda

[4] Alesina, Cohen e Roubini (1997) apresentam uma extensa revisão bibliográfica e testes empíricos para uma amostra de países industrializados. Ver também Persson e Tabellini (2002).

resultante dos impostos distorcionários [$L(\Delta_t)$] e de um elemento não pecuniário que reflecte a preferência pelo partido (η_t). Sendo todos os eleitores idênticos, a função de bem-estar social assume a seguinte forma:

$$\Omega_t = y - \tau_t - \Delta_t - L(\Delta_t) + \eta_t \quad (2)$$

Esta função utilidade torna claro que, para tudo o resto igual, os eleitores preferem o partido com maior nível de competência e, na ausência de assimetria de informação, qualquer governo, independentemente do seu grau de competência, opta por utilizar apenas impostos não distorcionários para financiar os gastos. O modelo assume ainda que a competência dos governantes é estocástica e persistente. Admite que se verifica um choque na competência em cada período, que o choque actual apenas é percebido pelos eleitores com um período de atraso e define a competência actual como a soma do choque ocorrido no período corrente e o choque verificado no período anterior.

Rogoff e Sibert (1988) demonstram que todos os governantes, excepto os menos competentes (ε=0), aumentam os impostos distorcionários na véspera das eleições para sinalizarem a sua competência.[5] Em suma, o modelo prevê que em períodos pré-eleitorais o nível de impostos não distorcionários seja inferior ao eficiente e que o nível de inflação seja superior ao óptimo. A intuição é a mesma se em vez de inflação considerarmos os défices. Neste caso, antes dos escrutínios teremos impostos menores e défices mais elevados. No modelo de Rogoff (1990), que enfatiza a composição das despesas públicas em vez das receitas de senhoriagem, os governantes sinalizam a sua competência aumentando as despesas públicas em bens e serviços e em transferências e diminuindo o investimento público, uma vez que este último apenas se torna visível e produtivo no período seguinte.

O modelo de ideologia foi criado por Hibbs (1977) e, de uma forma muito sucinta, afirma que uma vez no poder os políticos tentam favorecer os grupos da população que os elegeram, o que os leva a terem objectivos diferentes quanto à evolução das variáveis económicas. Tal como Nordhaus, Hibbs admite a existência de uma curva de Phillips explorável através das políticas orçamental e monetária. As expectativas racionais foram introduzidas no modelo de ideologia por Alesina (1987). De acordo com este autor, a incerteza quanto à ideologia do partido que ganhará as eleições poderá originar erros nas expectativas de inflação para o período imediatamente posterior às mesmas. Tornar-se-á desta forma possível aos governantes, em função da sua ideologia, desviar a taxa de desemprego do seu valor natural imediatamente após as eleições.

Apesar da literatura internacional sobre ciclos político-económicos ser já bastante extensa, a maioria dos trabalhos empíricos analisa o comportamento dos governos centrais. No entanto, já no início da década de 90, Rogoff (1990: 33-34) sugeria o interesse de investigar o poder local. Segundo este autor, os testes empíricos à existência de ciclos eleitorais deveriam centrar-se nos instrumentos de política orçamental em virtude de ser mais fácil aos governantes manipularem os mesmos que os agregados macroeconómicos. Afirmava, ainda, que para variáveis como os impostos e os gastos, seria possível recorrer a

[5] No caso de um político com competência mínima optar por utilizar impostos distorcionários, qualquer governante mais competente é capaz de adoptar uma política que envolva impostos não distorcionários menores com um nível de impostos distorcionários igual ou até inferior. Por conseguinte, o recurso a receitas de senhoriagem não aumenta a probabilidade de reeleição e origina um custo de bem-estar social, não podendo ser um equilíbrio. Note-se que no modelo, a função objectivo dos governantes tem em consideração o seu desejo de reeleição e a perda de bem-estar social causada pela senhoriagem.

dados referentes aos governos locais, com a vantagem de assim se obter um muito maior número de observações que em análises agregadas.

Um dos primeiros estudos a seguir a sugestão de Rogoff foi o de Blais e Nadeau (1992), que testou a existência de ciclos político-orçamentais em dez províncias Canadianas, de 1951 a 1984. Os resultados encontrados sugerem a existência de um ciclo eleitoral de pequena amplitude, apenas visível no ano do escrutínio, cujo efeito é mais notório nas despesas em serviços sociais e em estradas. Segundo estes autores, não existem diferenças significativas na amplitude do comportamento oportunista dos governos locais em função da sua ideologia, do tempo de duração dos seus mandatos, ou do facto de ser o seu primeiro mandato ou de terem sido reeleitos.

No mesmo ano, tomando como referência os governos locais israelitas, Rosenberg (1992) apresenta um modelo em que na determinação do montante de despesas públicas ao longo dos mandatos os governantes têm em consideração não só o objectivo de serem reeleitos, mas também, a sua situação no caso de perderem a eleição. Neste último caso, a manipulação das despesas públicas poderá ter em vista um aumento das oportunidades de emprego no sector privado, ou mesmo uma transferência directa de rendimentos através da alocação de contratos a empresas do sector privado.[6] De acordo com o modelo, os governantes que decidem não se recandidatar aumentam mais as despesas públicas antes dos escrutínios que aqueles que se recandidatam. Os testes efectuados nas despesas de desenvolvimento de dez cidades israelitas, utilizando dados anuais que vão de 1964 a 1982, confirmam esta hipótese.

Vários estudos se seguiram em países como os EUA, a Alemanha e a Suécia, entre outros.[7] No que concerne à realidade portuguesa, tanto ao nível do poder central como local os estudos realizados são ainda muito escassos.[8] Sendo Portugal uma democracia recente, o problema da escassez de observações para efectuar uma análise agregada é ainda mais acentuado que para outros países. Por conseguinte, optamos por investigar o comportamento dos governos locais ao nível dos municípios.

3. Municípios Portugueses: Breve Caracterização e Fontes de Dados Estatísticos

Os municípios portugueses foram formalmente criados na Constituição de 1976 e as primeiras eleições autárquicas ocorreram em Dezembro do mesmo ano. O painel de dados que serve de laboratório à nossa análise engloba a totalidade dos municípios de Portugal Continental (278), durante os anos de 1979 a 2000, o que nos permite cobrir seis períodos eleitorais. No que diz respeito a modificações na organização territorial durante o período analisado, há que destacar a criação de quatro novos municípios (Amadora, Vizela,

6 Desenvolvimentos destas ideias podem ser encontrados em Baleiras e Santos (2000).

7 Para uma extensa e actualizada revisão da literatura empírica sobre os EUA veja-se Besley e Case (2003). Para estudos sobre a Alemanha ver Seitz (2000) e Galli e Rossi (2002) e sobre a Suécia ver Petterson-Lidbom (2001).

8 Relativamente ao poder local, ver os documentos de trabalho de Baleiras e Costa (2001) e Veiga (2002).

Odivelas e Trofa) para os quais não existe, por conseguinte, informação para todos os anos.[9]

Actualmente, o quadro de transferência de atribuições e competências para as autarquias locais é regulado pela Lei n. 159/99[10]. De acordo com esta, os municípios dispõem de atribuições nos seguintes domínios: equipamento rural e urbano, energia, transportes e comunicações, educação, património, cultura e ciência, tempos livres e desporto, saúde, acção social, habitação, protecção civil, ambiente e saneamento básico, defesa do consumidor, protecção do desenvolvimento, ordenamento do território e urbanismo, polícia municipal e cooperação externa.

É importante realçar que não há diferenças nas instituições dos municípios de Portugal continental e que estes são todos regulados pelo mesmo regime financeiro, sendo o das regiões autónomas objecto de diploma próprio.[11] Deve, no entanto, ser mencionado que durante o período analisado vigoraram quatro leis das finanças locais.[12] Os municípios dispõem de património e finanças próprias, cuja gestão compete aos respectivos órgãos, sendo a tutela sobre a sua gestão meramente inspectiva. A autonomia financeira dos municípios resulta dos seguintes poderes que são conferidos aos seus órgãos:

a) Elaborar, aprovar e modificar as opções do plano, orçamentos e outros documentos previsionais.

b) Elaborar e aprovar os documentos de prestações de contas;

c) Arrecadar e dispor de receitas que por lei lhes forem destinadas e ordenar e processar as despesas legalmente autorizadas;

d) Gerir o seu próprio património, bem como aquele que lhes for afecto.

O objectivo do presente trabalho é testar a existência de ciclos eleitorais nos municípios e analisar as consequências macroeconómicas do eventual comportamento oportunista dos autarcas. Tomando em consideração que é mais fácil aos autarcas manipularem os instrumentos de política que a evolução das variáveis económicas no seu município, restringimos o nosso estudo às finanças locais.[13] Tendo presente que os

9 O concelho da Amadora foi criado em 8 de Agosto de 1979, segundo a Lei n. 45/79, incorporando freguesias antes pertencentes aos concelhos de Oeiras e Sintra. A Lei n. 83/98 de 14 de Dezembro criou o município da Trofa, englobando freguesias antes pertencentes ao concelho de Santo Tirso. A Lei n. 63/98 de 1 de Setembro criou o município de Vizela incorporando freguesias antes pertencentes aos concelhos de Guimarães, Lousada e Felgueiras. A Lei n. 84/98 de 14 de Dezembro criou o município de Odivelas com freguesias antes pertencentes ao concelho de Loures.

10 Esta Lei revogou o Decreto-Lei n. 77/84 de 8 de Março.

11 A Lei das finanças locais (Lei n. 46/98 de 6 de Agosto) define o regime financeiro dos municípios e das freguesias do continente, enquanto a Lei n. 13/98 de 24 de Fevereiro regulamenta as finanças das Regiões Autónomas.

12 Lei n. 1/79, Decreto-Lei n. 98/84, Lei n. 1/87 e, actualmente, a Lei n. 46/98.

13 Esta opção vai de encontro ao sugerido por Rogoff (1990: 33-34) e também aos resultados obtidos por Besley e Case (2003) para os EUA. Estes autores não encontraram evidência empírica favorável à existência de ciclos político-económicos no rendimento *per capita*, nem na taxa de desemprego dos Estados americanos, o que os leva a corroborar a sugestão de Kenneth Rogoff de que uma abordagem mais promissora é a análise dos instrumentos de política económica.

dirigentes locais possuem um maior controlo sobre as despesas que sobre as receitas[14] e dentro destas sobre a componente de capital,[15] incidimos a nossa análise sobre estas rubricas.

Os dados sobre as finanças e população municipais foram extraídos da publicação anual *Finanças Municipais*, da responsabilidade da Direcção Geral das Autarquias Locais (DGAL). Esta publicação existe para os anos compreendidos entre 1979 e 1983 e entre 1986 e 2000. Para os dois anos em falta, 1984 e 1985, os dados foram pedidos directamente às Câmaras Municipais, tendo já sido recolhida informação para 150 e 154 municípios para os anos de 1984 e 1985, respectivamente. Os dados referentes à área dos municípios foram obtidos através da base de dados *Sales Index* da Marktest e os referentes às percentagens da população com menos de 15 anos e com mais de 65 anos foram extraídos dos *Census* de 1970, 1981 e 1991 e do *Anuário Estatístico Regional* do INE (dados de 1996 a 2000).

Recorreu-se à Comissão Nacional de Eleições e ao Secretariado Técnico dos Assuntos para o Processo Eleitoral para obter as datas das eleições autárquicas e os resultados das mesmas. Convém referir a este propósito que as datas dos escrutínios são definidas de forma exógena ao poder local. Desde o restabelecimento da Democracia em 1974, verificaram-se eleições autárquicas nos anos de 1976, 1979, 1982, 1985, 1989, 1993, 1997 e 2001, sempre no mês de Dezembro.

4. Especificação do Modelo

Este artigo procura determinar o impacto de factores políticos nas finanças dos municípios de Portugal Continental, nomeadamente testando as hipóteses subjacentes à existência de ciclos político-económicos ao nível dos saldos de execução orçamental e de vários tipos de despesas municipais.

O primeiro modelo empírico testado tem como variável dependente o saldo de execução orçamental (*per capita*, a preços de 1995), definido da seguinte forma:[16]

$$\text{Saldo} = [\text{RCorr} + (\text{RCap-RAF-RPF})] - [\text{DCorr} + (\text{DCap-DAF-DPF})] \quad (3)$$

Em que, *RCorr* são as receitas correntes, *RCap* as receitas de capital, *RAF* as receitas em activos financeiros, *RPF* as receitas em passivos financeiros, *DCorr* as despesas correntes, *DCap* as despesas de capital, *DAF* as despesas em activos financeiros, *DPF* as despesas em passivos financeiros. Valores positivos de *Saldo* correspondem a superavites e valores negativos a défices.

---

14 As transferências da Administração Central e da U.E. representam uma fonte muito importante (em média, 70 por cento) das receitas dos municípios.

15 Ao nível das despesas correntes o pagamento de encargos pouco flexíveis tais como as despesas com o pessoal, electricidade, água, etc. assumem uma grande importância.

16 Para cada município, o valor do saldo foi dividido pelo índice de preços ao consumidor para o ano base de 1995 e, depois, pela sua população total. Embora fosse eventualmente preferível trabalhar com a Capacidade/Necesidade de Financiamento de cada município, calculada de acordo com o Sistema Europeu de Contas de 1995 (SEC95), tal não é possível por o Instituto Nacional de Estatística só disponibilizar dados agregados para essa variável. Assim, usamos o saldo de execução orçamental calculado com base na contabilidade pública, seguindo a metodologia adoptada pela Direcção Geral do Orçamento do Ministério das Finanças, que exclui dos totais de receitas e despesas as transacções em activos e passivos financeiros.

Tendo em conta que os autarcas têm pouco controle sobre as receitas dos seus municípios, é previsível que as rubricas de despesa estejam mais sujeitas a manipulações com intuitos políticos. Deste modo, foi estimado um segundo modelo que tem como variável dependente as despesas totais (reais *per capita*) do município, *DespTot*. Mas, como também há pouca margem de manobra ao nível das despesas correntes, que são fortemente condicionadas pela massa salarial, é previsível que a evidência de ciclos político-económicos seja maior ao nível das despesas de capital, *DespCap*, e, dentro destas, nas despesas de investimento, *DespInv* (ambas expressas em valores reais, *per capita*).[17] Assim, são também estimadas equações para estes tipos de despesa.

Nos quatro modelos acima referidos são usadas as seguintes variáveis explicativas:

- Valores desfasados da variável dependente, de forma a ter em conta a componente autoregressiva da série temporal;
- $TransfTot_{it}$ é o total de transferências reais *per capita* que o município *i* recebe durante o ano *t*. Dado o seu peso de cerca de 70 por cento nas receitas dos municípios, antevê-se que tenham forte influência sobre as despesas totais, *DespTot*. Nas equações para as despesas de capital (*DespCap*) e investimento (*DespInv*) utilizamos, em alternativa, as transferências de capital reais *per capita* que o município *i* recebe durante o ano *t*, $TransfCap_{it}$.[18] Estas variáveis, ao reflectirem a evolução da realidade macroeconómica do país, permitem-nos controlar para alterações na mesma. Prevê-se que maiores transferências permitam maior despesa;
- $AnoEleição_{it}$ é uma variável binária que assume o valor um em anos de eleições autárquicas e o valor zero nos demais.[19] Com esta variável testamos a hipótese de os défices orçamentais e as despesas dos municípios serem maiores em anos de escrutínio. Assim, um coeficiente de sinal negativo é esperado para *AnoEleição* na equação para o *Saldo* e sinais positivos são esperados nas equações para *DespTot*, *DespCap* e *DespInv*;
- $Direita_{it}$ é uma variável binária que iguala um quando o Presidente de Câmara do município *i* é de um partido de direita (PPD/PSD ou CDS/PP) e zero quando pertence a um partido de esquerda (PS, PCP/CDU ou PRD). Com esta variável testamos a existência de ciclos ideológicos nos saldos e nas despesas;
- Tendo em conta que as despesas *per capita* dos municípios também podem depender de factores como a estrutura etária da população, a densidade populacional, a localização geográfica, a população, etc., incluímos em todas as estimações as seguintes cinco variáveis de controlo:
    - $\%Pop<15$ - Percentagem da população com menos de 15 anos;

[17] Trabalhamos com valores reais *per capita* de forma a ser possível a comparação de dados ao longo do tempo e entre municípios de diferentes dimensões.

[18] As transferências de capital representam, em média, 72 por cento das despesas de capital. As estatísticas descritivas podem ser consultadas na tabela 1.

[19] Esta definição da variável eleitoral segue Alesina, Cohen e Roubini (1997).

- *%Pop>65* - Percentagem da população com mais de 65 anos;
- *Denspop* – Densidade populacional;
- *Litoral* – Variável aleatória que toma o valor de um para municípios que pertencem a distritos do litoral e o valor de zero para os restantes;
- *CatPop* – Categoria populacional: 1 – Lisboa e Porto; 2 – outros, com população igual ou superior a 40 mil habitantes; 3 – com população entre 10 mil e 40 mil habitantes; 4 – restantes municípios.[20]

[inserir Tabela 1]

O modelo empírico pode ser resumido da seguinte forma:

$$y_{it} = \sum_{j=1}^{p} \alpha_j y_{i,t-j} + \mathbf{X}_{i,t}^{'}\boldsymbol{\beta} + \nu_i + \varepsilon_{it} \qquad i = 1,...,N \quad t = 1,...,T_i \tag{4}$$

em que y_{it} é a variável dependente e p é o número de valores desfasados da mesma, $\mathbf{x}_{it}^{'}$ é um vector de variáveis explicativas, $\boldsymbol{\beta}$ é um vector de parâmetros a estimar, ν_i é o efeito individual do município i, e ε_{it} é o termo de erro.

Dada a presença de efeitos individuais dos municípios, ν_i, os modelos acima descritos poderiam ser estimados assumindo tais efeitos como fixos ou aleatórios. No entanto, o valor desfasado da variável dependente estaria correlacionado com o termo de erro, ε_{it}, mesmo que este não estivesse autocorrelacionado. Isto implicaria estimações inconsistentes dos parâmetros do modelo, quando na amostra existisse uma clara dominância do número de indivíduos face ao período temporal.[21] Este é exactamente o caso do nosso painel, em que o número de municípios (N=278) é cerca de 12 vezes maior que o número de anos considerado (T=22).

Arellano e Bond (1991) desenvolveram um estimador de método de momentos generalizado (*Generalized Method of Moments* - GMM) para ultrapassar os problemas acima referidos. Tirando primeiras diferenças de (4), os efeitos individuais (ν_i) desaparecem e o modelo resultante pode ser estimado por variáveis instrumentais:

$$\Delta y_{it} = \Delta \sum_{j=1}^{p} \alpha_j y_{i,t-j} + \Delta \mathbf{X}_{i,t}^{'}\boldsymbol{\beta} + \Delta \varepsilon_{it} \qquad i = 1,...,N \quad t = 1,...,T_i \tag{5}$$

São instrumentos válidos: os níveis da variável dependente, desfasados dois ou mais períodos ($y_{i1},\ldots,y_{it-2}$); os níveis das variáveis endógenas, desfasados dois ou mais períodos ($x_{i1},\ldots,x_{it-2}$); os níveis das variáveis pré-determinadas, desfasados um ou mais períodos ($x_{i1},\ldots,x_{it-1}$); e os níveis das variáveis exógenas, correntes e desfasados ($x_{i1},\ldots,x_{it}$) ou, simplesmente, as primeiras diferenças das variáveis exógenas (Δx_{it}).

---

20 Estas categorias populacionais correspondem às usadas para definir os salários dos Presidentes de Câmara.

21 Ver Arellano e Bond (1991), Baltagi (2001) e Bond (2002).

Mais condições de momentos podem ser usadas se assumirmos que as variáveis explicativas (x_{it}) não estão correlacionadas com os efeitos individuais (ν_i). Neste caso, valores desfasados um período destas variáveis (x_{it-1}) podem ser utilizados como instrumentos na equação em níveis. A estimação combina então o conjunto de condições de momentos das equações em primeiras diferenças com as condições adicionais especificadas para a equação em níveis.

Para o caso em que uma variável explicativa x_{it} está correlacionada com o efeito individual ν_i mas as suas primeiras diferenças (Δx_{it}) já não estão correlacionadas com aquele, valores desfasados das primeiras diferenças (Δx_{it-1}) podem ser usados como instrumentos na equação em níveis (ver Arellano e Bover, 1995). Nessa equação também incluímos como instrumentos os valores desfasados das primeiras diferenças da variável dependente ($\Delta y_{i,t-1}$). Blundell e Bond (1998) e Blundell, Bond e Windmaijer (2000) mostram que este estimador GMM estendido é preferível ao de Arellano-Bond (1991) quando a variável dependente e/ou as variáveis independentes são persistentes.[22] É esta versão estendida do estimador GMM para painéis dinâmicos, que Bond (2002) designa de sistema-GMM (*GMM-SYS*), que é utilizada nas estimações apresentadas na secção seguinte.

5. Resultados Empíricos

Na tabela 2 são apresentados os resultados das estimações dos modelos descritos na secção anterior usando o método sistema-GMM para modelos lineares dinâmicos de dados em painel.[23] Os resultados correspondem aos da estimação de 2º passo (*two-step results*), usando desvios padrão robustos a heterocedasticidade e corrigidos para amostras pequenas.[24] As estatísticas-t são apresentadas entre parêntesis e a significância estatística é assinalada com asteriscos. No fundo da tabela são indicados o número de observações e de municípios.

[inserir Tabela 2]

O valor desfasado um período das variáveis dependentes é sempre estatisticamente significativo, tendo sido necessário incluir um segundo na equação para as despesas de

22 Tendo em conta que se verifica alguma persistência ao nível das rubricas de despesas e de transferências, é conveniente a estimação deste sistema-GMM. Acresce o facto de os resultados da implementação dos testes de diferença de Sargan propostos por aqueles autores indicarem claramente que, para os nossos dados, o sistema-GMM é preferível ao GMM só com equações em primeiras diferenças.

23 Como na literatura não há consenso acerca do número de instrumentos a incluir num sistema-GMM, optamos por considerar sempre todos os instrumentos válidos. Excepto na equação para o *Saldo*, tal opção resultou em matrizes de cerca de 500 instrumentos. Embora a hipótese da validade das restrições sobre-identificadores não seja rejeitada quando usamos matrizes de instrumentos de menor dimensão, achamos preferível a inclusão de todas as condições de momentos disponíveis nas estimações.

24 Embora na literatura seja comum apresentar os resultados do 1º passo por os do 2º terem geralmente desvios padrão enviesados para baixo, tal problema não se verifica no nosso caso já que o programa econométrico PcGive 10.2 usa a correcção para amostras pequenas sugerida por Windmeijer (2000). Assim, usamos os resultados do 2º passo, já que estes têm a vantagem de serem mais consistentes na presença de heterocedasticidade.

investimento.[25] Tal como antecipado, quanto maiores as transferências que um município recebe num dado ano, maiores são as suas despesas: os coeficientes estimados associados a *TransfTot* e *TransfCap* têm sinais positivos nas três últimas equações.[26]

Há evidência de ciclos oportunistas racionais para as quatro variáveis consideradas na tabela 2, já que *AnoEleição* é sempre estatisticamente significativa e tem o sinal esperado. Assim, em anos de eleições autárquicas registam-se maiores défices e maiores despesas totais, de capital e de investimento do que noutros anos. Tal como antecipado, a evidência estatística é relativamente fraca para o saldo e para as despesas totais (para as quais *AnoEleição* apenas é marginalmente estatisticamente significativa) e bem maior para as despesas de capital e de investimento.[27] Os resultados indicam que, mantendo tudo o resto constante, as despesas de investimento aumentam cerca de 2170 escudos (de 1995) por habitante no ano do escrutínio, um aumento relativo (comparado com a média amostral) de 6,8 por cento. Para as despesas de capital, o aumento no ano da eleição é de 1138 escudos de 1995 (aumento relativo de 3,2 por cento).[28]

Quanto a efeitos ideológicos, os défices e as despesas de capital são mais elevados nos municípios liderados por políticos de direita, as despesas totais são maiores nos municípios governados por Presidentes de Câmara de esquerda e não há evidência de efeitos partidários nas despesas de investimento. Dada a diversidade e baixa significância estatística dos resultados para as diferentes rubricas, não nos parece possível chegar a conclusões gerais sobre a existência de efeitos ideológicos nas finanças municipais.

Algumas variáveis de controlo também ajudam a explicar as diferenças entre os saldos e despesas reais *per capita* dos municípios: concelhos onde a percentagem da população com menos de 15 anos é maior registam saldos mais positivos, menores despesas totais e maiores despesas de investimento; uma maior percentagem de idosos (idade superior a 65 anos) está associada a maiores défices e não parece afectar as rubricas de despesa; os saldos e as despesas não tendem a variar com a densidade populacional; municípios pertencentes a distritos do litoral tendem a registar maiores despesas reais *per capita*, mas não há diferenças ao nível dos saldos em relação aos municípios de distritos do interior; concelhos menos populosos tendem a registar menores despesas totais, mas não há diferenças estatisticamente significativas ao nível dos saldos e das despesas de capital e de investimento.

Nas estimações cujos resultados são apresentados na tabela 3 começou-se por averiguar se a magnitude do ciclo eleitoral nas despesas de investimento depende da

---

25 A escolha do número de valores desfasados a incluir em cada equação baseou-se na significância estatística daqueles e na necessidade de evitar autocorrelação de segunda ordem nos resíduos. Embora o segundo valor desfasado de *DespInv* não seja estatisticamente significativo, a regressão sofre de problemas de autocorrelação de segunda ordem dos resíduos quando não o incluímos.

26 Embora as transferências sejam exógenas face às despesas (na equação em primeiras diferenças), os seus níveis estão correlacionados com os efeitos individuais dos municípios. Assim, as primeiras diferenças das transferências, desfasadas um período, foram usadas como instrumentos na equação em níveis.

27 Foi também estimado um modelo para as despesas correntes, cujos resultados confirmaram a nossa hipótese de que para estas não haveria comportamento oportunista, dada a sua maior rigidez. Encontramos, no entanto, evidência a favor de efeitos ideológicos: autarcas de esquerda gastam mais que os de direita em despesas correntes. Estes resultados serão disponibilizados a quem os solicitar.

28 Os aumentos relativos para o saldo e para as despesas totais são de, respectivamente, 24 e 2,6 por cento. Embora o aumento do défice indicie um forte efeito oportunista, este resultado deve ser interpretado com precaução, já que o coeficiente estimado apenas é marginalmente significativo (valor-p de 0,10).

ideologia, do apoio na Assembleia Municipal, ou do facto de o autarca se recandidatar.[29] Na coluna 1, a variável *AnoEleição* foi multiplicada por variáveis binárias que representam a ideologia dos autarcas: *Direita* e *Esquerda* (=1-*Direita*). Os resultados indicam que todos os autarcas são oportunistas, mas os de esquerda aumentam mais as suas despesas no ano da eleição (em relação aos outros anos) do que os de direita: o coeficiente associado a *AnoEleição*Esquerda* é mais do dobro do de *AnoEleição*Direita*.[30]

[inserir Tabela 3]

Na coluna 2 averiguámos a possibilidade do apoio de uma maioria dos deputados da Assembleia Municipal afectar as despesas e a magnitude do ciclo oportunista. Assim, foi incluída no modelo a variável binária *Maioria*, que iguala um quando o partido do autarca dispõe de uma maioria na Assembleia Municipal e zero na situação contrária. Adicionalmente, a variável *AnoEleição* foi interagida com as variáveis *Maioria* e *Minoria* (=1-*Maioria*). Os resultados indicam que uma maioria induz a maiores despesas médias de investimento (cerca de 926 escudos por habitante), mas não afecta a magnitude do efeito oportunista.[31]

A hipótese avançada por Rosenberg (1992) de que os autarcas que não se recandidatam geram um ciclo oportunista de maior magnitude do que os que se recandidatam é testada na coluna 3. Para o efeito, *AnoEleição* foi interagida com as variáveis mudas *Recand*, que iguala um quando o Presidente da Câmara se recandidata e zero quando não se recandidata, e *N.Recand* (*=1-Recand*). Embora o coeficiente associado a *AnoEleição*Recand* seja maior do que o de *AnoEleição*N.Recand*, o que poderia indicar maior oportunismo por parte dos autarcas que se recandidatam, um teste de *Wald* não rejeita a igualdade dos coeficientes. Assim, os nossos resultados não confirmam a hipótese e os resultados empíricos de Rosenberg (1992) para Israel.[32]

Tendo em consideração que alguns investimentos podem demorar vários meses a estarem concluídos é de esperar que os governantes comecem a aumentar as despesas de investimento, com o objectivo de transmitirem competência ao eleitorado, já no ano que precede o escrutínio. Para testar esta hipótese acrescentamos ao modelo da última coluna da tabela 2 a variável binária *AnoAntesEleição*, que assume o valor de um no ano anterior ao da eleição e de zero nos restantes anos. Os resultados apresentados na coluna 4 da tabela 3 revelam que esta variável é estatisticamente significativa e tem um coeficiente um pouco menor que o de *AnoEleição.*[33] Assim, o aumento das despesas de investimento antes das

29 Em todas as estimações foram incluídas as cinco variáveis de controlo. Os coeficientes e as estatísticas-t não são mostrados na tabela 3 e nas tabelas seguintes de forma a poupar espaço e porque o objectivo do nosso estudo é averiguar a existência de ciclos politico-económicos, o que não obriga à análise dos resultados respeitantes às variáveis de controlo.

30 Um teste de *Wald* rejeita claramente a hipótese de igualdade dos coeficientes.

31 Um teste de *Wald* não rejeita a igualdade dos coeficientes associados a *AnoEleição*Maioria* e a *AnoEleição*Minoria*.

32 Embora o estudo de Baleiras e Costa (2001) para 30 municípios portugueses apresente resultados semelhantes aos de Rosenberg (1992), estes autores não tiveram em conta a autocorrelação dos resíduos, pelo que poderão ter uma regressão espúria. Desta forma, não é possível assegurar que os resultados obtidos são válidos. Para além disso, se a amostra não for representativa, as suas conclusões só serão aplicáveis aos 30 municípios que analisaram.

33 Um teste de *wald* rejeita a igualdade dos coeficientes.

eleições parece começar no ano anterior às mesmas, sendo reforçado no ano das eleições. O modelo da coluna 5 acrescenta a variável binária *AnoEleiçãoLegislativa*, que iguala um em anos de eleições legislativas e zero nos restantes anos. Procurámos aqui testar a hipótese dos Presidentes de Câmara aumentarem as despesas de investimento em anos de eleições legislativas de forma a contribuírem para um melhor desempenho do seu partido ao nível nacional. Os resultados da estimação realizada suportam esta hipótese.

Em consonância com os resultados da tabela 2, praticamente não há evidência a favor da existência de efeitos ideológicos nas despesas de investimento: a variável *Direita* apenas é marginalmente estatisticamente significativa na estimação da coluna 1 e não é significativa nas restantes estimações.

O passo seguinte da análise empírica consistiu em determinar em que tipos de despesas de investimento há maior evidência de ciclos eleitorais e efeitos partidários. Com esse objectivo, estimámos o modelo da coluna 4 da tabela 3 para as sete componentes das despesas de investimento (ver tabela 4). Os ciclos oportunistas registam-se nos investimentos em *Outros Edifícios*, *Construções Diversas* e *Outros Investimentos*, nos quais se verificam aumentos relativos no ano da eleição, face à média amostral, de 13,6, 11,5 e 16,6 por cento, respectivamente. Nas *Construções Diversas*, o aumento das despesas começa a verificar-se um ano antes do escrutínio. Quanto a efeitos de ideologia, os autarcas de direita tendem a gastar relativamente mais em *Aquisição de Terrenos* e *Construções Diversas* (aumentos relativos de 24,3 e 4,9 por cento, respectivamente), enquanto os de esquerda gastam relativamente mais em *Material de Transporte* e *Maquinaria e Equipamento* (aumentos relativos de 11,6 e 14,6 por cento, respectivamente).

[inserir Tabela 4]

Dada a forte evidência a favor da existência de ciclos eleitorais nas rubricas de *Outros Edifícios* e de *Construções Diversas*,[34] decidimos analisar as componentes das mesmas. De forma a economizar espaço, apresentamos na tabela 5 apenas os resultados de estimações para as rubricas em que esta evidência se verifica.[35] No que toca a *Outros Edifícios*, tal só acontece para a sub-componente *Outros* (que tem um peso de 53 por cento), em que as despesas aumentam no ano da eleição e, um pouco menos, no ano anterior (aumentos relativos de 21,6 e 17,6 por cento face à média amostral). No que diz respeito às subdivisões de *Construções Diversas*, encontramos clara evidência de ciclos oportunistas em *Viadutos, arruamentos e obras complementares*, *Viação rural*, e *Outros*, com aumentos relativos, face à média amostral, de 14,8, 16,4 e 36,8 por cento, respectivamente. Convém também notar que estas três rubricas representam cerca de 69 por cento das construções diversas. Para a primeira e última sub-componentes, as despesas aumentam também no ano anterior à eleição, embora em menor magnitude (aumentos relativos de 10,6 e 27,4 por cento, respectivamente). Para *Infraestruturas e tratamento de resíduos sólidos* a evidência estatística é bem menor, mas o aumento relativo no ano da eleição é de 47,5 por cento face à média (o maior de todos). No entanto, há uma redução das despesas nesta rubrica no ano antes da eleição. Para as rubricas de despesas de

[34] Estas são as duas componentes mais importantes das despesas de investimento, pois juntas representam cerca de 83 por cento do total (Outros investimentos = 17,3 por cento e Construções diversas = 65,6 por cento).

[35] Os autores facultarão os restantes resultados a quem os solicitar.

investimento incluídas na Tabela 5 não há qualquer evidência de efeitos partidários, já que a variável aleatória *Direita* nunca é estatisticamente significativa.

[inserir Tabela 5]

Em suma, os resultados confirmam a hipótese de que os ciclos oportunistas se manifestam em maior escala ao nível das despesas de investimento e suas componentes, ou seja, nas rubricas em que os autarcas têm mais margem de manobra. É também importante salientar que a gestão eleitoralista das despesas pelos dirigentes locais incide principalmente nas rubricas de investimento mais visíveis pelo eleitorado, tais como *Outros Edifícios* e *Construções diversas* e, dentro destas últimas, *Viadutos, arruamentos e obras complementares* e *Viação rural*. Ou seja, nota-se claramente a intenção dos autarcas de ganharem popularidade antes das eleições.[36]

6. Impacto Macroeconómico

Analisados os resultados do trabalho empírico importa agora entrar em considerações de ordem normativa. O facto dos autarcas portugueses aumentarem as despesas, com destaque para as de investimento, na véspera das eleições, com o objectivo de parecerem mais competentes perante o eleitorado e vencerem as eleições, sugere a existência de ineficiência na alocação dos recursos colocados à disposição do município. Assim, disposições legais que diminuam o poder discricionário dos autarcas e a sua capacidade de gerar ciclos oportunistas poderão melhorar o bem-estar social.

Importa também avaliar o impacto deste comportamento ao nível macroeconómico, tomando em consideração o conjunto dos municípios. O Gráfico 1 ilustra o peso das despesas e dos saldos dos municípios de Portugal Continental no Produto Interno Bruto a preços de mercado (PIBpm).[37]

[inserir Gráfico 1]

Em consonância com os resultados da secção anterior, o peso das despesas tende a ser maior, e o dos saldos menor, em anos de eleições autárquicas do que nos anos anteriores do mesmo mandato. Ou seja, nos dados agregados visualiza-se facilmente a gestão eleitoralista das finanças municipais. Convém salientar, no entanto, que o impacto macroeconómico é relativamente pequeno, já que, em média, as despesas totais dos municípios representam 3,76 por cento do PIB e as despesas de capital e de investimento equivalem a 1,88 e 1,58 por cento do PIB, respectivamente.[38] Como os défices das

[36] Apesar de entendermos que a estimação de um sistema-GMM é mais apropriada que a de um modelo de efeitos fixos, todas as regressões acima referidas foram também estimadas para este último. Os sinais dos coeficientes e a significância estatística dos mesmos quase não se alteram, registando-se as principais diferenças ao nível da magnitude dos coeficientes estimados.

[37] Como para os anos de 1984 e 1985 só temos dados para 150 e 154 municípios, respectivamente, não são apresentados os valores agregados dos saldos e das despesas nesses anos. Os saldos foram calculados com base na equação 1.

[38] Assim, o aumento relativo de 6,8 por cento nas despesas de investimento num ano de eleições (encontrado na tabela 2) teria pouco impacto macroeconómico: com uma média de 1,58 por cento do PIB, as despesas passariam para 1,69 por cento num ano de eleições (aumentando somente 0,11 por cento do PIB).

autarquias raramente passam de 0,3 por cento do PIB, o seu impacto macroeconómico também é pequeno.

A respeito dos saldos orçamentais interessa-nos sobretudo determinar em que medida um aumento do défice global das autarquias por motivos eleitoralistas poderá por em perigo o cumprimento do Pacto de Estabilidade e Crescimento por parte do Estado Português. Assim, é preferível analisar os défices com base na Capacidade/Necessidade de Financiamento, definida no Sistema Europeu de Contas, pois é essa que importa para o Procedimento dos Défices Excessivos. Os dados para a Administração Local e para o total das Administrações Públicas[39] (em percentagem do PIBpm) são apresentados na tabela 6. Podemos constatar que, excepto para 1985 e 2001, há maior necessidade de financiamento por parte da Administração Local no ano das eleições que em qualquer outro ano do mesmo mandato. Mesmo para 2001, a necessidade de financiamento é maior que a média dos três anos anteriores.

[inserir Tabela 6]

Apesar do eleitoralismo dos autarcas, o peso dos défices da Administração Local no PIB é pequeno: a maior necessidade de financiamento registada desde 1979 correspondeu a 0,43 por cento do PIB (em 2002). No entanto, tais défices poderão dificultar o cumprimento do limite fixado no Pacto de Estabilidade e Crescimento (PEC) para o défice orçamental se a necessidade de financiamento conjunta da administração central e dos fundos de segurança social já estiver perto de 3 por cento do PIB.

Uma forma de diminuir a ineficiência causada pelos ciclos político-económicos e a possibilidade destes dificultarem o cumprimento do Pacto de Estabilidade e Crescimento seria impor limites legais mais apertados à gestão das finanças locais, nomeadamente no que respeita ao seu endividamento. Vários estudos efectuados para os EUA apontam para a existência de menores défices e endividamento em Estados onde as regras orçamentais são mais apertadas.[40] No entanto, há sempre o perigo dos governantes encontrarem formas de contornar os limites impostos através de técnicas de "contabilidade criativa", desorçamentação, criação de empresas municipais, etc.[41] Assim, se não houver especial cuidado na elaboração das regras a impor aos municípios, as mesmas poderão ser pouco eficazes.[42]

---

39 As Administrações Públicas são compostas pela Administração Central, a Administração Local e pelos Fundos de Segurança Social. Embora na Administração Local também estejam incluídas as Freguesias, as despesas e os défices destas são diminutos quando comparados com os dos municípios. Assim, os dados para a Administração Local reflectem essencialmente a situação dos municípios.

40 Ver, por exemplo, Alt e Lowry (1994) e Poterba (1994).

41 Segundo Bunch (1991), os Estados dos EUA que estão sujeitos a apertados limites ao endividamento usam as autoridades/empresas públicas (*public authorities*) para contorná-los. Aquelas acabam depois por emitir parte da dívida pública do Estado, substituindo-se ao governo estadual. O mesmo poderia passar-se em Portugal, com os municípios a transferirem boa parte dos seus serviços e despesas para empresas municipais, de forma a evitarem os limites impostos aos défices e ao endividamento.

42 Besley e Case (2003) argumentam que, embora seja consensual a ideia de que as regras orçamentais afectam o comportamento dos governos, a forma como as mesmas são definidas é muito importante.

7. Conclusões e Implicações de Política Económica

Os resultados empíricos apontam claramente para a existência de ciclos político-económicos nos municípios de Portugal Continental. Em consonância com os ciclos oportunistas racionais de Rogoff e Sibert (1988) os autarcas portugueses gerem os instrumentos de política económica de forma a revelarem maior competência pouco antes das eleições. Há clara evidência de que os défices e as despesas municipais, com destaque para as de investimento, aumentam significativamente no ano das eleições e, em vários casos, no ano anterior. A gestão eleitoralista das despesas pelos dirigentes locais incide principalmente nas rubricas de investimento mais visíveis pelo eleitorado, tais como *Outros Edifícios* (particularmente na componente *Outros*) e *Construções Diversas* (com especial destaque para *Viadutos, arruamentos e obras complementares*, *Viação rural* e *Outros*) o que denota a intenção de ganhar popularidade.[43] A magnitude do ciclo eleitoralista não parece depender do apoio de que o Presidente da Câmara dispõe na Assembleia Municipal, nem da sua decisão de se recandidatar ou não às próximas eleições. No entanto, os resultados sugerem que os dirigentes de esquerda tendem a aumentar mais as despesas nos anos de eleições que os de direita.

Quanto a ciclos ideológicos, é difícil identificar tendências gerais, dada a inconsistência de resultados entre saldos, despesas totais e despesas de capital. Nos totais de despesas de investimento e na maioria das suas subdivisões, não há evidência de efeitos partidários.

Apesar do pequeno peso das despesas e défices dos municípios no PIB, a tendência para elevar a despesa e o endividamento pouco antes das eleições autárquicas poderá, ocasionalmente, dificultar o respeito do Pacto de Estabilidade e Crescimento. Acresce que a instabilidade nas despesas de investimento acarreta ineficiências na alocação dos recursos que prejudicam a economia portuguesa. Assim, poderá ser benéfica a imposição de regras à gestão das finanças públicas locais, expressas na fixação de limites ao endividamento e aos défices, que dificultem o comportamento eleitoralista dos autarcas.

---

43 Em investigação futura tencionamos avaliar em que medida este aumento das despesas antes das eleições influência os resultados das mesmas, conjuntamente com outros indicadores sócio-económicos que caracterizam a situação dos municípios. Estudos realizados ao nível agregado, Veiga e Veiga (2004a e 2004b), permitem-nos concluir que tanto o desemprego como a inflação afectam as intenções de voto dos portugueses e a popularidade das principais entidades políticas.

Tabela 1 (cont.)

Estatísticas Descritivas

Variáveis	Nº Obs.	Média	Desvio Padrão	Mínimo	Máximo
Saldo e rubricas de despesa:					
Saldo	5809	-3.51	14.71	-210.92	185.44
Despesas totais	5797	68.49	42.25	2.29	841.33
Despesas de capital	5749	36.04	24.48	0.72	287.96
Despesas de investimento	5743	31.81	22.51	1.78	288.51
Aquisição de terrenos	5472	0.91	1.84	0.00	40.99
Habitação	5473	1.89	4.64	0.00	100.98
Outros edifícios	4449	5.51	6.40	0.00	80.65
Instalações desportivas, recreativas e escolares	5473	1.98	3.71	0.00	59.73
Equipamento social	5468	0.32	1.37	0.00	41.75
Outros	4424	2.94	4.62	0.00	61.36
Construções diversas	4449	20.89	17.87	0.00	214.70
Viadutos, arruamentos e obras complementares	5470	4.41	5.83	0.00	84.23
Esgotos	4450	2.38	3.90	0.00	69.10
Captação, tratamento e distribuição de água	4450	2.95	4.91	0.00	100.28
Viação rural	5474	5.91	8.79	0.00	152.86
Infra-estruturas e tratamento de resíduos sólidos	4442	0.22	2.08	0.00	98.91
Outros	4175	4.02	6.88	0.00	122.25
Material de transporte	4449	0.99	1.32	0.00	15.64
Maquinaria e equipamento	5475	1.91	2.01	0.00	25.79
Outros investimentos	5195	0.85	2.87	0.00	52.99
Transferências:					
Transferências totais	5664	50.89	38.87	4.53	443.29
Transferências de capital	5741	26.06	19.48	3.04	249.74
Variáveis políticas:					
AnoEleição	6116	0.27	0.44	0.00	1.00

Tabela 1 (cont.)

Estatísticas Descritivas

Variáveis	Nº Obs.	Média	Desvio Padrão	Mínimo	Máximo
AnoAntesEleição	6116	0.27	0.44	0.00	1.00
Direita	6049	0.47	0.50	0.00	1.00
AnoEleição*Direita	6049	0.13	0.34	0.00	1.00
AnoEleição*Esquerda	6049	0.14	0.35	0.00	1.00
Maioria	6049	0.59	0.49	0.00	1.00
AnoEleição*Maioria	6049	0.15	0.36	0.00	1.00
AnoEleição*Minoria	6049	0.12	0.32	0.00	1.00
AnoEleição*Recand	5923	0.21	0.40	0.00	1.00
AnoEleição*N.Recand	5923	0.05	0.22	0.00	1.00
AnoEleiçãoLegislativa	6056	0.36	0.48	0.00	1.00

Fontes: DGAL, OCDE, STAPE e várias Câmaras Municipais.
Nota: O saldo, as rubricas de despesa e as transferências estão sempre expressas em milhares de escudos (a preços de 1995) por habitante.

Tabela 2

Ciclos Político-Económicos Nas Finanças Locais

	Saldo	Despesas Totais	Despesas de Capital	Despesas de Investimento
Saldo (-1)	0.126 (4.36)***			
DespTot (-1)		0.445 (8.59)***		
DespCap (-1)			0.279 (9.31)***	
DespInv (-1)				0.324 (10.90)***
DespInv (-2)				-0.036 (-1.45)
TransfTot		0.537 (6.41)***		
TransfCap			0.845 (15.50)***	0.802 (23.90)***
AnoEleição	-0.860 (-1.90)*	1.787 (1.83)*	1.138 (3.53)***	2.170 (6.16)***
Direita	-0.748 (-1.84)*	-2.464 (-1.95)*	1.314 (2.02)**	0.495 (0.90)
%Pop < 15	0.269 (3.58)***	-1.229 (-6.66)***	-0.102 (-1.04)	0.315 (3.61)***
%Pop > 65	-0.123 (-1.74)*	0.051 (0.22)	-0.053 (-0.43)	-0.025 (-0.23)
DensPop	-.0002 (-0.84)	0.001 (0.97)	0.0005 (0.70)	0.0009 (1.19)
Litoral	0.013 (0.04)	5.029 (3.00)***	2.458 (3.03)***	1.893 (2.55)**
CatPop	-0.534 (-1.59)	-3.244 (-2.47)***	-0.849 (-1.20)	-0.003 (-0.005)
Nº de Observações	5397	5246	5298	4858
Nº de Municípios	275	275	275	275

Fontes: DGAL, STAPE e OCDE.

Notas:

- Estimações de sistemas-GMM (que combinam a equação em primeiras diferenças com a equação em níveis) para modelos lineares de dados em painel, usando o programa econométrico PcGive 10.2;
- Resultados da estimação do 2º passo (2-step results) usando desvios padrão robustos a heterocedasticidade e corrigidos para amostras pequenas;
- Estatísticas-t entre parêntesis. Nível de significância para o qual a hipótese nula é rejeitada: ***, 1%; **, 5%, e *, 10%;
- A hipótese de existência de autocorrelação de segunda ordem dos resíduos é sempre rejeitada e o teste Sargan nunca rejeita a validade das restrições sobre-identificadoras (over-identifying restrictions).

Tabela 3

Ciclos Político-Económicos Nas Despesas De Investimento

DespInv	1	2	3	4	5
DespInv (-1)	0.326	0.322	0.350	0.326	0.331
	(11.10)***	(10.80)***	(10.50)***	(11.00)***	(11.20)***
DespInv (-2)	-0.035	-0.037	-0.052	-0.020	-0.029
	(-1.39)	(-1.50)	(-2.06)**	(-0.78)	(-1.06)
TransfCap	0.800	0.803	0.770	0.792	0.791
	(23.90)***	(23.80)***	(18.50)***	(23.40)***	(23.40)***
Direita	0.998	0.563	0.614	0.415	0.456
	(1.75)*	(0.99)	(1.10)	(0.76)	(0.83)
AnoEleição*Direita	1.154				
	(2.68)***				
AnoEleição*Esquerda	3.211				
	(6.21)***				
Maioria		0.926			
		(1.96)**			
AnoEleição*Maioria		2.079			
		(3.96)***			
AnoEleição*Minoria		2.294			
		(5.71)***			
AnoEleição*Recand			2.263		
			(5.91)***		
AnoEleição*N.Recand			1.567		
			(2.17)**		
AnoEleição				3.222	3.584
				(8.24)***	(8.66)***
AnoAntesEleição				2.406	2.896
				(8.03)***	(7.72)***
AnoEleiçãoLegislativa					0.824
					(2.29)**
Nº de observações	4858	4858	4806	4858	4858
Nº de Municípios	275	275	275	275	275

Fontes: DGAL, STAPE, OCDE e INE.
Notas:

- Estimações de sistemas-GMM (que combinam a equação em primeiras diferenças com a equação em níveis) para modelos lineares de dados em painel, usando o programa econométrico PcGive 10.2;
- Os coeficientes e estatísticas-t para as variáveis de controlo %Pop<15, %Pop>65, DensPop, Litoral e CatPop (incluídas em todas as estimações) não são mostrados de forma a economizar espaço na tabela;
- Resultados da estimação do 2º passo (2-step results) usando desvios padrão robustos a heterocedasticidade e corrigidos para amostras pequenas;
- Estatísticas-t entre parêntesis. Nível de significância para o qual a hipótese nula é rejeitada: ***, 1%; **, 5%, and *, 10%;
- A hipótese de existência de autocorrelação de segunda ordem dos resíduos é sempre rejeitada e o teste Sargan nunca rejeita a validade das restrições sobre-identificadoras (over-identifying restrictions).

Tabela 4

Ciclos Político-Económicos Nas Componentes Das Despesas De Investimento

	Aquisição de Terrenos	Habitação	Outros Edifícios	Construções Diversas	Material de Transporte	Maquinaria e Equipamento	Outros Investimentos
Var.Dep. (-1)	0.227	0.417	0.374	0.289	0.083	0.247	0.098
	(2.75)***	(8.91)***	(9.62)***	(9.69)***	(2.77)***	(5.19)***	(2.71)***
Var.Dep. (-2)		-0.139		-0.082			0.057
		(-3.78)***		(-3.42)***			(2.07)**
TransfCap	0.024	0.031	0.141	0.594	0.002	0.030	0.020
	(4.72)***	(3.48)***	(7.26)***	(13.70)***	(0.61)	(5.68)***	(4.04)***
AnoEleição	-0.099	0.094	0.749	2.396	-0.109	-0.137	0.142
	(-1.54)	(0.63)	(3.83)***	(6.60)***	(-2.77)***	(-2.60)***	(2.33)**
AnoAntesEleição	0.057	0.076	0.178	1.387	0.053	-0.046	0.095
	(1.00)	(0.61)	(0.99)	(4.70)***	(1.44)	(-0.87)	(1.62)
Direita	0.221	0.049	-0.041	1.013	-0.115	-0.279	-0.068
	(3.18)***	(0.27)	(-0.17)	(1.74)*	(-1.89)*	(-3.78)***	(-0.77)
Nº de observações	4805	4230	4083	3807	4080	4813	3713
Nº de Municípios	275	275	275	275	275	275	275

Fontes: DGAL, STAPE, OCDE e INE.

Notas:

- *Var.Dep.* representa a variável dependente descrita no título de cada coluna;
- Estimações de sistemas-GMM (que combinam a equação em primeiras diferenças com a equação em níveis) para modelos lineares de dados em painel, usando o programa econométrico PcGive 10.2;
- Os coeficientes e estatísticas-t para as variáveis de controlo *%Pop<15, %Pop>65, DensPop, Litoral e CatPop* (incluídas em todas as estimações) não são mostrados de forma a economizar espaço na tabela;
- Resultados da estimação de 2º passo *(2-step results*) usando desvios padrão robustos a heterocedasticidade e corrigidos para amostras pequenas;
- Estatísticas-t entre parêntesis. Nível de significância para o qual a hipótese nula é rejeitada: ***, 1%; **, 5%, and *, 10%;
- A hipótese de existência de autocorrelação de segunda ordem dos resíduos é sempre rejeitada e o teste Sargan nunca rejeita a validade das restrições sobre-identificadoras (*over-identifying restrictions*).

Tabela 5

Sub-Componentes das Despesas de Investimento para as Quais Há Evidência de Ciclos Político-Económicos

Componente:	Outros Edifícios	Construções Diversas			
Sub-componente:	Outros	Viadutos, arrua-mentos e obras complementares	Viação Rural	Infraestruturas e tratamento de resíduos sólidos	Outros
Var.Dep. (-1)	0.303 (6.01)***	0.389 (4.65)***	0.384 (10.40)***	0.362 (7.40)***	0.359 (6.03)***
TransfCap	0.065 (4.26)***	0.094 (3.93)***	0.146 (5.60)***	0.025 (1.53)	0.167 (5.68)***
AnoEleição	0.633 (4.12)***	0.653 (3.38)***	0.970 (4.04)***	0.105 (1.75)*	1.477 (5.87)***
AnoAntesEleição	0.518 (3.94)***	0.468 (2.90)***	0.169 (1.03)	-0.096 (-2.43)**	1.099 (4.79)***
Direita	0.023 (0.13)	0.336 (1.51)	0.316 (0.78)	-0.069 (-1.32)	-0.268 (-0.99)
Nº de observações	4080	4801	4809	4077	3682
Nº de Municípios	275	275	275	275	275

Fontes: DGAL, STAPE, OCDE e INE.

Notas:

- *Var.Dep.* representa a variável dependente descrita no título de cada coluna;
- Estimações de sistemas-GMM (que combinam a equação em primeiras diferenças com a equação em níveis) para modelos lineares de dados em painel, usando o programa econométrico PcGive 10.2;
- Os coeficientes e estatísticas-t para as variáveis de *controlo %Pop<15, %Pop>65, DensPop, Litoral e CatPop* (incluídas em todas as estimações) não são mostrados de forma a economizar espaço na tabela;
- Resultados da estimação de 2º passo (*2-step results*) usando desvios padrão robustos a heterocedasticidade e corrigidos para amostras pequenas;
- Estatísticas-t entre parêntesis. Nível de significância para o qual a hipótese nula é rejeitada: ***, 1%; **, 5%, and *, 10%;
- A hipótese de existência de autocorrelação de segunda ordem dos resíduos é sempre rejeitada e o teste Sargan nunca rejeita a validade das restrições sobre-identificadoras (*over-identifying restrictions*).

Tabela 6

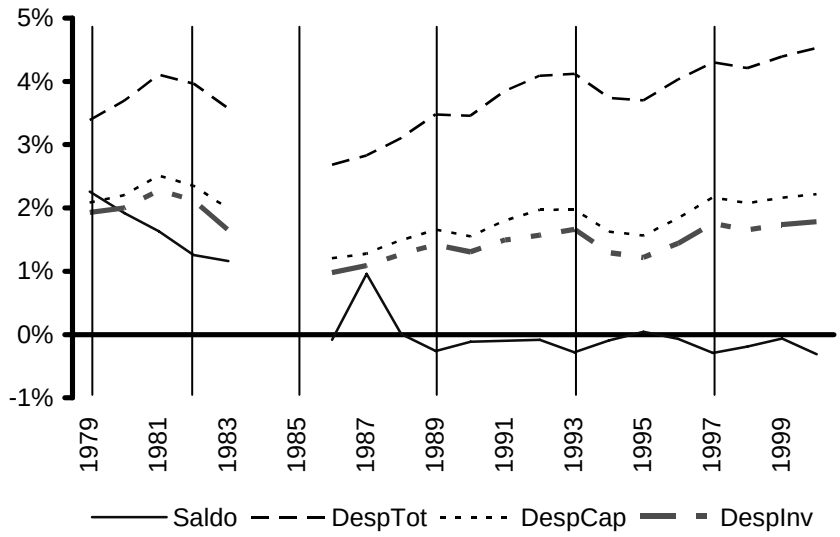
Capacidade (+) / Necessidade (-) de Financiamento (%PIBpm)

	Administração Local	Administrações Públicas
1979	0,33%	-5,80%
1980	-0,06%	-8,49%
1981	-0,34%	-12,46%
1982	-0,43%	-8,55%
1983	-0,24%	-6,71%
1984	-0,10%	-10,17%
1985	-0,17%	-10,28%
1986	0,04%	-5,90%
1987	0,33%	-5,53%
1988	0,25%	-3,17%
1989	-0,13%	-2,13%
1990	-0,17%	-5,03%
1991	-0,20%	-6,00%
1992	-0,17%	-2,95%
1993	-0,34%	-6,10%
1994	-0,14%	-6,07%
1995	0,05%	-5,74%
1996	-0,32%	-3,98%
1997	-0,42%	-2,68%
1998	0,42%	-2,39%
1999	0,26%	-2,84%
2000	-0,41%	-2,83%
2001	-0,36%	-4,22%
2002	-0,43%	-2,64%

Fontes: *Séries Longas*, Banco de Portugal (1979-95) e DGEP, Ministério das Finanças (1996-2002).
Nota: As linhas sombreadas correspondem a anos de eleições autárquicas.

Gráfico 1

Despesas e Saldos dos Municípios em Percentagem do PIBpm



Fontes: DGAL e OCDE (MEI).
Nota: As linhas verticais assinalam as datas de eleições autárquicas.

Bibliografia

Alesina, A. (1987), "Macroeconomic Policy in a Two-Party System as a Repeated Game," *Quarterly Journal of Economics*, 102: 651-678.

Alesina, A., Cohen, G. e Roubini, N. (1997), *Political Cycles and the Macroeconomy*, Cambridge and London, MA, MIT Press.

Alt, J. e Lowry, R. C. (1994), "Divided Government and Budget Deficits: Evidence from the States," *American Political Science Review*, 88: 811-828.

Arellano, M. e Bond, S. (1991), "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *The Review of Economic Studies*, 58: 277-297.

Arellano, M. e Bover, O. (1995), "Another look at the instrumental variable estimation of error-component models," *Journal of Econometrics* 68, 29-51.

Baleiras, R.N. e Costa, J.S. (2001), "To be or not to be in office again, that is the question – political business cycles with local governments," *Working Paper*, Universidade Nova de Lisboa - Portugal.

Baleiras, R.N. e Santos, V. (2000), "Behavioral and institutional determinants of political business cycles," *Public Choice*, 104: 121-147.

Baltagi, B.H. (2001), *Econometric Analysis of Panel Data*, 2nd ed., Chichester, John Wiley & Sons.

Besley, T. e Case, A. (2003), "Political institutions and policy choices: evidence from the United States," *Journal of Economic Literature*, XLI(1): 7-73.

Blais, A. e Nadeau, R. (1992), "The electoral budget cycle," *Public Choice*, 74: 389-403.

Blundell, R. e Bond, S. (1998), "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics* 87, 115-143.

Blundell, R.W., Bond, S.R. e Windmeijer, F. (2000). "Estimation in dynamic panel data models: improving on the performance of the standard GMM estimator." In Baltagi, B.H. (ed.), *Advances in Econometrics Vol. 15: Nonstationary Panels, Panel Cointegration and Dynamic Panels*, JAI Elsevier Science.

Bond, S.R. (2002), "Dynamic panel data models: a guide to micro data methods and practice," *Portuguese Economic Journal*, 1: 141-162.

Bunch, B.S. (1991), "The effect of constitutional debt limits on state governments' use of public authorities," *Public Choice*, 68: 57-69.

Direcção Geral as Autarquias Locais, *Finanças Municipais*, Direcção Geral das Autarquias Locais, Lisboa. 1979 a 1983 e 1986 a 2000.

Galli, E. e Rossi, S. (2002), "Political budget cycles: the case of the Western German Länder," *Public Choice*, 110: 283-303.

Hibbs, D. (1977), "Political Parties and Macroeconomic Policy," *The American Political Science Review*, 7: 1467-1487.

Mueller, D.C. (1997), "Public Choice in Perspective," in Mueller, D.C. (ed.), *Perspectives on Public Choice*, Cambridge University Press: 1-17.

Nordhaus, W. (1975), "The Political Business Cycle," *Review of Economic Studies*, 42: 169-90.

Persson, T. e Tabellini, G. (2002), "Political Economics and Public Finance," in Auberbach, A. e Feldstein, M. (2002), *Handook of Public Economics*, vol. 3, Amsterdam, North-Holland.

Pettersson-Lidbom, Per (2001), "An empirical investigation of the strategic use of debt," *Journal of Political Economy*, 109(3), 570-583.

Poterba, J. (1994), "State responses to fiscal crisis: the effect of budgetary institutions and politics," *Journal of Political Economy*, 102: 798-821.

Rogoff, K. (1990), "Equilibrium political budget cycles," *American Economic Review*, 80: 21-36.

Rogoff, K. e Sibert, A. (1988), "Elections and Macroeconomic Policy Cycles," *Review of Economics Studies*, 55: 1-16.

Rosenberg, J. (1992), "Rationality and the Political Business Cycle: The Case of Local Government," *Public Choice*, 73:71-81.

Seitz, H. (2000), "Fiscal policy, deficits and politics of subnational governments: The case of the German Laender," *Public Choice*, 102: 183-218.

Veiga, L.G. (2002), "The Political Economy of Local Governments' Expenditures," *NIPE - Working Paper* 8/2002.

Veiga, L.G. e Veiga, F.J. (2004a), "Popularity functions, partisan effects and support in Parliament," *Economics & Politics*, 16(1), 101-115.

Veiga, L.G. e Veiga, F.J. (2004b), "The Determinants of Vote Intentions in Portugal," aceite para publicação na revista *Public Choice.*

Windmeijer, F. (2000), "A finite sample correction for the variance of linear two-step GMM estimators," *Institute of Fiscal Studies Working Paper Series* Nº W00/19.

Sessão 2 - Funcionamento do Mercado de Trabalho
Session 2 - Labour Market Functioning

WAGE MOBILITY IN PORTUGAL [*]

Ana Rute Cardoso
IZA Bonn, Universidade do Minho, and CEPR [#]

December 2003

Abstract

This study analyzes mobility in the wage distribution in Portugal, and provides some international comparisons. In particular, a comparison with the UK is performed, replicating the work by Dickens (2000) and progressing to discuss the impact of differences in the institutional framework, which is more regulated and centralized in Portugal than in the UK, with minimum wages, employment protection, and collective bargaining widely applied. The overlapping period with information reported for both countries is 1986-94, and the analysis on Portugal is extended to 1999. The major questions to be answered are: What is the extent of wage mobility in Portugal, and how does it compare to the UK and other countries? Is a more strongly regulated labour market more likely to protect low-wage workers, sustaining their wages, while nevertheless reducing their chances of upward mobility? How has wage mobility evolved over time, in a country marked by sharp growth rates during the second half of the 1980s and by acute shortage of skills? Results indicate that Portugal and the UK became more unequal and less mobile labour markets, having departed from similar levels in mid-80s. Also in France, wage mobility has declined. Mid-90s seem to have been in Portugal a turning point, with rising wage dispersion and high mobility until then, and declining wage dispersion and lower mobility afterwards. Whereas in the UK, low wage workers experienced the sharpest decline in mobility, in Portugal and France, workers in the top half of the wage distribution and holding highest schooling levels experienced the sharpest decline in mobility. The evidence does not support the idea that a more regulated institutional framework reduces individual mobility within the wage distribution. Changes in mobility levels and patterns seem to have been mainly driven by broad trends taking place in these economies, namely the interplay between modernization and technological progress, on one hand, and the supply of highly-schooled workers, on the other. The period of economic growth provided in Portugal opportunity for a rising wage premium and for wage mobility for specific groups of workers, holding upper skills. Much of these changes were feasible under an institutional regime that, despite apparent rigidity, leaves ample scope for firm manoeuvre when bargaining over wages.

Keywords: Wage Mobility; Wage Dispersion.

JEL codes: J31, J60.

---

[*] I am grateful to Pedro Portugal and anonymous referees for their comments on earlier versions of this paper. The data used were provided by the Ministry of Employment, Statistics Department, whose cooperation is acknowledged. The usual disclaimer applies.

[#] IZA, Schaumburg-Lippe-Str. 7/9, 53113 Bonn, Germany. E-mail: cardoso@iza.org

1. Introduction

In recent years, wage mobility has been deserving increasing attention both in the literature and in the policy debate. This growing interest has been mainly driven by the search for policy responses to the widespread increase in labour market inequality. Indeed, in particular in the UK and the USA, the relative position of low-paid workers has worsened markedly, as their wages deteriorated either in absolute terms or relative to the rest of the economy. The relationship between trends in wage inequality and wage mobility became subject to intense debate, as two opposing views emerged. On one hand, some argue that increasing differences in the labour market, as captured by a sequence of snapshots, are not a matter of concern, because rising wage mobility would guarantee that a specific position in the wage distribution is a temporary situation in a worker's life cycle, with ample opportunity for his/her fortune to change over time. In particular, low wages could, in that case, be an entry gate into employment. On the contrary, others claim that rising inequality has been accompanied by declining mobility, in which case positions in the wage hierarchy have become more persistent. In particular, the low-paid would be trapped to remain in a situation of deteriorating wages over their life cycle, eventually revolving between low-pay, unemployment and welfare recipiency, a situation deserving more concern, from a welfare and Social Security point of view. The implications of rising wage dispersion therefore depend crucially on whether it reflects transitory fluctuations in earnings, or an increase in permanent differences across workers.

Even though the literature on the UK and the USA has focused mainly on mobility out of low-wages, the issue is also relevant at higher wage levels. For example, the incentives for acquisition of human capital depend, not just on the returns to that investment, but also on its risk. Whereas the returns to skills are reflected in the dispersion of the wage distribution, its degree of risk is reflected in the probability of (upward and downward) wage mobility. Understanding of wage mobility patterns and trends may also be helpful in designing pension schemes and other benefit or payment schemes that are contingent on earnings over the life cycle.

In Portugal, wage dispersion increased sharply during the 'eighties and the first half of the 'nineties, but no evidence has so far been gathered on wage mobility. A comparison with the UK could provide interesting insights, checking whether the common trends concerning wage inequality also applied to wage mobility, despite the contrasts in labour market institutions. The British labour market is recognized as a very flexible one in Europe, whereas Portugal presents a more regulated framework and a more centralized wage bargaining system, which could lead to lower wage mobility. In fact, collective bargaining sets wages for detailed categories of workers, defining specific wage hierarchies. Since the extension of contracts to non-unionized workers is widespread, most of the workforce is covered by collective bargaining and could therefore be subject to a rigid wage hierarchy. Nevertheless, high wage flexibility, with wages responding to the business cycle, namely the unemployment rate, and to conditions prevailing at the firm level, have been pointed out as characteristics of the Portuguese labour market, which could operate in the opposite direction.

The aim of this study is therefore to add comparable cross-country evidence on earnings mobility, discussing the impact of institutional differences in the labour market. Results on Portugal will be compared in particular with the UK, following closely the methodology by Dickens (2000) and, whenever feasible, with other countries, such as

France. The major questions to be answered are: What is the extent of wage mobility in Portugal, and how does it compare to other countries? Is a more strongly regulated labour market more likely to protect low-wage workers, sustaining their wages, while nevertheless reducing their chances of upward mobility, and reducing overall wage mobility in the economy? Are the opportunities for mobility larger in the longer-run than in the short-term, also in Portugal? How has wage mobility evolved over time in a country marked by sharp growth rates during the second half of the 1980s and by severe shortage of skills? Has it shared the decline in mobility that has taken place in the UK and in France? Which institutional forces can drive the similarities and differences between Portugal and other countries, concerning patterns, levels and trends in wage mobility?

The study relies on a remarkable longitudinal matched employer-employee data set, covering all the wage-earners in manufacturing and services private sector in Portugal. Problems commonly faced by panel data sets, such as under- or over-sampling of certain groups, and panel attrition, are thus reduced. Also, the legal requirement for the data to be posted in the company contributes to its reliability, reducing measurement errors.

Section 2 provides details on the data source and concepts used. Section 3 overviews trends in wage dispersion. Section 4 detects patterns in wage mobility, comparing it among groups of workers with different wage levels, and over different time horizons. Trends in short-term wage mobility are then analyzed, first for the overall labour market, in section 5, and then separately for groups of workers with different wage levels, in section 6. Section 7 deals with trends in longer-term wage mobility. Section 8 discusses some of the institutional forces that may have driven the similarities and contrasts identified between the Portugal and other economies, before concluding comments are presented in the final section.

2. Data Source and Concepts Used

2.1. Data Set

A longitudinal data set matching firms and workers in the Portuguese economy, was used. The data are gathered annually by the Ministry of Employment, based on an inquiry that every establishment with wage-earners is legally obliged to fill in. Reported data cover all the personnel working for the establishment in a reference week (in March until 1993, and in October from 1994 onwards). Public administration and domestic service are not covered, whereas the coverage of agriculture is low, given its low share of wage-earners. For the remaining sectors, the mandatory nature of the survey leads to an extremely high response rate, and in fact the population of firms with wage-earners in manufacturing and the services private sector is covered.

Reported data on the worker include the gender, age, schooling, monthly earnings (split into several components) and duration of work. A worker identification code, based on a transformation of the social security number, enables tracking him/her over time. The panel data set built covers the period 1986 to 1999[1]. Extensive checks have been performed to guarantee the accuracy of the data, using the variables gender, date of birth and highest schooling level achieved, according to the following steps: selection of valid worker

[1] Note, however, that no worker data were reported for the year 1990.

identification codes; handling of duplicate worker codes in a given year; correction of missing values when reported data for the rest of the periods was absolutely consistent; correction of inconsistent data for the same worker across years; dropping workers with remaining missing data on gender, age or schooling (see the appendix for a more detailed description of the procedure followed).

After the checks on the worker identification code, on average 87 percent of the original yearly data sets have been kept for analysis, yielding an initial panel of 5,537,232 workers and 25,134,881 observations worker-year. After the remaining data checks, the final panel included 22,444,601 observations worker-year and 4,882,908 workers, having therefore kept 89 percent and 88 percent of the initial panel, respectively.

2.2. Population Under Analysis

Following Dickens (2000), the analysis is carried out separately for males and females, and it focuses on workers aged 22 to 59. Table 1 reports, for each year, the number of workers included in the panel. For consistency with Dickens (2000), reported workers with missing wage (because they are not wage-earners, but instead owners of the company, unpaid family members or members of cooperatives, according to the Portuguese classification) were kept in the analysis of transition matrices.

Table 1

Size of the Data Set Under Analysis

	Males		Females	
year	wage-earner	missing wage	wage-earner	missing wage
1986	679.751	94.104	335.062	46.326
1987	683.846	95.309	342.133	49.679
1988	708.288	103.053	367.786	51.077
1989	756.500	113.269	410.058	55.870
1991	789.935	120.915	458.624	66.022
1992	806.977	123.114	484.945	70.295
1993	803.051	119.113	494.921	65.267
1994	811.444	119.254	535.842	64.823
1995	838.239	122.020	565.767	68.450
1996	838.151	122.158	579.290	70.031
1997	887.284	112.836	631.229	60.997
1998	909.893	119.717	655.933	70.610
1999	950.549	127.396	702.173	77.474

Source: Computations based on Portugal, MTSS (1986-1999).
Note: Workers who are not wage-earners (but instead owners of the company, unpaid family members or members of cooperatives) were assigned missing wages.

Unlike in the study by Dickens, data on unemployed workers is not available. Therefore, once a worker drops out of the data base, one of several situations may have occurred: (s)he may have become unemployed, inactive, or an independent worker, may have joined an uncovered sector (Public Administration) or the Army. Even though unemployment is the most likely situation, when analyzing transition matrices, the cell

unemployed must be replaced with out of the data base, and caution exerted in comparing the results between the UK and Portugal.

2.3. Wage concept

Gross actual hourly earnings were computed as

$$hw = \frac{bw + sen + reg + overtw}{normh + overth}, \tag{1}$$

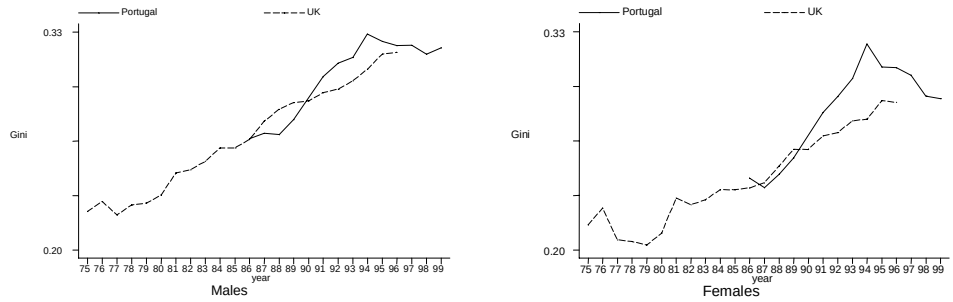
where bw stands for base-wage, sen are seniority-indexed components of pay, reg are other regularly paid components and $overtw$ is payment for overtime; $normh$ refers to the normal duration of work and $overth$ are overtime hours.

3. Trends in wage dispersion

The Gini index is the measure of inequality most widely used in the literature, taking account of the whole distribution of wages to compute a synthesis indicator of the distance among workers' wages. It ranges between 0 (complete equity, if everyone earned the same wage) and 1 (complete inequity, if one single worker were to earn the whole wage bill). Figure 1 reports this index for gross hourly earnings in Portugal and the UK, computed in either case from employer reported worker level data. Even though data for the UK are available for a wider period that includes the 1970s, comparison between the two countries is performed after 1985.

Figure 1

Trends in Wage Dispersion, Portugal and the UK.



Source: For Portugal – Computations based on Portugal, MTSS (1986-1999); for the UK – Machin (1988).
Note: For Portugal, workers aged 22 to 59; no age limits are mentioned for the UK. The values for Portugal in 1990 result from linear extrapolation of the terend, since no worker data are available.

The similarity between the level of wage dispersion in Portugal and the UK in 1986 is outstanding.[2] Moreover, in both countries inequality increased very sharply until

[2] Using data from the Family Expenditure Survey, instead of the New Earnings Survey, both reported in Machin (998), indicates a similar level of wage dispersion in Portugal and the UK in 1994, with lower departing values in Portugal in 1986 and a much faster increase in inequality during that period, specially among women.

mid-90s. However, in Portugal it increased at a faster pace up to 1994, but declined afterwards, whereas in the UK it kept increasing.

Figure 2 considers separately dispersion at the bottom and at the top halves of the distribution. It reports the ratios between: the wage of the person who is in the middle of the wage distribution (percentile 50 or the median), to the wage of the person who has only 10 percent of the workers in the economy paid less than him/her (percentile 10); the wage of the person who has only 10 percent of the workers paid above him/her (percentile 90), to the middle wage.

The contrast between the pattern of wage dispersion in the UK and Portugal is striking. The bottom part of the distribution is much more compressed in Portugal, whereas, conversely, the top is much more stretched. Note that, in Portugal in 1986, the male at the middle of the wage distribution earned one and a half times as much as the 10 percentile male, whereas in the UK he earned almost twice as much. On the other hand, a worker at the top, at the 90 percentile in Portugal, earned 2.2 times the wage of the worker in the middle of the distribution, whereas in the UK that ratio was 1.9. Moreover, by 1994, the ratio had grown to 2.6 in Portugal and remained at 1.9 in the UK. Similar levels of overall wage dispersion therefore result from a very different pattern, with a compressed bottom and a stretched top in Portugal. The compression at the bottom reflects the impact of the national minimum wage, pulling low-wages closer to the middle of the distribution.[3] The wider dispersion at the bottom in the UK is a major reason justifying why concern with declining wage mobility has focused so much on its implications for low-wage workers.

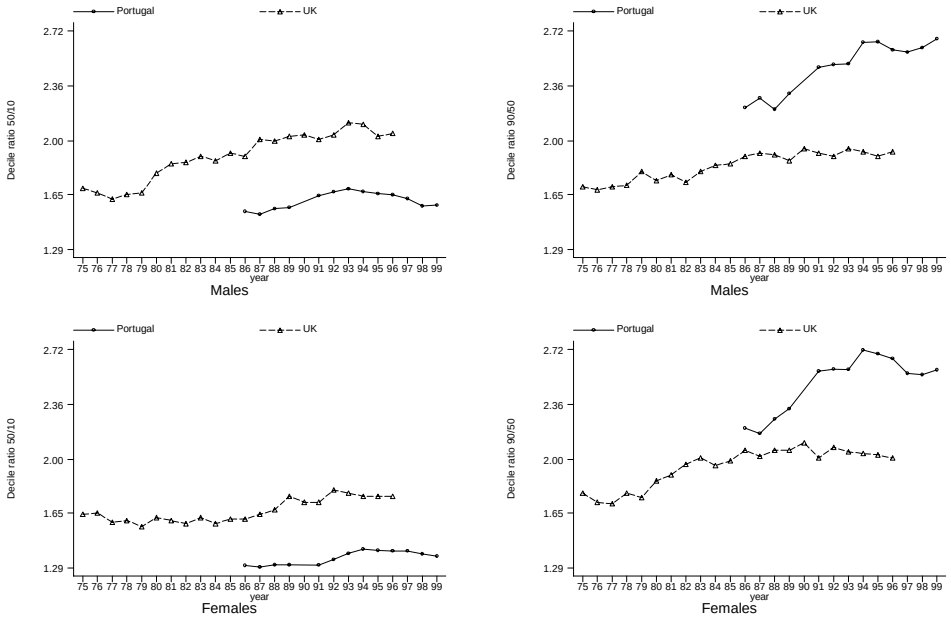
In the UK, the wage distribution became more stretched, both at the top and at the bottom. The same trend took place, but at a faster pace, in Portugal until mid-90s, but it was reversed afterwards. In fact, until mid-90s, the lowest wages were increasingly left behind relative to the middle wages, and the top wages jumped ahead, growing faster than middle wages. To be in the top decile in Portugal therefore meant earning a wage that was higher in 1994 than in 1986 relative to the median. Conversely, being in the lowest decile meant earning a wage that was in 1994[4] lower than in 1986, when compared to the median wage. On the contrary, from 1994 onwards, the Portuguese wage distribution embarked on a compression process, both at the bottom and at the top, as both the lowest and the highest wages in the economy got closer to the middle of the distribution. Nevertheless, this trend was not enough to offset the rise in inequality that had taken place until mid-90s.

[3] In the UK, the national minimum wage was introduced in 1999, and its impact is therefore beyond the time frame of the data presented.

[4] For males, 1993.

Figure 2

Trends in Wage Dispersion at the Bottom and Top of the Distributions, Potyugal and the UK



Source: For Portugal – computations based on Portugal, MTSS (1986-1999); for thr UK – Machin (1988).
Note: For Portugal, workers aged 22 to 59; no age limits are mentioned for the UK.

4. Patterns of Wage Mobility

4.1. Transition Matrices

The previous section relied on a static perspective of the wage distribution, comparing snapshots taken at different moments in time, which reported a measure of the distance between workers' wages. Introducing the time dimension to follow a worker's wage path would provide a more complete picture of the wage distribution or, in fact, it would provide the movie instead of a series of pictures.

To compare the level of mobility over different time horizons, Dickens focused on the following time spans: 1993-94 (one-year), 1991-94 (three-year) and 1989-94 (five-year). When choosing the time spans under analysis for Portugal, two constraints should be kept in mind: in 1994, the reference week for the inquiry changed from March to October, and therefore 1993-94 captures changes over a longer period, 19 months, rendering it infeasible for comparison with the remaining periods; in 1990, no worker data were reported. Given these constraints, to compare levels of mobility over different time horizons, the following time spans will be analyzed for Portugal: 1991-92 (one-year), 1989-92 (three-year) and 1987-92 (five-year), thus sliding backwards Dickens's windows by two years.

Decile transition matrices provide a description of wage mobility patterns. To build them, consider the wage distribution in a given year and rank the workers according to the wage they receive, from the worst to the best paid in the economy. Then partition this sequence of workers into ten groups of equal size deciles , with the first group including the ten percent worst paid workers in the economy, and the last group including the ten percent best paid ones. The same exercise is repeated for a different year. Transition matrices report the share of workers in a given decile in the initial year that moved to a given decile in the final year. They therefore indicate the extent of mobility in the wage distribution, revealing, not just the share of workers who changed position, but also how far they traveled. If absolute immobility were the rule, then the diagonal which reports the share of workers who were in a particular decile in the initial period, and in that same decile in the final year would equal 100, and no other positive values would be found on the matrix. Similarly, low mobility would be represented by a concentration of values along or close to the diagonal.

4.2. Mobility By Wage Group

Mobility patterns in the Portuguese labour market are very similar to those that have been detected by Dickens for the UK.

First of all, the mobility matrix in table 2 for males from 1991 to 1992 in Portugal reports a high degree of earnings persistence, with the proportion of workers who does not change decile ranging from 31 percent to 66 percent, depending on the decile of origin. Most of those who travel in the wage distribution, do so a short distance. Immobility is apparently higher for the lower and upper tails of the distribution, a pattern found, not just for the UK, but also for Germany and the USA (Burkhauser et al, 1997). Consider first workers in the lowest wage bracket. Out of those: 41 percent remain in that group one year later; only 22 percent move up the wage distribution, with over half of these climbing up just one decile; hardly any of these workers makes it to the upper half of the distribution; one third drops out of the database, with unemployment or inactivity as likely outcomes. As to the group of the best paid workers in the economy, persistence is even higher, with two thirds of them remaining in that situation one year later; most of those moving down, travel only one decile. In the middle deciles, on the other hand, immobility rates are lower, at approximately 31 to 38 percent. It should nevertheless be noted that part of the immobility detected for the tails of the distribution results from the methodology used. Due to the shape of the earnings distribution, the upper decile covers a wider range of wages, and therefore it would take a larger wage variation to move the worker into a different decile; also, at either end of the distribution, there is less room to escape, as workers are bound to move only in one direction.

Table 2

Transition matrix, males, Portugal, 1991-92.

	State in 92											
State in 91	out db*	miss w*	1	2	3	4	5	6	7	8	9	10
out db*		16.1	13.9	11.7	10.7	9.7	8.5	7.7	6.2	5.4	4.8	5.3
miss w*	32.4	47.6	2.3	2.0	2.0	2.0	2.1	1.9	1.8	1.8	2.2	2.1
1	32.8	4.4	41.4	12	3.7	2.0	1.2	0.8	0.5	0.3	0.3	0.8
2	29.2	3.9	10.4	33.7	12.8	4.6	2.3	1.3	0.8	0.4	0.3	0.5
3	27.4	3.7	2.1	12.4	31.1	13.8	4.8	2.1	1.1	0.7	0.4	0.5
4	25.6	3.5	1.1	3.0	13.4	31.2	14.2	4.4	1.7	0.9	0.5	0.6
5	22.9	3.6	0.7	1.4	3.6	13.5	33.7	14.1	3.6	1.4	0.7	0.6
6	20.8	3.5	0.5	0.8	1.4	2.9	13.1	37.5	14	3.6	1.2	0.7
7	17.7	3.5	0.4	0.5	0.7	1.1	2.6	12.7	43	13.9	3.0	1.1
8	15.4	3.3	0.2	0.3	0.5	0.7	0.9	2.4	13.4	47.9	12.7	2.3
9	14.0	2.7	0.2	0.2	0.2	0.4	0.5	0.9	2.0	12.7	55.3	10.9
10	15.1	3.5	0.5	0.4	0.4	0.5	0.6	0.6	0.9	1.8	10.4	65.3

Source: Computations based on Portugal, MTSS (1991-1992).
Notes: (*) out db – out of the database; w -missing wage (workers who are not wage-earners, but instead owners of the company, unpaid family members or members of cooperatives.

Table 3

Transition matrix, females, Portugal, 1991-92.

	State in 92											
State in 91	out db*	miss w*	1	2	3	4	5	6	7	8	9	10
out db*		13.2	13.5	10.3	9.5	8.8	8.7	8.7	7.9	7.6	6.2	5.7
miss w*	34.0	37.1	3.2	3.0	3.2	3.0	3.0	2.9	3.1	2.7	2.4	2.5
1	33.9	6.5	33.2	11.7	5.8	2.8	2.3	1.4	1.0	0.6	0.3	0.6
2	29.7	6.0	10.5	30.4	11.6	5.4	2.9	1.5	0.8	0.5	0.2	0.6
3	26.2	6.1	4.4	11.4	27.5	13.8	5.7	2.6	1.0	0.5	0.3	0.5
4	24.9	5.9	1.5	5.7	13.3	28.8	12.2	4.5	1.6	0.8	0.4	0.5
5	24.1	5.6	1.3	2.4	5.8	13.0	29.7	12	3.6	1.3	0.6	0.6
6	24.1	5.3	0.9	1.2	1.7	4.0	13.2	32.5	12.6	3.0	0.8	0.7
7	22.9	5.0	0.6	0.5	0.8	1.0	2.6	13.2	38.6	12.2	2.0	0.8
8	21.4	4.9	0.4	0.3	0.4	0.5	0.8	2.0	12.1	45.3	10.3	1.6
9	18.1	4.1	0.3	0.1	0.2	0.2	0.4	0.6	1.3	9.1	56.4	9.1
10	15.9	3.5	0.5	0.3	0.4	0.4	0.5	0.5	0.6	1.2	9.2	67.0

Source: Computations based on Portugal, MTSS(1991-1992).
Notes: (*) out db - miss w - missing wage (workers who are not wage-earners, but instead owners of the company, unpaid Family members or members of cooperatives).

Secondly, similar to the UK, the lowest paid workers are more likely to end up, the following year, in a situation of unemployment.[5] Likewise, those that were unemployed/out of the data base, are more likely to enter employment with low wages.

All of these patterns hold for females as well, even though persistence in the lowest decile is not as high, just like in the UK, with one fourth of the women moving up the wage distribution over the period of one year in Portugal.

4.3. Mobility over different time horizons

Over longer time horizons, the opportunities for wage variation and therefore for workers to swap positions in the wage distribution are higher. So, mobility over a three-year period is higher than over one year, and it is still higher over a five-year period (see tables 6 to 9 in appendix). This pattern of increasing mobility over longer time periods is expectable, and it holds for several countries, for example Germany and the USA (Burkhauser et al, 1997) or Austria (Hofer and Weber, 2002), besides the UK (Dickens, 2000).

Over three-years, the immobility rates are in Portugal 23 percent, 45 percent and 19 percent, respectively for the lowest, top and middle deciles of the male distribution. Comparable values for females are 17 percent, 49 percent and 15 percent, with females again revealing a higher probability to escape low wages. Transitions over a five-year period are still higher. For the first, top and middle deciles, immobility rates for males are, respectively, 15 percent, 38 percent and 12 percent. For females, those values are 11 percent, 41 percent and 13 percent, respectively.

Whereas mobility patterns have been compared between the UK and Portugal, comparison of mobility levels across countries using the transition matrices previously reported would be inaccurate, due to the different definition of the cell unemployed in the UK and out of the data base in Portugal. Another approach must therefore be followed.[6] Moreover, it has been widely acknowledged that cross-country comparisons of levels of wage mobility are hard to perform, since the ranking of countries is highly sensitive to the (inequality) measure used (see for example OECD (1997) or Hofer and Weber (2002))[7]. Exactly the same approach should thus be used for both countries.

5. A More Complete Picture: Trends in Wage Mobility Based on Individual Ranks in the Distribution

To check changes in wage mobility over time and make comparisons of both levels and trends across countries, the position of individuals in the wage distribution will not be aggregated into ten groups, but instead each worker will be assigned his own wage position, n_i : from 1, the lowest paid worker in the economy, to N , the best paid one. Instead of allowing the measure to fall in the range 1 to N , it will be standardized to fall

5 More rigorously, out of the data base in Portugal.

6 Even though it can be remarked that the levels of immobility for different deciles are very similar in Portugal and the UK when evaluated using the British Household Panel Survey.

7 Their approach follows Shorrocks (1978), to measure mobility as the extent to which an inequality index is reduced when wages are averaged over a longer time horizon.

between 0 and 1, being computed as $rank_i = n_i / N$. A synthesis measure M will now translate the degree of mobility existing in the wage distribution (Dickens, 2000)[8]:

$$M = \frac{2\sum_{i=1}^{N} | rank_{it} - rank_{is} |}{N} \quad (2)$$

where the subscripts t and s refer to the year and N refers to the total number of workers. M is twice the average absolute jump in rank (percentile), and it ranges between 0 (no mobility) and 1 (maximum mobility). It is computed over workers who are wage-earners in both periods under analysis, therefore overcoming the limitation of having a cell that has a different meaning in the previous transition matrices.

5.1. Declining mobility

Comparison between the UK and Portugal for the overlapping period (1987 to 1994) highlights, once again, a notable similarity mobility levels were very similar at the start, either for males or females (see figure 3). Also, in both countries mobility has declined, for males and for females. However, the pace of decline was faster in the UK, such that by 1994 the degree of mobility was considerably lower there. It can however be noted that male wage mobility kept declining in Portugal, to reach in 1996 values almost as low as the UK two years earlier.[9]

Concentrating on Portugal, it is noticeable that from mid-80s to late-90s wage mobility declined for males, having fluctuated sharply for females. However, within the period, different trends can be identified. Up to 1993, mobility in the male earnings distribution fluctuated around 0.15, whereas after 1994 it fluctuated around 0.13, therefore revealing a noticeable decline at mid-90s. In the female wage distribution, the final value was slightly lower than the initial one, with a peak in 1993.

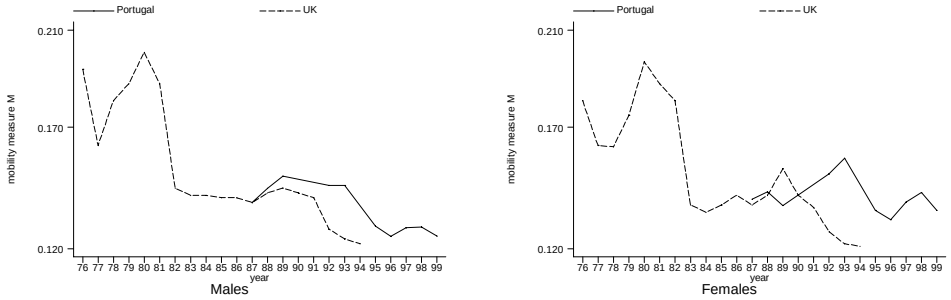
---

8 Since wages tend to rise over the life cycle, in Dickens (2000) part of the analysis of mobility is conducted removing the age effect from wages, so that actual mobility is evaluated after accounting for the life cycle trend. For comparability, the analysis from here onwards also considers age-adjusted wages, by using the residuals from a wage regression on age dummies. The patterns and trends are not affected by this change, but the mobility levels are lower.

9 Unfortunately, we do not have information on the trend in the UK from 1994 to 1999.

Figure 3

Trends in Wage Mobility, Portugal and the UK.



Source: For Portugal – Computations based on Portugal, MTSS (1986-1999); for the UK – Dickens (1997).
Note: For Portugal, information for 1989-90 are not available, and for 1993-94 it has not been used in the plots. Thus the values for those periods result from linear extrapolation of the trend.

A curious aspect in Portugal therefore concerns the turning point at mid-90s in the evolution of the male wage distribution. Until 1994, wage inequality was rising, and afterwards it has been declining; until 1993, mobility was high, and after 1994 it has been lower. Therefore, rising inequality was associated with relatively high mobility, whereas declining inequality coexisted with lower mobility. Information on a wider time horizon would be required to enable strong statements about this link between changes in inequality and the level of mobility.

Nevertheless, there is indication that the implications of rising inequality may have been different in Portugal and the UK. In the UK, there is strong evidence that rising differences among workers at a moment in time are now inflated over the life cycle, since the opportunities for positions to swap, declined (rising inequality and declining mobility throughout the period). Evidence on Portugal is less compelling. Despite the rise in inequality and the decline in mobility between the initial and final years under analysis, the evolution within that period suggests that, differently from the UK, rising inequality was accompanied by high mobility, and declining inequality was accompanied by lower mobility.

The crucial question that follows is therefore: Which workers were enjoying the higher mobility that existed until 1993? The low-paid workers? The top wage-earners? This is exactly the issue addressed in the next section.

5.2. Which workers experienced much higher mobility in Portugal until mid-90s than afterwards? Comparing trends across countries

If the high mobility existing until mid-90s was mainly driven by low-wage workers moving upwards, then, despite the deterioration of their relative wage level, the lifetime prospects for someone entering the labour market at a low wage would tend to get brighter after entry. On the contrary, if the lowest wage workers were not enjoying that higher overall mobility, then, at best, they would be trapped to remain in a wage level that was deteriorating relative to the rest of the economy; at worst, they would be cycling

between low employment and no employment. The latter situation would deserve more attention from policy makers.

What about top wage workers? If the mobility of top wage workers were higher until mid-90s than afterwards, it would mean that reaching the top entitled the worker to a rising premium, but at a price the risk of falling down again (higher than after mid-90s). On the contrary, being as immobile before 1993 as afterwards, would mean that, once the top was reached, the worker would stay there and become increasingly better off relative to the rest of the working population.

Table 4 explores which groups of workers experienced gains in mobility when wage inequality was increasing. It reports the yearly mobility index M by decile of origin, distinguishing between the period when earnings inequality was rising and the period when it was declining.[10] The last column reports the change between the two periods in the average yearly mobility index.

Over all the working population considered, the table confirms the patterns and trends previously detected: mobility is higher in the middle of the distribution while, on the contrary, higher persistence is a characteristic of the bottom and top deciles; yearly mobility in the male wage distribution was, on average, 15 percent higher until mid-90s than in the period that followed; in the female distribution, it was 8 percent higher until mid-90s. That trend was common to every decile in the male wage distribution, and almost every decile in the females'.

However, table 4 reports a very clear pattern regarding the identification of the groups of workers who experienced much higher mobility until mid-90s than afterwards. For the half best paid group of men in the economy, mobility was until mid-90s considerably higher than afterwards (around 20 percent higher). Their earnings instability was therefore higher during that period, with a higher probability of swaps in relative positions.[11] Males on lowest wages enjoyed, until 1993, somewhat higher mobility than afterwards (9 percent higher, for the first decile). Therefore, while the lowest wages were deteriorating relative to the rest of the male labour market, the possibility of escaping that situation was not remarkably better than after mid-90s. For the female wage distribution, the decline in mobility was more homogenous across deciles.

Also for France, wage mobility declined (during the 1970s and 1980s). The analysis by Buchinsky, Fougere and Kramarz (1998) identified the group of workers most affected by the decline in wage mobility those holding a higher education diploma. Simultaneously, they were experiencing the highest wage growth. Unfortunately, the study does not progress to explain the mechanisms generating these outcomes.

Repeating for Portugal the computations in table 4, this time separately by schooling levels (below Higher Education and Higher Education diploma), shows that the reduction in wage mobility has hit in particular the workers with highest schooling in the top half of the wage distribution[12].

---

10 In 1987, inequality was not yet increasing for females, and in 1999 it was no longer decreasing for males, so those periods were not included.

11. Note however that this did not affect the very best paid males (top decile). Indeed, those in the last decile enjoyed until mid-90s, not just a remarkable increase in their relative wages but, once that stage had been reached, considerable stability, as the mobility rates are, not just very low, but close to their post-93 levels.

12 Table available from the author on request.

Therefore, whereas in France and Portugal the holders of Higher Education diplomas saw their wage mobility decline the most, in the UK mobility declined especially for low wage workers. A similar result has been reported for the USA (Buchinsky and Hunt, 1999). Section 7 discusses possible causes for these contrasts.

As to the level of mobility one finds, relying on information for the period 1988-89, that overall mobility was very similar for males in the UK and Portugal (respectively 0.145 and 0.150), as figure 3 had already reported. Nevertheless, mobility at the tails is slightly higher in Portugal. Indeed, whereas for deciles 2 to 8, the difference between both countries mobility rates never exceeds 0.01[13], for the first decile and the two top ones, it does. For females, overall mobility during that period was higher in the UK, but nevertheless, women on the lowest wages in Portugal (deciles 1 and 2) were slightly more likely to escape that situation. The probability of escaping low wages therefore seems to be slightly higher in Portugal than in the UK, for either males or females, contrary to what a more regulated framework would lead one to expect.

[13] Comparing table with Dickens (2000: table 4).

Table 4

One-Year Mobility Index M, By Decile of Origin, Portugal

MALES	Period rising inequality			Period declining inequality				change in av. mobility
decile origin	88-89	91-92	92-93	94-95	95-96	96-97	97-98	between the two periods
All	0.15	0.146	0.146	0.129	0.125	0.129	0.129	1.151
1	0.128	0.129	0.131	0.114	0.12	0.111	0.131	1.087
2	0.157	0.151	0.153	0.132	0.138	0.138	0.151	1.098
3	0.178	0.165	0.164	0.144	0.147	0.151	0.158	1.126
4	0.186	0.171	0.168	0.147	0.149	0.154	0.157	1.15
5	0.185	0.172	0.169	0.144	0.146	0.15	0.151	1.185
6	0.175	0.165	0.164	0.137	0.135	0.142	0.14	1.212
7	0.157	0.151	0.148	0.13	0.124	0.132	0.128	1.184
8	0.134	0.135	0.135	0.116	0.11	0.115	0.108	1.197
9	0.11	0.109	0.109	0.094	0.087	0.091	0.086	1.222
10	0.088	0.114	0.122	0.136	0.094	0.101	0.079	1.054
FEMALES	Period rising inequality			Period declining inequality				change in av. mobility
decile origin	88-89	91-92	92-93	94-95	95-96	96-97	97-98	between the two periods
All	0.138	0.151	0.157	0.136	0.132	0.139	0.143	1.081
1	0.157	0.169	0.175	0.152	0.152	0.149	0.179	1.058
2	0.163	0.175	0.184	0.149	0.153	0.163	0.178	1.083
3	0.161	0.186	0.198	0.155	0.157	0.17	0.183	1.092
4	0.171	0.193	0.198	0.162	0.164	0.177	0.182	1.094
5	0.167	0.183	0.189	0.152	0.156	0.166	0.168	1.119
6	0.152	0.166	0.166	0.138	0.142	0.148	0.153	1.112
7	0.135	0.138	0.143	0.121	0.127	0.131	0.13	1.091
8	0.116	0.117	0.119	0.107	0.107	0.113	0.109	1.077
9	0.087	0.083	0.087	0.087	0.076	0.079	0.074	1.084
10	0.069	0.098	0.113	0.136	0.085	0.096	0.075	0.95

Source: Computations based on Portugal, MTSS (1988-1998).
Notes: For 1989-90 and 1990-91; mobility measures cannot be computed because no worker data are available for 1990; similarly, for 1993-94, mobility measures are not presented, because they refer to wider time horizon, and are thus not comparable to the rest of the series. Different from Dickens (2000), no standard-errors were computed, because the set includes the population of the wage-earners in the private sector.

5.3. Trends in longer-term mobility

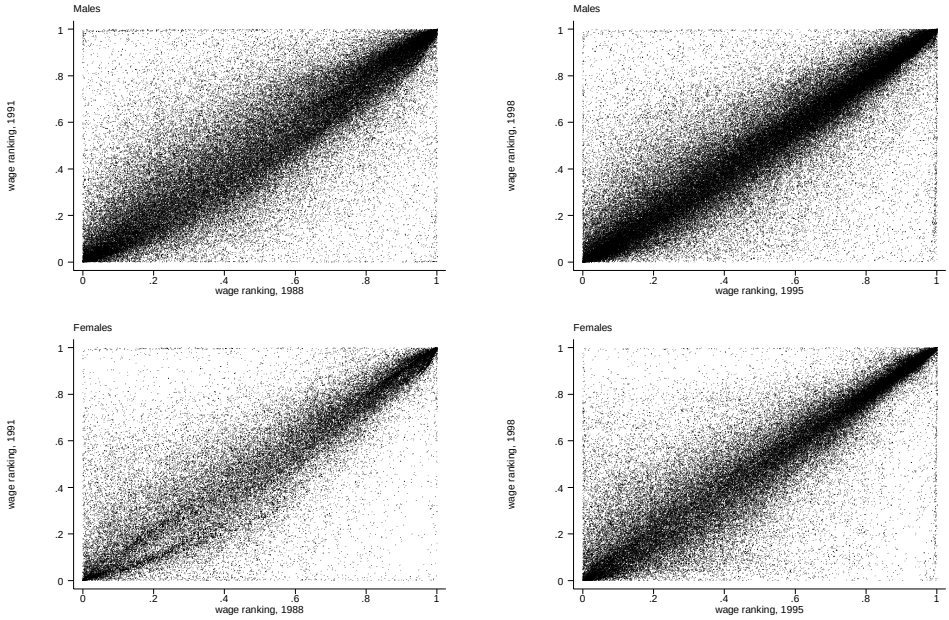
The previous sections have concentrated on changes in the level of mobility occurring over one-year intervals. This section focuses on trends in longer-term mobility, over three and five years. To complement the mobility measure M, this part of the study uses simple inspection of an informative graphical device.

Consider the wage ranks of each worker, in the initial and in the final year. Then plot, on the horizontal axis, the rank of the worker in the initial year, and on the vertical axis his/her rank in the final year. If the worker remained in the same relative position (rank), then (s)he would be represented by a dot on the diagonal of the graph. A higher concentration of points along or close to the diagonal therefore provides indication of lower mobility. On the contrary, if the wage earned by a worker in a particular year were

unrelated to his/her wage in a different year, the dots would be scattered randomly all over the graph. Figure compares mobility in the three-year periods 1988-91 and 1995-98.[14] For longer-term mobility, over five years, the periods 1988-93 and 1993-98 are compared in figure .[15]

Figure 4

Three-Year Mobility, Males and Females, Portugal



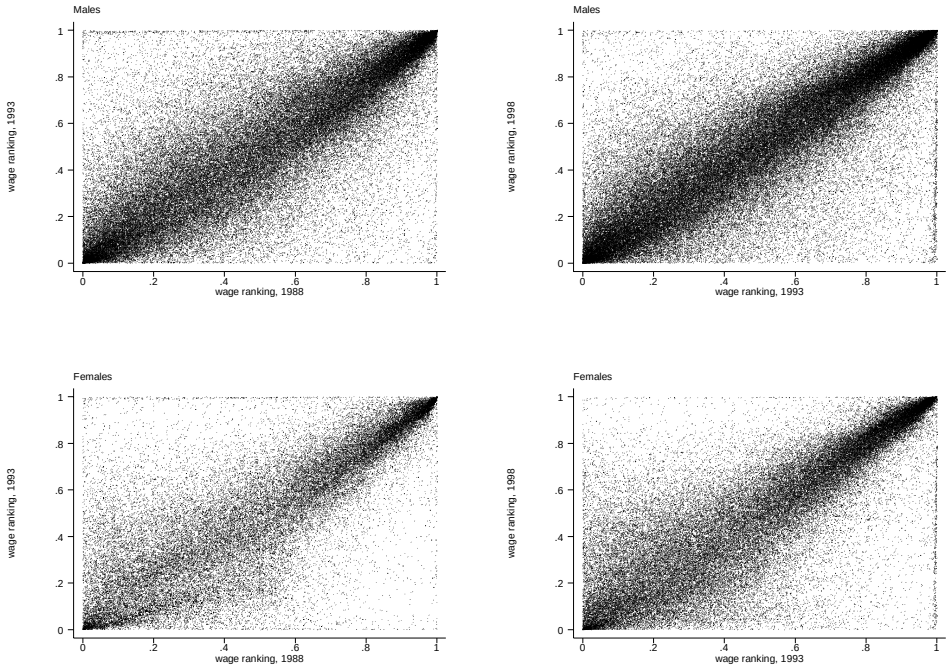
Source: Computations based on Portugal, MTSS (1988-1998).
Note: The plots refer to 25 per cent samples of the data set.

14 The requirement to compare periods with similar inflation rates is less of a constraint for Portugal than it was for the UK. In fact, Dickens highlights that, having wages at a particular point in time, the timing of wage increases is very relevant, in particular in periods of high inflation, when nominal wage adjustments will be larger, and therefore having workers whose wage has not yet been updated, together with workers whose wage has already been updated, provides a misleading idea of the true level of wage mobility. However, in Portugal, most wages are bargained over annually, with the wage change becoming effective almost invariably in January, such that the timing of wage bargaining does not undergo change depending on the inflation rate. Indeed, in periods of very high inflation such as the 80s, the Government has set minimum periods for the duration of each collective agreement: from 1980 to 1989 it was imposed, as a general rule, that any collective agreement signed should last for at least two years, while its wage clauses could be updated annually; from 1989 onwards, the general rule is the annual duration of the agreements. On the other hand, all the reported data refer to the same reference week.

15 Note however that the period 1993 to 1998 covers more than five years, because the inquiry took place in March 93 and October 98, therefore overstating the mobility that would have taken place over a five-year horizon.

Figure 5

Five-year Mobility, Males and Females, Portugal



Source: Computations based on Portugal, MTSS (1988-1998).
Note: The plots refer to 25 per cent samples of the data sets.

It can be noted, in figure 4, that the concentration of dots along the diagonal increased from 1988-91 to 1995-98, both for males and females. Just like short-term mobility, that over a three-year horizon also seems to have declined. The same trend holds for longer-term mobility, over five years. Thus, between mid-80s and late-90s, wage mobility declined in Portugal, irrespective of the time horizon considered (one-, three- or five-year mobility). This result is confirmed by the mobility measures reported in table 5.

Table 5:

Mobility Over Three- and Five-Year Periods, Males and Females, Portugal

	1988-91	1995-98	1988-93	1993-98
Males	0.230	0.179	0.260	0.231
Females	0.217	0.190	0.256	0.244

Source: Computations based on Portugal, MTSS (1987-1998).

6. *Discussing the impact of different institutions on labour market inequality and mobility*

Major similarities and some contrasts have been found between the British and the Portuguese wage distributions, which call for some discussion on underlying causes and the role played by institutional forces. The limited evidence available for France can as well contribute to the discussion.

First of all, both the UK and Portugal became more unequal labour markets, having departed from similar levels of wage inequality in mid-80s. However, inequality at the bottom is striking in the UK, whereas in Portugal it is striking at the top. In other words, in the UK the poor fare – worst relative to the rest of the economy – than in Portugal. Conversely, in Portugal the rich fare better – relative to the rest of the economy – than in the UK.

Secondly, the possibility to escape low wages seems to be higher in Portugal than in the UK, either for males or females.

Thirdly, both countries became less mobile labour markets, having departed from similar levels of wage mobility in mid-80s. A similar trend of declining wage mobility has taken place in France.

Fourth, whereas in the UK rising inequality and declining mobility seem to have been uninterrupted trends, Portugal presented in mid-90s a turning point. Until then, rising inequality was associated with high wage mobility; afterwards, declining inequality coexisted with lower wage mobility. Mobility declined in particular for workers in the top half of the wage distribution and, among these, those holding a Higher Education diploma. This pattern resembles France, where Buchinsky, Fougere and Kramarz (1998) have detected that the declining wage mobility was sharpest for workers holding a University diploma.

Differences between the institutional framework in Portugal and France, on one hand, and the UK, on the other hand, can contribute to explain these differences in outcomes. The minimum wage certainly had a role compressing the bottom of the wage distribution in Portugal, as it sustained the wages of less skilled workers, leading to a remarkable contrast with the UK in the pattern of wage dispersion.

On the other hand, collective bargaining might be expected to lead to a rigid wage hierarchy and low mobility in Portugal, as it defines strict wage hierarchies that are then applied to both unionized and non-unionized workers. Also, stricter regulations on employment protection might be expected to lead to lower worker turnover, therefore reducing one major source of wage mobility, the transition between employers.

However, evidence suggests instead that these different institutional architectures did not lead to sharp contrasts in patterns, levels or trends in wage mobility. In particular, mobility out of low-wages is slightly higher in Portugal than in the UK. Persistence in a situation of minimum wage does not seem to be in Portugal any higher than persistence in a situation of lowest wages in the UK economy.

Part of the explanation may be that in fact, despite the regulated wage bargaining system, wages are in Portugal quite flexible. So does the OECD (1992) highlight, when finding that, at the macro level, wages in Portugal are among the most responsive to changes in the unemployment rate, and so do studies of wage adjustment at the micro level indicate. Indeed, Cardoso and Portugal (2003) found that wage adjustments by firms, beyond the levels defined by collective bargaining, provide ample scope for the firm to

adapt the wage levels to firm- and worker-specific conditions. Similarly, Cardoso (2000) had found that firms diverge widely in the way they reward different worker attributes.

Evidence on other cross-country comparisons of wage mobility had pointed in a similar direction. Very similar patterns of wage inequality and wage mobility had been found by Burkhauser et al (1997) for the USA and Germany, despite the sharp contrasts in their labour market institutions. Also, the OECD (1997) found that less regulation in labour and product markets, in countries such as the UK and the USA, does not result in higher levels of wage mobility.

The major difference between Portugal and the UK probably concerns the pace of economic modernization they were undergoing during the period under analysis, and the availability of a skilled workforce. There is by now a certain consensus over the idea that technological progress has been the major force determining rising wage inequality. Despite increasing supply of more skilled workers, the returns to skill, namely formal schooling, increased sharply.[16] During the second half of the 1980s, growth and modernization took place at a fast pace in the Portuguese economy, and unemployment decreased from 9 percent in 1985 to 4 percent in 1991, substantially below the European Union average. This framework of growth, in an economy with a shortage of skilled labour, led to rising wage premiums for the most skilled. The rise in wage dispersion was driven in particular by rising returns to education, as documented by previous studies (Hartog et al, 2001) (Cardoso, 1999). But more skilled workers were precisely those whose wage mobility was higher until mid-90s than afterwards. The pace of change in the Portuguese economy during the second half of the 1980s is thus likely to provide the major clue explaining both the rising inequality, and the higher mobility existing until then for skilled workers. After mid-90s, inequality increased at a slower pace in the UK, and it decreased in Portugal. In both countries, that has been probably linked to the remarkable expansion of higher education and the rising supply of a highly educated labour force, which slowed down the pace of rising returns to education.

It is interesting to note that also in France a sharp decline in wage mobility has been detected, during the 1970s and 1980s, and that the tentative explanation put forth by authors analyzing the issue also focused on the relative supply of workers holding a Higher Education degree. Indeed, Buchinsky, Fougere and Kramarz (1998) found that the younger cohorts entering the labour market with a Higher Education diploma enjoy a higher starting wage (higher decile in the wage distribution) than previous cohorts, but lower wage mobility. However, the study does not progress to an overall explanation of the forces driving these trends. The tentative explanation presented highlights the impact of the expansion of Higher Education in the reduction of the wage mobility for workers holding a University degree.

7. Conclusion

This study compared earnings mobility in Portugal and other countries, in particular the UK, for which overlapping information is available from 1986 to 1994, and extended the analysis of the Portuguese case to 1999. Common patterns and common

[16] See for example Gottschalk and Smeeding (1997) for an overview of the discussion on the forces driving rising inequality.

trends are the major message to be highlighted from the comparison between the British and Portuguese wage distributions.

Striking similarities concerning wage dispersion and wage mobility, include: the two countries shared similar levels of wage dispersion; either one became a more unequal society, as earnings inequality increased sharply, both at the top and at the bottom of the distribution; mobility is, in either country, lower at the tails of the wage distribution; women reveal nevertheless a higher probability than men to escape low-wages; mobility is higher over longer time horizons, since the opportunities for wage variation, and therefore for relative positions to change, are larger; either country became a less mobile society. The trend towards lower wage mobility was shared by France.

However, one striking difference stands out between Portugal and the UK. The similar level of earnings dispersion results from very different patterns, with the bottom of the wage distribution much more compressed in Portugal than in the UK, due to the impact of the minimum wage which at the time had not yet been introduced in the UK. The top of the distribution is much more stretched in Portugal.

Other differences, more subtle, can be identified: in the UK the increase in inequality and the decline in mobility were uninterrupted, whereas in Portugal mid-90s seems to have been a turning point. Until then, rising wage dispersion was associated with high mobility, and afterwards declining wage dispersion was accompanied by lower mobility. Workers in the top of the wage distribution were the ones who experienced the most pronounced decline in mobility after mid-90s. Further disaggregating the analysis indicates that, in Portugal like in France, workers holding a University diploma were the ones whose mobility declined the most.

Therefore, changes in mobility levels and patterns seem to have been mainly driven by broad trends taking place in these economies, namely the interaction between modernization and technological progress, on one hand, and the supply of workers holding highest schooling levels, on the other, and not so much by different institutional architectures in the labour market or their change. In fact, the evidence found does not support the idea that a more regulated institutional framework in Portugal, with minimum wages, employment protection legislation and collective bargaining widely applied, reduces the degree of individual mobility within the wage distribution.

The sharper rise in inequality during the second half of the 80s in Portugal and its slower decline in mobility are likely to be a result of the faster changes in an economy where, moreover, qualified human resources are more scarce. The period of economic growth seems to have provided opportunities for wage mobility for specific groups of workers, holding upper skills. Much of these changes were feasible under an institutional regime that, despite apparent rigidity, leaves ample scope for firm manoeuvre when bargaining over wages.

In terms of welfare implications, it can be noted that the role played by minimum wages in Portugal, the slightly higher mobility out of low-wages, the less abrupt decline in mobility over the 80s and 90s, and the fact that the decline in mobility affected mainly the higher-paid, render trends in this country less worrying than in the UK. Nonetheless, the low mobility at the bottom of the wage distribution is of concern, and the opportunities for upward wage mobility and to escape the low-employment-no-employment cycle would improve if the provision of general skills (formal schooling), training and lifelong learning were to improve.

Appendix: Panel of workers

Merging data across years

The following procedures have been implemented before merging data across years to build the panel data set on workers.

Selection of valid worker identification codes: The worker identification code is based on a transformation of the Social Security number and, after contacts with the National Social Security Office, codes made up of 6 to 10 digits were considered for analysis and subsequent data checks. This restriction led to dropping an average 9.1 percent of the observations in the original yearly data sets.

Handling of duplicate worker identification codes in a given year: Workers whose code was unique or duplicate just once, corresponding in fact to the same worker (same gender, date of birth and schooling level) reported working for two different companies, were kept for analysis. This restriction led to dropping on average each year from the original data sets: 0.4 percent of the observations, because the worker code was duplicated more than once; 3.1 percent of the observations because the worker code, though duplicate just once, did not correspond to the same worker (different gender, age or schooling); 0.2 percent of the observations because they corresponded to the same worker being mistakenly reported twice by the same company.

After the data checks on the worker identification code, on average 87 percent of the original yearly data sets have been kept for analysis, yielding an initial panel of 5,537,232 workers and 25,134,881 observations worker-year.

Checks on the consistency of data for each worker across years

Inconsistencies were identified if the worker gender or date of birth was reported changing, or if the highest schooling level achieved was reported decreasing over time. The following procedures were implemented to correct these problems.

Correction of missing values when reported data for the rest of the periods was absolutely consistent: Whenever the gender, age or education of a worker was reported consistently but missing in some year(s), it was straightforward to assign the reported value to the missing observation.[17] These corrections affected 0 percent[18], 2.2 percent and 1.6 percent of the observations in the initial panel, respectively for gender, age and schooling.

Correction of inconsistent data on gender, birth date or schooling across years: A different problem arised when the information was reported inconsistently across years. In that case, the information reported over half the times has been taken as the correct one.[19] Inconsistent values on gender were replaced, after checking that the date of birth in the observation to be corrected was the same as the most reported date of birth for that worker. A similar procedure was followed for the birth date and education, replacing inconsistent values with that reported over half the times. According to this procedure, 0.8 percent, 3

---

17 If schooling was consistently reported (possibly increasing over time) and the values it achieved before and after a missing were the same, the missing value was corrected; similarly, missings in the initial/final period(s) were extrapolated from the earliest/latest reported value.

18 Two observations.

19 Note that this requirement is more demanding than just considering the modal value as the accurate one.

percent and 5.7 percent of the observations in the initial panel have been corrected, respectively for gender, birth date and education. Workers with inconsistent data after the introduction of the previous corrections were dropped. Note that the whole information on the worker was dropped, whichever the incorrect number of observations identified: for gender, 2.3 percent of the observations and 2.1 percent of the workers in the initial panel; for the birth date, 1.4 percent and 1.8 percent, respectively; and for education, 4.5 percent and 3.5 percent, respectively.

Dropping workers with remaining missing data on gender, age or schooling: Workers with missing data after the introduction of the previous corrections were dropped. The share of observations dropped due to missing age or schooling was, respectively, 0.6 percent and 1.9 percent of the initial panel, corresponding to 3 percent and 2.2 percent of the workers.

The checked panel included 22,444,601 observations worker-year and 4,882,908 workers, having therefore kept 89 percent and 88 percent of the initial panel, respectively.

Appendix: Additional tables

Table 6

Three-year transition matrix, males, Portugal, 1989-92

	State in 92											
State in 89	out db*	miss w*	1	2	3	4	5	6	7	8	9	10
out db*		14.7	12.7	10.8	9.9	9.1	7.9	7.2	6.1	5.5	7.8	8.4
miss w*	46.7	33.2	2.1	1.8	1.8	1.9	1.8	1.8	2.6	2.5	1.7	2.1
1	46.6	4.5	23.0	11.3	5.4	3.2	2.2	1.5	0.9	0.6	0.3	0.6
2	41.5	3.8	9.3	18.4	11.4	6.5	3.9	2.4	1.3	0.7	0.4	0.5
3	40.2	3.9	3.3	11.0	15.8	11.6	6.4	3.7	2	1.1	0.6	0.5
4	37.7	3.9	1.8	4.3	12	16.3	11.6	6.4	3.2	1.5	0.7	0.6
5	34.5	4.0	1.3	2.4	5.7	12.7	18.5	11.4	5.4	2.3	1.2	0.7
6	31.6	4.1	0.8	1.3	2.5	5.2	14.8	20.8	11.7	4.7	1.6	0.9
7	27.6	3.9	0.5	0.7	1.2	1.9	4.6	16.6	25.2	12.6	3.7	1.5
8	24.8	3.8	0.4	0.4	0.7	0.9	1.5	3.8	18.8	31.5	10.9	2.5
9	28.2	4.0	0.3	0.3	0.4	0.5	0.8	1.4	3.7	18.2	32.7	9.7
10	31.0	4.5	0.4	0.3	0.3	0.5	0.5	0.6	1.0	2.6	13.7	44.7

Source: Computations based on Portugal, MTSS (1989-1992).
Notes: (*) out db - out of the database; miss w – missing wage (workers who are not wage-earners, but instead owners of the company, unpaid family members or members of cooperatives).

Table 7

Three-Year Transition Matrix, Females, Portugal, 1989-92

	State in 92											
State in 89	out db*	miss w*	1	2	3	4	5	6	7	8	9	10
out db*		12.7	12.5	10.1	9.2	8.7	8.4	8.4	7.9	7.5	6.5	8.0
miss w*	49.8	24.4	2.5	2.4	2.8	2.8	2.8	2.7	3.0	2.4	2.0	2.5
1	48.0	6.0	16.6	9.0	6.9	4.2	3.4	2.5	1.6	0.9	0.4	0.5
2	42.5	5.8	7.7	11.8	10.9	8.4	5.8	3.4	1.9	1.1	0.5	0.5
3	35.8	6.5	3.0	15.5	11.9	14.4	6.3	3.3	1.6	0.8	0.3	0.5
4	35.7	6.3	2.7	6.7	12.6	14.3	11.5	5.6	2.5	1.1	0.6	0.5
5	35.2	6.9	1.8	3.2	8.6	10.2	15.3	10.7	4.9	2.1	0.7	0.5
6	35.5	5.8	1.2	2.0	2.8	4.7	12.2	17.3	12.0	4.5	1.4	0.7
7	34.0	5.5	0.8	0.7	1.0	1.3	4.0	13.5	22.5	12.4	3.3	1.1
8	32.7	5.4	0.5	0.3	0.3	0.5	0.9	3.2	14.0	29.2	10.9	2.1
9	29.1	4.6	0.3	0.2	0.3	0.5	0.6	0.7	2.0	13.3	40.7	7.9
10	28.6	3.7	0.3	0.2	0.2	0.2	0.3	0.3	0.5	1.4	15.2	49.3

Source: Computations based on Portugal, MTSS (1989-1992).
Notes: (*) out db - out of the database; miss w – missing wage (workers who are not wage-earners, but instead ownersof the company, unpaid family members or members of cooperatives).

Table 8

Five-Year Transition Matrix, Males, Portugal, 1987-92

	State in 92											
State in 87	out db*	miss w*	1	2	3	4	5	6	7	8	9	10
out db*		14.4	11.8	10.4	9.7	9.0	8.0	7.3	6.7	7.0	7.7	8.1
miss w*	57.2	24.7	1.9	1.5	1.6	1.8	1.7	1.7	1.7	1.9	1.9	2.4
1	53.5	4.4	14.9	9.5	5.5	4.0	2.9	2.1	1.3	0.9	0.5	0.6
2	48.7	3.9	7.1	12.6	9.0	6.8	4.8	3.2	1.8	1.0	0.6	0.6
3	47.0	3.7	3.0	8.8	11.2	9.4	6.9	4.6	2.7	1.5	0.7	0.7
4	45.1	3.9	1.6	4.2	9.3	11.1	9.8	7.0	4.1	2.1	1.0	0.7
5	43.3	3.8	1.1	2.3	5.8	10.0	11.9	10.3	6.3	3.1	1.4	0.9
6	39.5	4.4	1.0	1.6	2.5	5.6	13.1	13.5	9.9	5.2	2.4	1.2
7	36.9	4.4	0.6	0.8	1.2	2.1	5.7	15.0	16.6	10.6	4.2	2.0
8	33.5	4.4	0.4	0.4	0.6	0.9	1.5	5.4	19.5	20.6	9.6	3.2
9	35.6	4.3	0.3	0.3	0.4	0.5	0.6	1.2	4.2	18.3	25.7	8.7
10	39.0	4.8	0.3	0.3	0.2	0.3	0.3	0.4	0.8	1.8	14.0	37.9

Source: Computations based on Portugal, MTSS (1987-1992).
Notes: (*) out db - out of the database; miss w – missing wage (workers who are not wage-earners, but instead owners of the company, unpaid family members or members of cooperatives).

Table 9

Five-Year Transition Matrix, Females, Portugal, 1987-92

	State in 92											
State in 87	out db*	miss w*	1	2	3	4	5	6	7	8	9	10
out db*		12.5	11.4	9.8	9.2	8.7	8.3	8.5	8	7.5	7.7	8.3
miss w*	58.5	16.8	2.1	2.0	2.2	2.4	2.6	2.8	2.7	2.6	2.7	2.6
1	55.9	5.6	10.5	7.3	6.0	4.7	3.5	2.6	1.9	1.2	0.5	0.5
2	50.5	5.2	5.4	8.4	8.7	7.0	6.0	4.0	2.4	1.4	0.6	0.5
3	43.0	6.5	2.5	12.5	9.6	12.2	6.1	3.5	2.1	1.0	0.5	0.5
4	43.8	6.3	2.4	5.6	10.3	11.2	9.8	5.3	2.9	1.4	0.7	0.4
5	40.5	7.2	1.7	3.4	7.6	9.0	13.3	8.7	4.9	2.3	1.0	0.6
6	42.5	5.7	1.1	1.9	2.5	5.0	9.8	13.3	10.4	5.2	1.8	0.8
7	41.7	5.4	0.8	0.8	1.1	1.4	4.4	10.9	16.4	11.6	4.2	1.4
8	40.1	5.4	0.5	0.4	0.5	0.5	1.1	3.5	13.2	22.2	10.2	2.5
9	38.8	5.1	0.4	0.3	0.2	0.3	0.5	0.9	3.0	15.2	27.9	7.5
10	37.5	4.1	0.3	0.1	0.2	0.3	0.4	0.5	0.6	1.5	13.5	41.2

Source: Computations based on Portugal, MTSS (1987-1992).
Notes: (*) out db - out of the database; miss w – missing wage (workers who are not wage-earners, but instead owners of the company, unpaid family members or members of cooperatives).

References

Atkinson, A. B., Bourguignon, F. and Morrisson, C. (1992) *Empirical Studies of Earnings Mobility*. Reading, UK: Harwood Academic Publishers.

Buchinsky, M. and Hunt, J. (1999) Wage mobility in the United States. *Review of Economics and Statistics*, 81(3), 351-368.

Buchinsky, M., D. Fougere and F. Kramarz (1998) La mobilite salariale en France: 1967-1987. *Revue Economique*, 49(3), 879-890.

Burkhauser, R.V., Holtz-Eakin, D. and Rhody, S.E. (1997) Labour earnings mobility and inequality in the United States and Germany during the growth of thew 1980s. *International Economic Review*, 38(4), 775-794.

Cardoso, A.R. (1999) Firm's wage policies and the rise in labour market inequality: the case of Portugal. *Industrial and Labour Relations Review*, 53(1), 87-102.

Cardoso, A.R. (2000) Wage differentials across firms: an application of multilevel modelling. *Journal of Applied Econometrics*, 15(4), 343-354.

Cardoso, A.R. and Portugal, P. (2003) Bargained wages, wage drift and the design of the wage setting system. Mimeo.

Dickens, R. (2000) Caught in a trap? Wage mobility in Great Britain: 1975-1994. *Economica*, 67(268), 477-97.

Gardiner, K. and Hills, J. (1999) Policy implications of new data on income mobility. *Economic Journal*, 109, F91-F111.

Gittleman, M. and Joyce, M. (1996) Earnings mobility and long-run inequality: an analysis using matched CPS data. *Industrial Relations*, 35(2), 180-196.

Gottschalk, P. and Smeeding, T.M. (1997) Cross-national comparisons of earnings and income inequality. *Journal of Economic Literature*, 35, 633-687.

Hartog, J., Pereira, P.T. and Vieira, J.A.C. (2001) Changing returns to education in Portugal during the 1980s and early 90s: OLS and quantile regression estimators. *Applied Economics*, 33, 1021-1037.

Hartog, J., Pereira, P.T. and Vieira, J.A.C. (2000) Inter-industry wage dispersion in Portugal. *Empirica*, 27, 353-364.

Hofer, H. and Weber, A. (2002) Wage mobility in Austria 1986-1996. Labour Economics, 9, 563-577.

Machin, S. (1998) Recent shifts in wage inequality and the wage returns to education in Britain. *National Institute Economic Review*, 166, 87-96.

OECD [Organisation for Economic Co-operation and Development] (1997) Earnings mobility: taking a longer-run view. In his *Employment Outlook*, July 1997. Paris, OECD, 27-61.

OECD [Organisation for Economic Co-operation and Development] (1992) OECD *Economic Surveys*, Portugal, 1991/92. Paris, OECD.

Portugal, Ministerio do Trabalho e da Seguranca Social (1986 to 1999). Quadros de Pessoal. *Data in magnetic media*.

Shorrocks, A. (1978) Income inequality and income mobility. *Journal of Economic Theory*, 19(2), 376-393.

THE ECONOMIC CONSEQUENCES OF WORKER DISPLACEMENT IN PORTUGAL

Anabela Carneiro[*]

July 2004

Abstract

This paper examines the long-term earnings losses of displaced workers in Portugal, using a nationally representative longitudinal linked employer-employee data set. The results show that four years after displacement the earnings of displaced workers remain around 9% (women) to 12% (men) below their counterfactual expected levels. The post-displacement earnings losses are mainly associated with the loss of tenure within the firm and, to a lesser extent, to the loss of sector-specific features. Further-more, workers who experienced a spell of nonemployment are the most affected by job displacement. Finally, this study points to the importance of controlling for employers' characteristics in this type of wages-dynamic analysis, since there are systematic differences in earnings between displaced and non-displaced workers that stem from differences in firm characteristics. Ignoring them may confound the evaluation of the earnings losses.

JEL classification: J31; J63; J65

Keywords: Displacement; Earnings Losses; Nonemployment; Firm Characteristics

[*] Faculdade de Economia da Universidade do Porto and CETE.
Corresponding author. Faculdade de Economia do Porto, Rua Dr. Roberto Frias, 4200-464 Porto, Portugal. Tel.: 351-225571231; fax: 351-225505050; e-mail: anacar@fep.up.pt.

1. Introduction

Worker displacement has been the subject of an extensive and growing literature[1]. The costs of job loss in terms of unemployment (incidence and duration) and earnings-changes (magnitude and persistence) have been the most studied aspects of job displacement.

During the 1980s a number of empirical studies emerged analyzing workers' post-displacement wages and unemployment in the U.S. (see, for instance, Podgursky and Swaim,1987; Kruse, 1988; Addison and Portugal, 1989; Kletzer, 1989, 1991)[2]. Basically, these studies provide a snapshot view of short-term earnings losses, defined as the difference between pre- and post-displacement earnings of displaced workers. This literature established a basic stylized fact - high-tenure workers suffered the greatest losses in terms of subsequent wages and time employed.

However, this type of analysis, focusing solely on workers who have been displaced, is likely to underestimate the magnitude of wage losses, since it does not account for the earnings growth that would have occurred in the absence of job loss. A simple comparison of pre- with post-separation earnings for displaced workers is insufficient. Following the studies of Ruhm (1991) and Jacobson *et al.* (1993) the empirical investigation on worker displacement and wages has been examining how the wages of displaced workers vary compared to workers who were not displaced. The emphasis in worker displacement research has shifted from short-term wage losses to long-term dynamics. In fact, in recent years the existence of suitable longitudinal data sets in the U.S. and Europe that match workers and firms enables the comparison of wage patterns for displaced and identical non-displaced workers.

More specifically, two main issues are addressed by the recent empirical literature on worker displacement and wages. The first issue examines the existence of earnings losses (relative to non-displaced workers) prior to and after a displacement, as well as their magnitude. The second, which is directly related with the first, is to analyze the persistence of these losses over the time.

For a variety of data sets and methodologies the studies for the U.S. have established that displaced American workers usually experience short spells of unemployment, but also substantial and persistent reductions in earnings, on the order of 10-25%[3]. These losses begin to occur at least three years before displacement (see Jacobson *et al.*, 1993). For Europe the evidence is less clear-cut. Some studies have concluded for the existence of significant and large earnings losses, while others have concluded for the existence of reduced wage losses. In one point, however, these studies seem to be in agreement. A displaced worker who takes more than a year to find a new job suffers a large penalty on earnings.

It is to be expected that reemployment wages of displaced workers be lower than the wages of those workers that remain employed. As mentioned by Fallick 1996), there are at least four reasons that can explain this pattern. First, the lost of human capital specific to the firm or industry. To the extent that these skills are

---

1 According to previous research, displacement is usually defined as the permanent and involuntary separation of workers from their jobs without cause (i. e. for economic reasons). Displacements occur when a firm doses or downsizes.

2 See Hamermesh (1989), Fallick (1996) and Kletzer (1998) for surveys.

3 See Kletzer for a survey (1998).

non-transferable, their contribution to worker's productivity is permanently lost when a job loss occurs. Second, payments by seniority in order to provide incentives not to shirk may delay higher earnings to the latter part of the career. In this case, a permanent separation reduces lifetime earnings. Third, there is the loss of a high quality job match between the worker and the firm. In fact, some authors claim that standard estimates of the return to job-specific training are biased upward by job match and individual unobserved heterogeneity[4]. A long job tenure may signal a high quality match between the firm and the worker and/or a high ability worker, because more able workers and workers in good jobs are less likely to separate. In this line of research, Addison and Portugal (1989) and Kletzer (1989, 1991) showed that tenure in the pre-displacement job is positively associated with post-displacement earnings, reflecting heterogeneity in worker ability and the transferability of skills. Fourth, to the extent that firm's and/or industry characteristics also play a role in the process of wage determination, a displaced worker may lose some wage premium that he/she was previously receiving such as insider rents, union premiums or efficiency wage differentials.

Finally, a fifth reason should be added if a displaced worker experiences some period of unemployment. It is widely argued that a spell of unemployment causes the depreciation of general or transferable worker skills and that prolonged unemployment accelerates the process of depreciation of human capital (Phelps, 1972 and Pissarides, 1992). On the other side, some authors also argue that a period of unemployment may signal an individual as a low productivity worker, leading employers to pay, at least initially, less than his/her marginal productivity (Lockwood, 1991).

The main goal in this study is to analyze the costs of worker displacement in terms of earnings losses. Do displaced workers in Portugal suffer pre- and post-displacement wage losses? If so, what are the magnitude and persistence of these losses? What are its main determinants?

Two main objectives will drive the investigation. The first objective is to analyze the long-term impact of a displacement on the earnings evolution of displaced workers in Portugal. For this purpose a rich data set that links employers and employees will be exploited. The recent empirical research on this issue has been based on individual or household data sets with little information on the employers in which the individuals work. The possibility to link workers with their employers constitutes a tremendous advantage of this data set, reinforced by the fact that recent empirical work on wage determination has been showing that employers' characteristics are an important determinant of workers' wages (see, for example, Nickell and Wadhawani, 1990; Blanchflower *et al.*, 1990; Hildreth and Oswald, 1997). At this point, a quantile regression analysis will be used in order to determine if earnings evolution varies significantly according to the worker's position in the wage distribution. The second objective is to decompose the earnings losses according to their sources/causes.

The issue of job displacement and wages is relevant for several reasons. For this study the most important is that it may help to shed some light on the mechanism of wage determination. If wages reflect primarily general skills we would

[4] See, among others, the studies of Abraham and Farber (1987, 1988), Altonji and Shakotko (1987), Topel (1991) and Dustmann and Meghir (1999).

not expect considerable wage losses upon reemployment; if, on the contrary, wages reflect a large investment in firm (industry)-specific human capital or a high quality job match, then a substantial drop in earnings will be expected. Moreover, if nonemployment implies some depreciation of the stock of human capital or acts as a signal for employers of low productivity workers, then larger earnings losses would be expected for workers experiencing longer spells of nonemployment.

The remainder of this study proceeds as follows. In Section 2 a brief literature review on worker displacement and wages is presented. In Section 3 the longitudinal data are described. In Section 4 the empirical model and regressions estimates are reported. Section 5 presents the contribution of worker and firm characteristics for the difference in earnings. Some concluding remarks follow in Section 6.

2. Literature Review

The methodology of Jacobson et al. (1993) introduced a different approach on the study of worker displacement and earnings losses. Basically, their methodology consists in comparing the earnings changes of displaced workers over a long-term period with the earnings changes that would have occurred if the displaced had not lost their jobs. Since this latter variable is not observable, a comparison group of non-displaced workers is used. Using this methodology, the empirical research on this issue has been growing with the appearance of a considerable number of studies for the U.S. and some European countries such as Germany, France, Britain, Italy and Estonia. In this Section we present a brief summary of the main results obtained in these empirical studies.

Using administrative data on Pennsylvanian workers and their firms over the 1974-86 period, Jacobson *et al.* (1993) found considerable losses both prior to and after displacement for displaced workers when compared with continuously employed workers. High-tenure displaced workers suffer long-term (quarterly) earnings losses averaging 25% per year. These losses start to appear approximately three years before separation and are substantial even for workers reemployed in similar firms. These findings are widespread across gender, age, local labour market conditions, industry and firm size.

Ruhm (1991) obtained comparable results using the PSID (Panel Study of Income Dynamics) data set. He found that weekly wages of U.S. displaced workers in the 1970s were lower by 16% in the year following displacement and, after four years, still 14% lower than those of their non-displaced counterparts. Displacement is also associated with substantially elevated initial unemployment. However, and contrary to earnings losses that are large and persistent, unemployment effects seem to dissipate after a few years.

Stevens (1997) using a national sample of displaced workers from the PSID (1968-88), also found large and persistent earnings losses. In her analysis special attention is given to the effects of multiple job losses on future earnings. Stevens shows that workers who avoid additional displacements have a stronger recovery on annual earnings and hourly wages. Hourly wages are reduced by 9% six or more years after displacement, but only 4% when controlling for the most recent displacement. Moreover, hourly wages began to decline one year prior to job displacement by around 8%.

Bender *et al.* (1999) investigated worker displacement in France and Germany in the 1980s using administrative records. Contrary to the evidence for the U.S. they found, for both countries, that displacement does not induce large (daily) earnings losses. Instead, large penalties were found for displaced workers who were out of employment for more than a year. In fact, a displaced worker in France that takes more than one year to find a new job earns 5% less relative to other displaced workers who are reemployed within the calendar year following displacement. In Germany this loss is around 19%.

Burda and Mertens (2001) studied the effects of displacement on reemployment earnings of German workers using the German Socioeconomic Panel and the Social Security file (1985-94). They found that, on average, a displaced worker subsequently reemployed does not suffer a significant wage reduction. Indeed, in the year following the displacement a full-time displaced worker experiences a wage growth reduction of only 3.6% compared with a similar group of non-displaced workers. They also concluded that large wage losses are associated with changes of industry, but not of firm. Kouch (2001) also used the German Socioeconomic Panel (1986-96) in order to examine the effects of displacement due to plant closing on annual earnings and unemployment duration. He reported an estimated loss of around 13.5% of predisplacement earnings in the year of displacement and a loss of only 6.5% two years later. No measurable difference in earnings is reported in the years that precede the displacement.

Arulampalam (2001) used the British Household Panel (1991-97) to examine the effect of an interruption of employment on re-entry wages in Britain. After a spell of unemployment a reemployed individual will face a wage penalty of about 6% on re-entry. After three years this penalty rises to 13.5%. Another important conclusion pointed out by the author is that the first spell of nonemployment causes the largest scar on wages. Gregory and Jukes (2001) also examined the effects of unemployment (incidence and duration) on subsequent earnings of a large sample of British men over the 1984-94 period. They found consistent evidence in favour of an unemployment scarring for British workers, due mainly to the effect of unemployment duration. In fact, unemployment incidence has only a temporary effect on wages that vanishes after two years of continuous reemployment. The effect of unemployment duration, by contrast, is permanent and proportional to the length of the spell. In addition to the average penalty of around 10% resulting from the occurrence of an unemployment spell, a worker that experiences a one-year spell of unemployment will suffer an additional penalty of 10 percentage points on first-year earnings after reemployment.

Rosolia (2002) used the Italian Social Security records (1974-97) to analyze the weekly wage patterns prior to and after job separation of displaced Italian workers. His results reveal that displaced workers experience a wage loss both prior to and following separation, with the magnitude of these losses depending on the duration of the nonemployment spells. Workers experiencing a short nonemployment spell (1-2 quarters) and those staying out of employment for more than a year are the ones who bear the largest losses. Four years after separation their earnings are still around 10% below their expected levels. If workers are immediately reemployed in the quarter after displacement the wage gap is negligible. Some evidence was found that high tenure workers at separation are the ones who lose more in the post-

displacement period, suggesting that earnings losses are mainly determined by the destruction of a specific match between the worker and his employer. On the contrary, general human capital does not seem to affect the estimated pattern of earnings losses.

Lehmann *et al.* (2002) using data from the Estonian labour force surveys, examined the costs of job loss in Estonia during the period of transition from a planned to a market economy (1989-99). The authors conclude that costs of displacement in Estonia are mainly associated with the risk of nonemployment which affects a minority of workers who experience long-term nonemployment, rather than with a potential penalty in the re-entry wage. In fact, little evidence was found in favour of any wage change differential between displaced and non-displaced workers.

More recently, Bowlus and Vilhuber (2002) and Lefranc (2002) employed a job search framework to analyze the effects of worker displacement on wage outcomes. Specifically, Bowlus and Vilhuber (2002) analyzed the displacement effects on reemployment wages using a longitudinal data set for several U.S. states covering the 1990s that allows one to identify, at the same firm, displaced workers as well as early-leavers. Adopting a search model of displacement they showed that, other things equal, workers who left a "distressed" firm up to two quarters prior to the displacement have higher reemployment wages than do those who stayed with the same firm until the displacement.

Finally, Lefranc (2002) analyzed the magnitude and sources of wage losses of displaced workers in France (1990-97) and the U.S. (1981-92) using micro data from labour force surveys. He shows that while the magnitude of the wage losses are very similar in the two countries (around 10-15%), the sources of wage adjustment differ considerably. In France, most of the earnings losses result from the loss of accumulated firm-specific human capital, while in the U.S., earnings losses mostly stem from the loss of search rents on the displacement job.

3. The Data

3.1. Description

The data set of this study comes from *Quadros de Pessoal (QP)* and includes all workers that lost their jobs in 1994, 1995 or 1996 due to firm closure and were present in the *QP* registers in at least one of the three years that precedes the displacement event. Indeed, for reasons of completeness we also included in the sample those workers that left the firm (voluntarily or involuntarily) before closing.

QP is an annual mandatory employment survey collected by the Portuguese Ministry of Employment, and covers virtually all firms employing paid labour in Portugal[5]. The data set includes both firm-specific information [location, industry (SIC codes), legal setting, foreign capital, employment, sales, ownership type] and workforce characteristics (labour earnings, worker qualifications, gender, age, tenure, hours of work, etc).

[5] Hence, this source does not cover operated family businesses without wage-earning employee and self-employment. Public administration is also excluded.

Moreover, the survey has a longitudinal dimension which makes it particularly well suited for analyzing the issues of firms' entry and exit. Each firm entering the database is assigned a unique identifying number and the Ministry implements several checks to ensure that a firm that has already reported to the database is not assigned a different identification number. By merging the yearly data sets by this identifier it is possible to pinpoint all firms that have entered and exited economic activity. In particular, an exit from the database should signal a firm that has ceased its activity. This criteria, however, is not entirely accurate, due to the fact that some of the firms may temporarily exit the database. A temporary exit may occur for a number of reasons other than cessation of activity, a very likely reason being that the survey form was not received in the Ministry of Employment before the date when the recording operations were closed. Almost all of these temporary exits last less than two years, but can still cause an identification problem if they occur in the terminal years. In order to account for this problem, the information on the last two years after displacement was used solely to control for temporary exits in the intermediate years[6]. Thus, a firm is an exiting plant in year *t* if it is present in year *t-l,* but absent in *t, t*+1 and *t*+2. Beyond this control, and to ensure that we are in the presence of firms' true closures and not mergers or acquisitions, we excluded from the sample those workers that appeared in the database in the period after displacement with a year of admission in the new job less than the year of displacement minus one[7].[r] These exclusions reduced the sample by around 20%.

It is worth noting that workers also have an identification number based on a transformation of his/her social security number which allows us to follow them over the years and to match workers and their firms.

In sum, there are three reasons that make this survey an appropriate source for analyzing the effects of job displacement on earnings. The first is its representativeness. Each year every establishment with wage earners is legally obliged to file a standardized questionnaire, which, by law, is posted in a public space of the firm. Indeed, the administrative nature of the data and their public availability imply a high degree of coverage and reliability. The second results from its longitudinal dimension that enables one to identify firm closings and to follow workers over the years[8]. The third advantage is to have information at both individual and firm level and to match workers with their employers. In fact, almost all empirical research on job displacement and wages has been based on individual or household data sets with little information on the employers for which the individuals work.

Of course, this data set also has disadvantages. The most important is that it is impossible to know if a worker is really in a situation of nonemployment when he/she does not appear in the *QP* registers. In fact, workers who exit the *QP* database, besides being unemployed or out of the labour force, may be self-employed or employed in the public administration sector. To construct a proxy for the

---

6 The last year for which the survey is available is 1998.

7 If, for example, a worker's displacement year is 1994 and he (she) appears in the database in the post-displacement period with a year of admission in the new job of 1992 or less, he (she) is excluded from the sample.
It should also be noted that when a worker is present in the QP files for more than one time in a given year, the register in the firm in which he had worked a higher number of hours was selected.

8 Eight spells of QP, from 1991 to 1998, were used for this study.

existence of a nonemployment spell, we have used the year of admission in the last job as a criterion to classify the workers. Thus, a worker is classified as having experienced a positive spell of nonemployment if he (she) is employed in year *t* and is not immediately reemployed in year *t* + 1.

As mentioned above, to be included in the sample a worker must be present in the *QP* registers at least in one of the three years that precedes the displacement. Additionally, and in order to ensure that all workers are employed with the same employer three years before separation, the year of admission in the pre-displacement job should be equal to or less than the year of displacement minus three. Even though this data set includes information on usual hours of work, we decided to include only workers employed full-time in the period before separation. In the period after displacement the individual may be reemployed in a part-time job. We also limited the sample to workers aged between 18 and 59 in the year prior to the firm closing. We have also excluded those observations for which some explanatory variable was not available for a particular year and the extreme values for wages and sales (0.1% top and bottom observations). After these exclusions, we obtained a sample of 106,549 workers that were displaced between 1994-96 due to firm closing.

The control group includes three sub-samples and was constructed in the following way. For each year prior to the displacement year we obtained a random sample of around 300,000 workers who were employed in firms that did not close[9]. These individuals were followed in the three years before displacement and in the years after displacement until 1998. The control group of the 1994 displacements includes a sample of workers employed in 1993 in firms that did not close in 1994. These workers were followed over the 1991-98 period if they remained with the same employer over that period. The control group of the 1995 displacements includes a sample of workers employed in 1994 in firms that did not close in 1995. These workers were followed over the 1992-98 period. Finally, the control group of the 1996 displacements includes a sample of workers employed in 1995 in firms that did not close in 1996. These workers were followed over the 1993-98 period if they remained with the same employer over that period. To be included in the sample a worker must be present in the *QP* files at least in the year that precedes the displacement and in the year of displacement. In order to guarantee that the worker was employed with the same employer three years before separation, we control for worker's admission year in the firm[10]. It should be noted, contrary to previous studies, that we do not restrict our control group only to continuously employed workers over the entire period of analysis, avoiding the potential selection of better than average workers in the control group that a longer time span may induce (e. g. Jovanovic, 1979). The sample of non-displaced individuals includes all workers continuously employed in the time span of interest, but also workers that were employed with the same employer until the displacement year but absent from the *QP* files in the post-displacement years (or at least in part of that period). This sub-sample were also limited to full-time workers in the pre-displacement period aged between 18-59 in the year prior to displacement. After excluding those observations with missing values in the explanatory variables and the extreme values in

9 The sample was drawn according to a normal random number generator

10 In the year prior to displacement the worker must have at least two years of tenure with the employer.

wages and sales, we obtained a control group composed by 172,624 non-displaced workers.

3.2. Some Basic Characteristics

Table 1 presents average real hourly earnings of displaced workers around the date of displacement[11]. The year of displacement is denoted as year zero. The years that precede the displacement are denoted as years minus three to minus one and the following years as years one through four. The pattern of average hourly earnings is shown for all these periods, separately for men and women. Comparable figures for non-displaced workers are also reported. In columns 5 (for men) and 10 (for women) the ratio of average hourly earnings of displaced and non-displaced workers is reported. Figure 1 displays the evolution of the ratio over the entire period of analysis.

As can be shown in Table 1, displaced workers earn, on average, lower wages than non-displaced workers. This gap is more pronounced for male workers. Moreover, and as is made clear in Figure 1, the earnings differential between displaced and non-displaced workers tends to increase with the approximation of the displacement year and deepens until two years after displacement, when a pattern of recovery appears to emerge. This analysis seems to reveal that a displacement event tends to deepen the earnings gap between displaced and non-displaced workers. However, it can always be argued that displaced workers would have had a slow earnings growth even without being displaced. A regression analysis will be required in order to control for observed individual and firm characteristics and to be able to determine the wage growth that displaced workers would have experienced if they had not been displaced.

[11] Average hourly earnings are defined as the ratio between total regular payroll and the total number of usual hours worked.

Table 1

Average Real Hourly Earnings Before and After Displacement

	Men					Women				
	Displaced		Non-displaced		Ratio	Displaced		Non-displaced		Ratio
Years	AHE	Obs.	AHE	Obs.	(1)/(3)	AHE	Obs.	AHE	Obs.	(6)/(8)
Displacement	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
3 years before	421.1	38850	561.0	75954	0.751	322.2	28845	399.8	45459	0.806
2 years before	435.3	37441	583.7	89110	0.746	332.9	28829	419.9	55330	0.793
1 year before	430.8	31276	597.0	105506	0.722	337.7	25411	431.7	67118	0.782
Displacement year	448.3	6442	623.2	105506	0.719	348.3	4433	451.0	67118	0.772
1 year after	445.1	10302	646.3	88994	0.689	346.1	7660	465.1	55325	0.744
2 years after	460.7	12077	675.6	83548	0.682	356.1	8897	485.6	51790	0.733
3 years after	475.0	9422	682.0	49626	0.697	368.2	7168	485.8	29924	0.758
4 years after	500.7	5679	694.3	23371	0.721	379.4	4510	497.8	13677	0.762

Notes:

(i) AHE is average real hourly earnings [PTE (escudo): 1 EUR=200.482 PTE]; CPI deflator (base=1991).

(ii) Obs. denotes number of observations.

Finally, some additional considerations should be made concerning the composition of the sample of displaced workers. First, it should be observed that of the total number of displaced workers, 47% left the firm one or two years prior to the displacement year (see Table 2). Of the total number of workers that left the firm one year before displacement, 30% were reemployed. This rate is only 22% for workers who left their firms two years before displacement. These results seem to suggest that early leavers have lower reemployment rates. It would be interesting to know if these workers left the firm voluntarily or because they were laid off. In this data set, however, it is not possible to know the causes of workers' separations.

Second, of the total number of reemployments 58%, experience a positive spell of nonemployment and only 4% two or more spells (see Table 3). Furthermore, and according to Table 3, 35% of the reemployed individuals changed industry[12]. After being reemployed, 22% of the reemployed individuals changed job again (19% one time and 3% two or more times).

Third and finally, it should be noted that of the total number of workers that were displaced between 1994 and 1996, only 45% were reemployed in the following period[13]. Of this total, 44% were immediately reemployed in the displacement year,

[12] Here industry is defined at the one-digit level. According to the Portuguese Classification of Economic Activities (CAE), at the one-digit level there are nine sectors. We decided to define the industry at the one-digit level in order to avoid measurement errors, since in 1995 the CAE codes were altered.

[13] It is worth noting that even though the reemployment rate is 45%, the total number of reemployments represents solely one third of the sample of displaced workers. This is due to the fact that observations with missing values on the explanatory variables in the post-displacement period were dropped.

26% one year after, 17% two years after, 9% three years after and the remaining 4% only four years after the displacement occurred (see Table 4). This low rate of reemployment may be justified, on the one hand, by the existence of a high replacement rate of the unemployment benefit and, on the other hand, by the long period of duration of the unemployment benefit[14]. Thus, it is possible that the current profile of the unemployment benefit system, which brings together a high replacement rate and a perspective of a long maximum period for the granting of the benefit, may contribute to reducing the transitions from unemployment to employment.

Nonetheless, these figures point to a potential problem of selectivity that may arise when we analyze the earnings losses of only those displaced workers that reenter employment. In fact, for an individual's wage to be observed in the post-displacement period he/she must be employed and registered in the QP files at least one time over that period. This requirement may give rise to non-randomness in the sample of wage observations, and if displaced reemployed workers have (un)observed characteristics that are systematically different for those who stay out of employment, this sample selection may bias the earnings losses estimates. In order to check that the earnings losses estimates are not biased due to sample selection, in the next Section a selectivity test for the sample of displaced workers will be presented.

Table 2

Early-leavers - Timing and Frequency

Displaced Workers [N= 106549]			
Early Leavers [N=49862 (47%)]			
2 Years Before [N=23304 (47%)]		1 Year Before [N=26558 (53%)]	
Reemp.	Non eemp.	Reemp.	Non Reemp.
5235 (22%)	18069 (78%)	7942 (30%)	18616 (70%)

Table 3

Reemployments – Some Basic Characteristics

Displaced Workers [N= 106549]				
Reemployed [N=34825]				
Out of Employment		Changed Job After Reemp.		Changed
1 spell	2 or more	1 time	2 or more	Industry
20311 (58%)	1221 (4%)	6641 (19%)	1016 (3%)	12155 (35%)

[14] In Portugal the unemployment benefit represents approximately 65 per cent of the wages that an individual expects to obtain in the labour market. The duration of this benefit varies from a maximum of 10 months for individuals under 25 years of age to a maximum of 30 months for those aged 55 or over.

Table 4

Reemployments by Years since Displacement

Displaced Workers [N= 106549]				
Reemployed [N=34825]				
Disp. Year	1 Year After	2 Years After	3 Years After	4 Years After
15263 (44%)	8896 (26%)	6057 (17%)	3257 (9%)	1352 (4%)

4. Empirical Model and Results

4.1. The Econometric Model

As mentioned before, the main goal in this study is to compare the earnings paths of displaced and non-displaced workers in the period before and after the displacement has occurred. The empirical model is drawn from Jacobson *et al.* (1993) and Margolis (1999). An equation capturing the difference in earnings across the displaced and comparable non-displaced workers in the sample can be written as:

$$\ln W_{it} = \beta_1 X_{it} + \beta_2 Z_{ijt} + \sum_{k=-3}^{4} D_{it}^{k} \delta_k + \varphi \sum_{t=1}^{4} OE_{it} + \gamma_t + \varepsilon_{it} \qquad (1),$$

where $\ln W_{it}$ is the natural log of average real hourly earnings for individual *i* at period *t*. Average hourly earnings are defined as the ratio between total regular payroll and the total number of usual hours worked[15]. Average hourly earnings were deflated using the Consumer Price Index (CPI; base=1991). X_{it} is a vector of individual characteristics such as education and age. Z_{ijt} includes a set of characteristics from the firm in which the individual works such as size and sales per worker.

D_{it} is a dummy variable that takes the value one if at time *t* worker *i* is *k* years after separation or -*k* years before separation. The δ_k parameters reflect the difference in earnings *k* years before or after separation between displaced workers and the corresponding reference group. Hereinafter, we will refer to these dummies as displacement dummies. The variable OE_{it} takes the value one in the after-separation period for those individuals who experience at least one spell of nonemployment (zero otherwise). The parameter φ captures a negative permanent effect on earnings for those individuals who experienced a nonemployment spell.

γ_t is a set of time dummies that control for year-specific effects. ε_{it} is a disturbance term with zero mean and constant variance.

[15] Total regular payroll includes base wages, seniority payments and regular benefits.

Equation (1) will be estimated by weighted least squares (WLS) separately for the samples of male and female workers[16]. The validity of this estimation method relies on the assumption that the included independent variables control for all the individual heterogeneity that could be correlated with the occurrence of a displacement. In this study the fact that job displacement occurs for exogenous reasons (firm closing) makes this assumption reasonable. If unobserved permanent individual heterogeneity affects the probability of displacement and is correlated with wage determinants, WLS estimates of equation (1) will be biased.

If unobserved individual heterogeneity is of concern, the model could be estimated by fixed effects. Even though the fixed effects model allows us to control for unobserved heterogeneity, the inexistence of time variability does not allow us to identify the coefficients of the time-invariant variables, invalidating the decomposition of the effects of workers and firm characteristics in the differences in earnings between displaced and continuously employed workers.

4.2. Empirical Results

4.2.1. Pre- and Post-displacement Earnings Losses - WLS Regressions

In Tables 5 and 6 the WLS regression results of model (1) are presented for men and women, respectively. Column 1 of Tables 5 and 6 provides results for a parsimonious specification in which average real hourly earnings depend on the displacement dummies ($Disp_{it}$) and a set of individual characteristics that do not change with the job, such as age (and its square) and education. Education is defined as the number of years of schooling completed. Two dummy variables were added to the model: one that takes the value one if the worker has a part-time job in the post-displacement period (zero otherwise), and the other that takes the value one for displaced workers who left the firm one or two years before closing (zero otherwise). A set of time dummies is also included in order to control for macroeconomic conditions.

As can be shown in column 1 of Table 5, men that will be displaced in year zero earn, three years before separation, 11.8% less than their non-displaced counterparts, conditional on age, education, and macroeconomic conditions[17]. In these same conditions women earn 8.3% less (see column 1 of Table 6). This gap increases for both male and female with the approximation of the displacement year. Two years after the displacement year, Portuguese men earn 22.0% less than their continuously employed counterparts, and women earn 17.5% less. Only in the latter years does a slower

---

[16] Since the experiment group corresponds to the population of displaced workers through firm closing between 1994 and 1996 and the corresponding control group to a random sample of non-displaced workers, we decided to use the weighted least squares method of estimation. Thus, each observation was weighted by its representativeness in the population.

[17] To calculate the percentage change in y induced by Δx, i.e. $\frac{yt+1-yt}{yt}$, the exact formula is $\frac{yt+1-yt}{yt} = \exp(\delta) - 1$.

recovery in earnings seems to emerge. In any event, three years after displacement men's earnings differential has risen by around 11 percentage points (p.p.) and women's differential by around 8 p.p., when compared to the earnings differential three years before displacement[18]. This pattern is, in general, consistent with Figure 1. However, the introduction of a control for worker characteristics considerably decreases the earnings differential, suggesting that differences in worker characteristics between the two groups explain a sizable part of the wage gap between displaced and non-displaced workers.

The coefficient estimates of the variable early-leaver are negative but not statistically different from zero. Even though early-leavers represent almost half of the sample of displaced workers, the pattern of earnings losses for these workers does not differ substantially from the pattern of earnings losses of those workers who remain in the firm until the closing down[19].

Differences in earnings between displaced and similar non-displaced workers may be explained by differences in employer characteristics where the individuals work. It is well established that closing firms have some characteristics (such as size and age) that are systematically different from those that remain in activity. It is also well known that differences in employer's performance may affect wages. In order to analyze to what extent the earnings gap between displaced and non-displaced workers is determined by differences in employer characteristics, a set of controls for employers observing heterogeneity was included in column 2 of Tables 5 and 6. The first variable is size, which is measured as the natural log of total employment in the firm. Sales per worker is defined as the ratio of annual real sales and total employment (in logs). Finally, eight industry (one-digit level) and six regional (NUTs II) dummies were added to the model.

The effects of size and sales per worker on average hourly earnings are positive and strongly significant. The results reveal that a large part of the relative annual earnings differential may be explained by differences in employers observed characteristics. Three years before separation the earnings gap between displaced workers and the reference group is almost negligible and not statistically different from zero (+1.1% for men and +1.2% for women). Nonetheless, and after controlling for firms' characteristics, it is still possible to observe a very similar pattern on earnings evolution over the entire period of analysis. A pre-displacement dip in earnings is observed, followed by a drop in earnings in the displacement year and a slower recovery four years after separation (see column 2 of Tables 5 and 6). Not surprisingly, three years after displacement the relative earnings differential has risen by around 12 p.p. for men and 10 p.p. for women (only slightly higher than those obtained in specification 1).

---

[18] In Table B.1 of Appendix B this same parsimonious specification was estimated with individual-specific fixed effects. It should be noted that the fixed effects estimates of the coefficients of the displacement dummies do not have a straightforward interpretation in terms of earnings changes of displaced workers relative to non-displaced workers, since the coefficients represent within-individual earnings changes. Nevertheless, the fixed effects estimates also exhibit significant and negative effects of past displacement on hourly earnings.

[19] The same regressions presented in Tables 5 and 6 were re-estimated for a subsample of workers in which early-leavers were excluded. The results are quite similar to those obtained for the overall sample, and are not recapitulated here (available upon request).

These results point to the importance of controlling for observed employers' characteristics when analyzing the effects of a displacement event on earnings. In fact, firms' characteristics such as size and sales per worker have a significant impact on earnings, and not controlling for them may confound the evaluation of the earnings losses. In the next Section we shall decompose the difference in earnings between displaced and non-displaced workers according to firm characteristics.

Table 5

Weighted Least Squares Regressions - Men Dependent variable: log of Average Real Hourly Earnings (N=773104)

	(1)		(2)		(3)		(4)	
Independent Variables	Coef.	t-ratio	Coef.	t-ratio	Coef.	t-ratio	Coef.	t-ratio
$Disp_{-3}$	-0.126	(-14.0)	0.011*	(1.5)	0.004*	(0.5)	0.010*	(1.3)
$Disp_{-2}$	-0.151	(-17.7)	-0.006*	(-0.9)	-0.012*	(-1.7)	-0.007*	(-1.0)
$Disp_{-1}$	-0.193	(-22.9)	-0.029	(-4.1)	-0.029	(-4.0)	-0.026	(-3.8)
$Disp_0$	-0.203	(-10.8)	-0.073	(-4.6)	-0.039*	(-2.3)	0.010*	(0.6)
$Disp_1$	-0.234	(-15.6)	-0.087	(-7.0)	-0.042	(-3.0)	0.008*	(0.6)
$Disp_2$	-0.248	(-17.8)	-0.109	(-9.4)	-0.052	(-3.7)	0.0002*	(0.0)
$Disp_3$	-0.254	(-16.2)	-0.114	(-8.7)	-0.052	(-3.3)	0.002*	(0.1)
$Disp_4$	-0.252	(-12.6)	-0.113	(-6.7)	-0.046*	(-2.4)	0.009*	(0.4)
Age	0.066	(195.3)	0.046	(161.8)	0.046	(161.8)	0.043	(144.4)
Age Squared[a]	-0.059	(-141.1)	-0.042	(-120.1)	-0.042	(-120.1)	-0.042	(-114.0)
Education	0.103	(679.4)	0.069	(479.1)	0.069	(479.1)	0.071	(488.6)
Part-time Job	0.101	(27.7)	0.091	(29.9)	0.091	(30.0)	0.092	(30.4)
Early-leaver	0.012*	(-1.4)	-0.025	(-3.4)	-0.011*	(-1.5)	-0.009*	(-1.2)
Size			0.063	(254.5)	0.063	(254.5)	0.057	(226.2)
Sales per Worker			0.113	(254.9)	0.113	(254.9)	0.114	(258.2)
OE (Out of Employment)					-0.066	(-5.2)	-0.060	(-4.7)
JC (Job Changes)					-0.012*	(-0.7)	-0.011*	(-0.6)
CIND (Changed Industry)					-0.062	(-4.9)	-0.066	(-5.3)
Tenure							0.006	(35.6)
Tenure Squared[a]							-0.001*	(-2.5)
Constant	3.970	(604.1)	3.399	(521.1)	3.399	(521.2)	3.443	(522.7)
$\overline{R}^2$	0.44		0.61		0.61		0.62	

Notes:
(i) all specifications include a set of time dummies;
(ii) specifications (2), (3) and (4) include a set of industry and regional dummies;
(iii) [a] variables divided by 100; (iv) all estimates are significant at 1%, except those with an *.

Table 6

Weighted Least Squares Regressions – Women
Dependent variable: log of Average Real Hourly Earnings (N=501494)

	(1)		(2)		(3)		(4)	
Independent Variables	Coef.	t-ratio	Coef.	t-ratio	Coef.	t-ratio	Coef.	t-ratio
$Disp_{-3}$	-0.087	(-9.7)	0.012*	(1.7)	0.006*	(0.7)	0.016*	(2.1)
$Disp_{-2}$	-0.103	(-12.2)	0.004*	(0.6)	-0.001*	(-0.2)	0.008*	(1.1)
$Disp_{-1}$	-0.130	(-15.8)	-0.006*	(-0.9)	-0.006*	(-0.9)	-0.003*	(-0.4)
$Disp_0$	-0.151	(-7.5)	-0.044	(-2.7)	-0.001*	(-0.1)	0.068	(4.1)
$Disp_1$	-0.174	(-11.4)	-0.056	(-4.5)	-0.005*	(-0.4)	0.064	(4.7)
$Disp_2$	-0.192	(-13.5)	-0.084	(-7.2)	-0.022*	(-1.6)	0.047	(3.5)
$Disp_3$	-0.191	(-12.1)	-0.089	(-6.9)	-0.018*	(-1.2)	0.051	(3.3)
$Disp_4$	-0.185	(-9.4)	-0.083	(-5.1)	-0.011*	(-0.6)	0.057	(3.0)
Age	0.040	(108.2)	0.026	(85.8)	0.026	(85.8)	0.019	(57.6)
Age Squared[a]	-0.031	(-63.4)	-0.020	(-50.2)	-0.020	(-50.2)	-0.014	(-32.5)
Education	0.102	(587.8)	0.069	(408.4)	0.069	(408.4)	0.070	(417.4)
Part-time Job	0.144	(41.8)	0.108	(38.0)	0.108	(38.2)	0.110	(39.1)
Early-leaver	-0.011*	(-1.2)	-0.026	(-3.5)	-0.012*	(-1.5)	-0.009*	(-1.1)
Size			0.059	(231.3)	0.059	(231.4)	0.054	(209.3)
Sales per Worker			0.098	(200.4)	0.097	(200.4)	0.095	(196.1)
OE (Out of Employment)					-0.049	(-3.8)	-0.034	(-2.6)
JC (Job Changes)					-0.018*	(-1.0)	-0.011*	(-0.6)
CIND (Changed Industry)					-0.131	(-9.7)	-0.132	(-9.9)
Tenure							0.013	(63.6)
Tenure Squared[a]							-0.025	(-38.0)
Constant	4.239	(622.6)	3.713	(555.3)	3.713	(555.4)	3.850	(566.1)
$\overline{R}^2$	0.46		0.63		0.63		0.64	

Notes: see notes to Table 5.

In column 3 of Tables 5 and 6 controls for time spent out of employment and for post-displacement mobility were added to specification 2. The variable OE takes on the value one in all years following the displacement if the individual experienced one or more spells of nonemployment. A worker experiences a positive spell of nonemployment if he (she) is employed in year *t* and is not immediately reemployed in year $t+1$[20].

Since our sample includes displaced workers that may change job several times after being reemployed, a variable that controls for the effects of changing a job on subsequent wages was also included in equation (1). An individual experiences a job change if the identification number of his recent employer is different from the identification number of his previous employer. The variable JC takes the value one in all years after a job change has occurred, if the individual changed job again after being reemployed. Two categories were defined: 1 or more job changes and no job change (the omitted category). Finally, a dummy variable for industry changers was included in column 3. The variable CIND takes the value one in the post-displacement period for reemployed displaced workers who changed industry (at the one-digit level).

The results reported in column 3 of Tables 5 and 6 show that workers who experienced one or more spells of nonemployment are those most affected by displacement. In fact, once a control for time spent out of employment is included, the displacement dummies reflect the difference in average hourly earnings between the control group and those displaced workers who do not experience any spell of nonemployment. For the latter, the annual earnings gap three years after displacement increased by 6 and 3 p.p. for men and women, respectively. Male workers who experienced at least one spell of nonemployment have an additional penalty on hourly earnings of 6.6 p.p. and female workers of 4.9 p.p.

This result seems to suggest that earnings losses for displaced workers in Portugal reflect a major difference between those who find a new job quickly and those who do not. This empirical evidence is in line with that obtained by Bender *et al.* (1999) and Margolis (1999) for displaced workers in Germany and France.

Furthermore, the results reveal that workers reemployed in the same one-digit industry suffered lower post-displacement earnings losses. In fact, displaced male workers who changed industry earn 6.2% less and displaced women 13.1% less. These results are in line with those obtained by Addison and Portugal (1989), Jacobson *et al.* (1993), Carrington (1993) and Stevens (1997), who showed that industry changers in the U.S experience larger earnings losses. Finally, the coefficient estimates of the variable JC are negative but not statistically different from zero.

Post-displacement earnings losses may stem from the loss of accumulated firm-specific human capital and/or a "good" job match. In column 4 of Tables 5 and 6 a control for worker's tenure was added to specification 3. Tenure is measured as the total number of years with the employer. The inclusion of a control for the returns to accumulated tenure will enable us to determine the extent to which post-displacement earnings losses result from the loss of firm-specific capital and/or a high quality match.

[20] Note that we are also controlling in the regressions for subsequent separations after the first displacement through firm closing.

The results reveal that tenure has an important impact on the magnitude of the relative annual earnings differential. As expected, the post-separation earnings differentials are much smaller (by around 5 to 7 p.p.) when a control for job tenure is included in the model. According to column 4, once losses of firm tenure are accounted for, no negative annual earnings differential is observed over the post-displacement period for either male or female workers. These results suggest that the loss of match-specific characteristics has a significant impact on earnings, deepening the average hourly earnings gap between displaced and their similar non-displaced counterparts.

To sum up, we can decompose the earnings losses of displaced workers relative to non-displaced workers into three components: (i) the loss associated with nonemployment duration; (ii) the loss that stems from the loss of tenure in the job; and (iii) the loss related with industry changing. Overall, the results for the male sample reveal that the increase in the earnings gap of 12 p.p. is mainly due to the loss of tenure in the job and to joblessness. Indeed, joblessness and tenure account equally (45.5% each) for the deepening in the earnings gap. Changing industry explains only 9% of the increase in the earnings gap. For female workers, the increase in the earnings gap of 10 p.p. is mainly due to the loss of accumulated returns to tenure (50%), joblessness accounts for 33% of that increase and changing industry accounts only for 17%.

The four regressions presented in Tables 5 and 6 were re-estimated for a sub-sample of workers in which minimum wage earners were excluded and for a sub-sample of workers in which displaced workers that changed job again after being re-employed were dropped (separately for men and women). These results are quite similar to the ones obtained in Tables 5 and 6 and are not recapitulated here (available upon request). Finally, in Appendix C we present the results for the selectivity test that was carried out for the sample of displaced workers[21].

4.2.2. Pre- and Post-displacement Earnings Losses - Quantile Regressions

In this Section we use quantile regression techniques in order to examine if displacement effects on wages vary along the wage distribution. Unlike the conditional mean regression, the quantile regression analysis allows us to determine the effect of each of the covariates along the whole distribution of hourly wages and to examine to what extent the marginal effect of displacement on wages depends on the point of the conditional wage distribution where the individual is located.

Table 7 presents the quantile regressions estimates for male workers according to specifications 2 and 4 of Table 5. Table 8 presents the same estimates for female workers. The selected quantiles are 10-th, 50-th and 90-th. The estimated coefficients measure the impact of each covariate on wages for each specific quantile.

According to column 1 of Tables 7 and 8, even though the effect of each displacement dummy differs across quantiles, the earnings gap between displaced and similar non-displaced workers increased over the period by almost the same amount at the top, median and bottom of the wage distribution. Indeed, three years after

[21] The results reveal that even though the selectivity term is significantly different from zero, it does not affect appreciably the parameter estimates or t-statistics of the independent variables and, in particular, the displacement dummy variables.

displacement, and compared to the earnings gap three years before displacement, the annual earnings differential between displaced male workers and their counterparts has increased by 11 p.p. at the first and ninth decile and by 10 p.p. at the median. For female workers, the same earnings gap increased by 8 p.p. at the first decile, by 7 p.p. at the median and by 9 p.p. at the top decile. These results seem to suggest that the magnitude of the full effect of displacement on earnings is fairly similar across the whole wage distribution. Yet, once controls for job tenure and for time spent out of employment are included in the model, some differences across quantiles seem to emerge, namely for male workers.

According to column 2 of Table 7, once losses of accumulated returns to tenure and loss of sector-specific factors are accounted for, no negative post-displacement earnings differential appears to exist between displaced and similar non-displaced male workers at the 10-th and 50-th quantile. A dip in earnings is observed only in the pre-displacement period. At the 90-th decile, however, male workers seem to suffer post-displacement earnings losses even after controlling for job tenure and industry changing[22]. This might be justified by the fact that the returns to tenure decrease for higher quantiles, suggesting that tenure is more valued at relatively low paid jobs. On the contrary, the returns to education are higher for the top quantiles[23].

In sum, the results for the sample of male workers seem to suggest that even though the magnitude of the global effect of displacement on wages is quite similar across quantiles, the sources of these losses seem to differ along the wage distribution. While for low paid male workers, tenure within the firm and industry changing explain most of the earnings losses experienced by these workers, for high paid male workers there are other (unobserved) factors beyond tenure/industry changing that account for a non-negligible amount of these losses. Changing industry imposes a larger penalty on male earnings at the bottom than at the top quantiles. For low and high paid male workers the effects of being out of employment are quite similar, even though weaker for the former.

For the sample of female workers, as before, once the returns to accumulated tenure are accounted for, a positive earnings differential between displaced (immediately reemployed) and non-displaced workers is observed in the post-displacement period. In this case, and contrary to the sample of male workers, the displacement effects are quite similar across all the selected quantiles (see Table 8).

Overall, the regression quantile results reinforce the OLS results obtained in the previous Section, showing that the full effect of displacement on earnings is quite similar across quantiles. Low and high paid workers seem to be equally affected, in terms of magnitude and persistence, by job displacement. The only exception concerns the different sources of wage adjustment. For female and low paid male workers who find a new job quickly, the main source of earnings adjustment stems from the loss of accumulated returns to tenure. For high paid male workers, and since tenure on the job is less valued than for low paid workers, there are other

[22] According to column 2 of Table 7, only the coefficients of the displacement dummies one year before displacement and two years after displacement are statistically different from zero at the 5% level of significance.

[23] See Machado and Mata (1998) for a detailed description of the conditional wage distribution for Portugal and its evolution from 1982 to 1994.

unobserved sources of earnings losses. One possibility is that these latter losses may stem from the loss of search rents earned on the job lost. This type of explanation was advanced by Lefranc (2002) using a job search approach. The basic idea of the model is that wages will fall after displacement because before displacement workers have, on average, access to better jobs than when exiting unemployment due to the process of on-the-job-search that leads to the accumulation of "search rents". According to his model, after controlling for the accumulation of job-specific human capital, wage losses can be used to assess the level of reservation wages. Indeed, as workers increase their reservation wages, the range of acceptable wage offers shrinks and, consequently, the range for profitable job-to-job mobility will decline. Thus, loss of search rents is negatively related to the level of the reservation wage. Since in Portugal the unemployment benefit is considerably high, low paid workers may have high reservation wages relative to high paid workers, which leads to high unemployment duration for the former, limiting the range of profitable job-to-job mobility and consequently the losses of search rents after displacement[24]. On the contrary, for high paid workers, and due to the existence of relatively low reservation wages, the loss of search rents may be considerably higher.

[24] The amount of benefit paid represents around 65% of the average monthly work pay during a 6-month period prior to the second month before becoming unemployed. This amount can neither be less than the legal minimum wage nor exceed three times its value.

Table 7

Quantile Regressions – Men
Dependent variable: log of Average Real Hourly Earnings (N=773104)

	(1)						(2)					
	Q=0.1		Q=0.5		Q=0.9		Q=0.1		Q=0.5		Q=0.9	
Independent Variables	Coef.	t-ratio	Coef.	t-ratio	Coef.	t-ratio	Coef.	t-ratio	Coef.	t-ratio	Coef.	t-ratio
$Disp._{-3}$	-0.012*	(-1.4)	0.002*	(0.9)	0.014*	(1.5)	-0.010*	(-1.2)	0.003*	(1.6)	0.011*	(1.2)
$Disp._{-2}$	-0.025	(-3.2)	-0.015	(-9.0)	-0.004*	(-0.4)	-0.022	(-2.8)	-0.013	(-7.9)	-0.004*	(-0.5)
$Disp._{-1}$	-0.042	(-5.5)	-0.035	(-22.3)	-0.017*	(-2.0)	-0.038	(-5.0)	-0.032	(-20.5)	-0.017*	(-2.0)
$Disp_0$	-0.089	(-5.5)	-0.076	(-22.8)	-0.062	(-3.4)	0.033*	(1.9)	0.016	(4.4)	-0.014*	(-0.7)
$Disp_1$	-0.103	(-7.9)	-0.082	(-30.9)	-0.07	(-4.8)	0.036*	(2.5)	0.018	(6.0)	-0.020*	(-1.2)
$Disp_2$	-0.115	(-9.5)	-0.1	(-40.1)	-0.106	(-7.8)	0.04	(2.8)	0.011	(3.8)	-0.034*	(-2.0)
$Disp_3$	-0.124	(-9.1)	-0.102	(-36.3)	-0.103	(-6.7)	0.035*	(2.2)	0.014	(4.2)	-0.029*	(-1.5)
$Disp_4$	-0.12	(-6.9)	-0.104	(-28.9)	-0.106	(-5.4)	0.040*	(2.0)	0.016	(4.0)	-0.030*	(-1.3)
Age	0.038	(37.7)	0.044	(213.1)	0.053	(46.9)	0.032	(31.0)	0.041	(189.6)	0.054	(45.5)
Age Squared[a]	-0.037	(-29.8)	-0.042	(-161.3)	-0.047	(-33.4)	-0.034	(-26.1)	-0.041	(-152.4)	-0.051	(34.3)
Education	0.042	(79.6)	0.063	(581.6)	0.091	(153.4)	0.043	(82.4)	0.064	(594.1)	0.093	(155.6)
Part-time Job	-0.0003*	(-0.0)	0.044	(19.5)	0.251	(20.2)	0.005*	(0.4)	0.045	(20.1)	0.26	(21.0)
Early-leaver	-0.05	(-6.8)	-0.035	(-23.2)	0.008*	(1.0)	-0.025	(-3.4)	-0.021	(-13.1)	0.021*	(2.5)
Size	0.059	(67.2)	0.063	(347.6)	0.061	(60.7)	0.053	(59.4)	0.057	(306.9)	0.056	(55.0)
Sales per Worker	0.107	(69.9)	0.104	(329.8)	0.092	(53.7)	0.107	(71.1)	0.105	(335.9)	0.093	(54.1)
OE (Out of Employment)							-0.051	(-4.0)	-0.05	(-19.1)	-0.065	(-4.4)
JC (Job Changes)							-0.037*	(-2.2)	-0.015	(-4.3)	0.009*	(0.4)
CIND (Changed Industry)							-0.09	(-7.2)	-0.059	(-22.8)	-0.044	(-3.1)
Tenure							0.012	-19.5	0.008	(60.5)	0.001*	(1.6)
Tenure Squared[a]							-0.016	(-8.7)	-0.005	(-13.4)	0.011	(5.0)
Constant	3.462	(152.0)	3.545	(754.9)	3.708	(144.5)	3.537	(155.9)	3.59	(762.2)	3.705	(143.4)
$\overline{R}^2$	0.58		0.59		0.58		0.58		0.6		0.59	

Notes:
(i) all specifications include a set of industry, regional and time dummies;
(ii) [a] variables divided by 100;
(iii) all estimates are significant at 1%, except those with an *.

Table 8

Quantile Regressions – Women
Dependent variable: log of Average Real Hourly Earnings (N=501494)

	(1)						(2)					
	Q=0.1		Q=0.5		Q=0.9		Q=0.1		Q=0.5		Q=0.9	
Independent Variables	Coef.	t-ratio	Coef.	t-ratio	Coef.	t-ratio	Coef.	t-ratio	Coef.	t-ratio	Coef.	t-ratio
$Disp_{-3}$	-0.003*	(-0.3)	0.005	(2.8)	0.012*	(1.3)	0.001*	(0.2)	0.007	(4.2)	0.016*	(1.7)
$Disp_{-2}$	0.003*	(0.5)	-0.004	(-2.9)	0.002*	(0.3)	0.009*	(1.2)	-0.002*	(-1.1)	0.007*	(0.8)
$Disp_{-1}$	-0.014*	(-2.0)	-0.014	(-9.5)	0.002*	(0.3)	-0.010*	(-1.5)	-0.013	(-9.0)	0.009*	(1.1)
$Disp_{0}$	-0.057	(-3.6)	-0.038	(-11.6)	-0.023*	(-1.2)	0.043	(2.7)	0.066	(19.4)	0.049	(2.6)
$Disp_{1}$	-0.066	(-5.5)	-0.045	(-17.7)	-0.034*	(-2.4)	0.044	(3.3)	0.066	(23.7)	0.051	(3.3)
$Disp_{2}$	-0.083	(-7.3)	-0.066	(-27.8)	-0.072	(-5.4)	0.032*	(2.4)	0.052	(18.7)	0.023*	(1.5)
$Disp_{3}$	-0.091	(-7.2)	-0.077	(-29.0)	-0.076	(-5.2)	0.029*	(1.9)	0.054	(16.9)	0.023*	(1.3)
$Disp_{4}$	-0.082	(-5.2)	-0.062	(-18.6)	-0.073	(-3.9)	0.038*	(2.1)	0.063	(16.4)	0.030*	(1.4)
Age	0.022	(22.0)	0.021	(98.9)	0.024	-20.8	0.016	(15.5)	0.015	(69.0)	0.021	(17.4)
Age Squared(a)	-0.023	(-17.3)	-0.017	(-58.7)	-0.013	(-8.2)	-0.017	(-12.2)	-0.012	(-41.0)	-0.013	(-8.1)
Education	0.032	(56.8)	0.06	(502.0)	0.092	(139.0)	0.033	(58.1)	0.061	(512.6)	0.094	(143.1)
Part-time Job	-0.011*	(-1.2)	0.039	(19.8)	0.256	(23.1)	-0.003*	(-0.4)	0.051	(25.6)	0.264	(24.1)
Early-leaver	-0.029	(-4.2)	-0.025	(-17.3)	-0.009*	(-1.1)	-0.017*	(-2.4)	-0.009	(-6.2)	0.005*	(0.6)
Size	0.05	(59.7)	0.054	(305.3)	0.055	(56.0)	0.049	(57.8)	0.05	(278.8)	0.05	(50.9)
Sales per Worker	0.061	(39.4)	0.082	(252.6)	0.09	(50.1)	0.06	(39.6)	0.079	(245.7)	0.087	(48.8)
OE (Out of Employment)							-0.029*	(-2.4)	-0.028	(-11.1)	-0.032*	(-2.3)
JC (Job Changes)							-0.024*	(-1.4)	-0.015	(-4.3)	0.016*	(0.8)
CIND (Changed Industry)							-0.124	(-9.8)	-0.143	(-54.0)	-0.109	(-7.4)
Tenure							0.012	(17.7)	0.013	(86.0)	0.009	(10.9)
Tenure Squared(a)							-0.031	(-14.0)	-0.023	(-50.7)	-0.006*	(-2.5)
Constant	4.131	(190.6)	4.012	(878.9)	3.977	(157.3)	4.192	(191.2)	4.11	(895.7)	4.046	(159.0)
$\overline{R}^2$	0.58		0.60		0.60		0.59		0.61		0.60	

Notes: see notes to Table 7.

5. The Contribution of Worker and Firm Characteristics to Earnings Difference

In this Section we examine how worker and firm characteristics contribute to the difference in earnings between displaced and non-displaced workers before and after displacement. The average earnings differential between displaced and non-displaced workers before and after displacement can be written as:

$$\begin{aligned}\overline{\ln W}^{D} - \overline{\ln W}^{ND} &= \left(\overline{\ln W}^{D}_{AD} - \overline{\ln W}^{ND}_{AD}\right) - \left(\overline{\ln W}^{D}_{BD} - \overline{\ln W}^{ND}_{BD}\right) = \\ &= \beta_1\left[\left(\overline{X}^{D}_{AD} - \overline{X}^{ND}_{AD}\right) - \left(\overline{X}^{D}_{BD} - \overline{X}^{ND}_{BD}\right)\right] + \\ &+ \beta_2\left[\left(\overline{Z}^{D}_{AD} - \overline{Z}^{ND}_{AD}\right) - \left(\overline{Z}^{D}_{BD} - \overline{Z}^{ND}_{BD}\right)\right] + \left[\delta_{AD} - \delta_{BD}\right]\end{aligned} \quad (2),$$

where $\overline{\ln W}^{D}_{AD}$ is the mean of the log of average real hourly earnings of displaced workers in the after-displacement period and $\overline{\ln W}^{ND}_{AD}$ the mean of the log of average real hourly earnings of non-displaced workers in the post-displacement period. $\overline{\ln W}^{D}_{BD}$ and $\overline{\ln W}^{ND}_{BD}$ represent the same means but in the pre-displacement period. $\overline{X}$ are the means of variables that measure worker characteristics such as education and age, and $\overline{Z}$ the means of variables that control for firm characteristics such as size and sales per worker. These means are defined for displaced *(D)* and non-displaced workers *(ND)* in the period after *(AD)* and before displacement *(BD)*. β_1 and β_2 are the coefficient estimates of the variables included in $\overline{X}$ and $\overline{Z}$, respectively δ_{AD} and δ_{BD} are the coefficient estimates of the displacement dummies in the post- and pre-displacement period respectively.

In order to obtain the estimates of the $\beta's$ and the $\delta's$ parameters the time periods around the date of displacement were divided into two groups: the pre-displacement period, which includes the three years that precede the displacement year; and the post-displacement period, which includes the displacement year and the following four years. The estimates of an earnings regression equation similar to the one presented in column 4 of Table 5 are presented in Table B.2 of Appendix B for both men and women. Since we are interested in examining the contribution of worker and firm characteristics for earnings difference before and after displacement, we chose a specification that includes all the controls for observed workers' and firms' heterogeneity.

In Tables 9 and 10 the contribution of each independent variable to the difference in earnings before and after displacement is presented for male and female workers, respectively[25]. The results report that the post-displacement earnings differential between displaced and non-displaced workers increased by 2.8 log points

[25] The variables early-leaver and JC were not included in Tables 9 and 10 because their coefficients are not statistically different from zero.

for male workers and by 3.9 log points for female workers (see column 7). This increase in the earnings differential between displaced and non-displaced workers is related to the displacement event in two distinct ways. On one side, job displacement might affect the composition of the sample of displaced workers, since only 45% of the total number of displaced workers are reemployed in the post-displacement period. On the other side, and supposing that the sample of displaced workers remains unchanged concerning time-invariant characteristics, the displacement has an effect on variables that are associated with the job such as tenure, industry and firm size. The decomposition results show that both forces are at work in explaining the difference-in-differences in earnings.

The results reported in Tables 9 and 10 show that the earnings differential between displaced and non-displaced workers due to differences in education became positive. Indeed, whereas in the pre-displacement period displaced workers are, on average, less educated than non-displaced workers, in the post-displacement period displaced workers have, on average, more education (6.7 years against 6.4 for men and 6.9 against 6.6 for women). In fact, the data reveal that around 23% of the displaced workers that were reemployed changed their educational level[26]. At the first glance, this empirical evidence seems to suggest that displacement might have a positive impact on displaced workers, in the sense that it encourages the investment in human capital[27]. However, we have to put some caution in the interpretation of this result, since it may merely reflect the existence of measurement errors in the variable education that stem from the alteration of the education codes in 1994 and the alteration of the mandatory schooling from 4 to 9 years.

For both samples the earnings differential between displaced workers and the reference group due to differences in age increased. The negative contribution of age may result from the fact that older workers who were displaced left the labour force, taking advantage of early retirement schemes.

Besides the difference in earnings that stem from changes in the composition of the displaced workforce, there are other factors directly related to the loss of the job that cause post-displacement earnings losses. The results show that for both men and women the loss of tenure on the job explains why the earnings differential between displaced and non-displaced workers deepened in the post-displacement period. Changing industry after a displacement also con-tributes considerably to the increase in the earnings gap between displaced and non-displaced workers. Moreover, workers who experienced a spell of nonemployment have an additional and significant penalty on earnings. Once again, the decomposition of the contribution of each variable to the difference in earnings between displaced and non-displaced workers shows that the loss of tenure on the job, joblessness and industry changing are the main factors that explain the deepening in the earnings gap in the post-displacement period (see column 7).

---

26 Of the total number of workers that obtained more education after the displacement, 45% changed from 4 to 6 years of completed schooling and 14% from 6 to 9 years. It should be noted that the mandatory schooling is 9 years.

27 In Portugal there are some training programs promoted by the Institut for Employment and Vocational Training (IEFP) that allow individuals with low educational levels (especially, young individuals) to obtain a complete degree of schooling. Beyond that, workers who left school a long time ago may come back again to learn with a special educational program (the so-called "ensino recorrente").

Finally, the results reported in Tables 9 and 10 also reveal that there are differences in earnings even before displacement between displaced and non-displaced workers due to differences in firm characteristics in which the individuals work, and that tend to persist in the post-displacement period. In fact, displaced workers are employed in the pre-displacement period in firms with a smaller average size and with a lower average productivity (measured by sales per worker). In the post-displacement period, even though the earnings differential that stem from differences in average size is reduced, it remains considerably high (see column 6). Concerning sales per worker, the earnings differential in the post-displacement period is reduced for male workers and increased for female workers (see column 6).

This evidence reinforces the importance of controlling for employers' attributes when analyzing the displacement effects on earnings, since the pools of displaced and non-displaced workers seem to be very heterogeneous concerning firm characteristics. In fact, there are systematic differences in earnings between displaced and non-displaced workers due to systematic differences in firm characteristics that exist even before displacement, and tend to persist after the displacement.

The empirical literature on plant closings has been showing that firms that will soon close are usually of small size and lower productivity, so that the differences in firm characteristics for the two groups of workers in the pre-displacement period are not surprising. In the post-displacement period the differences in the two groups are not so justifiable, unless it can be argued that workers displaced through firm closing carry with them a "displacement stigma" that lead firms (namely, high productivity firms) facing a queue of job applicants to prefer those individuals who did not experience a displacement and probably no spell of unemployment (Blanchard and Diamond, 1994).

In sum, the decomposition of the contribution of each variable to the difference in earnings corroborates the regressions results that the post-displacement increase in the relative earnings differential for displaced workers who are immediately reemployed mostly stems from the loss of tenure within the firm and, to a lesser extent, from the loss of sector-specific features. For workers who take a long time to find a new job the earnings differential is considerably higher. For these workers the additional losses may result from the depreciation of the stock of human capital (general or specific) that a spell of nonemployment imposes or from lower wage offers received as employers perceive that these workers are of lower productivity.

Table 9

The Contribution of Worker and Firm Characteristics to the Difference-in-Differences in Earnings

Men (N=773104)

	Displaced Before Disp.	Non-disp. Before Disp.	Displaced After Disp.	Non-disp. After Disp.	Differential Before Disp.	Differential After Disp.	Difference-in -Differences
					[(1)-(2)]*ß	[(3)-(4)]*ß	(6)-(5)
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age	35.526	37.968	36.518	40.858	-0.036	-0.043	-0.007
Education	5.738	6.18	6.718	6.433	-0.031	0.020	0.051
Part-time Job	0.000	0.000	0.062	0.03	0.000	0.003	0.003
Size	3.321	4.892	3.521	4.849	-0.09	-0.076	0.014
Sales per Worker	8.446	8.923	8.574	8.982	-0.054	-0.047	0.007
OE	0.000	0.000	0.520	0.000	0.000	-0.034	-0.034
CIND	0.000	0.000	0.359	0.000	0.000	-0.024	-0.024
Tenure	7.308	10.736	1.263	13.751	-0.02	-0.072	-0.052
$\delta_{AD} - \delta_{BD}$							0.014
							-0.028

Note: the values of columns 5 and 6 for the variables age and tenure correspond to the net contribution of these variables.

Table 10

The Contribution of Worker and Firm Characteristics to the Difference-in-Differences in Earning

Women (N=501494)

	Displaced Before Disp.	Non-disp. Before Disp.	Displaced Before Disp.	Non-disp. Before Disp.	Differential Before Disp.	Differential After Disp.	Difference-in -Differences
					[(1)-(2)]*ß	[(3)-(4)]*ß	(6)-(5)
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age	32.760	34.212	33.403	37.249	-0.015	-0.033	-0.018
Education	5.955	6.373	6.943	6.643	-0.029	0.021	0.050
Part-time Job	0.000	0.000	0.080	0.043	0.000	0.004	0.004
Size	3.417	4.583	3.612	4.591	-0.063	-0.053	0.010
Sales per Worker	8.103	8.527	8.086	8.579	-0.04	-0.047	-0.007
OE	0.000	0.000	0.517	0.000	0.000	-0.020	-0.020
CIND	0.000	0.000	0.287	0.000	0.000	-0.038	-0.038
Tenure	6.930	9.212	1.290	12.306	-0.018	-0.09	-0.072
$\delta_{AD} - \delta_{BD}$							0.052
							-0.039

Note: the values of columns 5 and 6 for the variables age and tenure correspond to the net contribution of these variables.

6. Conclusions

6.1. Main Results

In this study we have analyzed the long-term earnings costs of displacement for Portuguese displaced workers through firm closing. For this purpose a large representative data set that links employers and its employees was used.

The findings of this research suggest four primary conclusions. First, we find that the effects of displacement on earnings are considerable in magnitude and quite persistent. Three years after displacement the average hourly earnings differential of displaced workers relative to non-displaced workers has increased by around 10 and 12 percentage points for women and men respectively. A slight recovery in earnings is observed only in the last year. These estimates are near the lower bound for the U.S.

Second, the empirical results revealed that the deepening in the earnings differential between displaced and non-displaced workers is in large part due to the loss of tenure within the firm and, to a lesser extent, to industry changing, suggesting that job displacement leads to the destruction of firm (industry)-specific features that positively influence wages.

Third, aside from the losses that stem from the loss of job tenure, workers who take a long time to find a new job suffer an additional penalty on earnings that is directly related with time spent out of employment. Indeed, workers who experienced a spell of nonemployment have an additional annual penalty on average hourly earnings of 3 to 6 percentage points.

In sum, the results revealed that for male workers the loss of tenure on the job explains about 45.5% of the increase in the post-displacement earnings gap relative to non-displaced workers; joblessness accounts also for 45.5% of that increase; and the remaining 9% is explained by the loss of sector-specific features. As for female workers, the loss of accumulated returns to tenure account for 50% of the increase in the earnings differential; joblessness accounts for 33%; and the loss of sector-specific features explains the remaining 17%.

Fourth, the decomposition of the contribution of worker and firm characteristics to the difference in earnings showed that there are significant differences in earnings between displaced and non-displaced workers that are related to differences in firm characteristics between these two groups of individuals. In fact, the pre-displacement difference in earnings between displaced workers and the control group due to differences in firms characteristics such as size and sales per worker tend to persist in the post-displacement period. The regression results resemble this finding and point to the importance of controlling for employer characteristics when estimating the earnings differential between displaced and similar non-displaced workers, since a gap in earnings exists even before displacement due to systematic differences in firm characteristics for the two groups of workers. Ignoring them may confound the evaluation of the earnings losses.

6.2. Policy Implications

Some policy implications may be advanced from this study. Job displacement causes considerable and quite persistent earnings losses. In particular, we concluded that displaced workers often experience substantial adjustment costs that are mainly associated with the lengthy of nonemployment spells and with the loss of firm (industry)-specific skills. In this context, public policies that aim to improve the reemployment prospects of displaced workers such as job search assistance and training programs seem to have an appropriate role in order to minimize the earnings losses that job displacement imposes on workers. The objective of job search assistance is basically to assist job-ready workers to develop effective job-seeking skills, which may contribute to increase the probability of reemployment of displaced workers and to minimize the negative effects on earnings caused by unemployment. Training or retraining programs have been emphasized as a tool for increasing the long-term earnings of displaced workers by enhancing their skills. Our results seem to reveal that displaced workers have some incentives to be enrolled in skill training programs while unemployed. Thus, public training programs that aim to assist in the replenishment of spe-cific human capital lost when a displacement occurs or/and to enhance general skills may improve the reemployment prospects of displaced workers and may increase their long-term earnings.

Altering the advance notice legislation by increasing the period of notification and imposing severe penalties to firms that do not compliance with the law could also avoid some losses, mainly through a reduction in the incidence of unemployment. In the same sense, a more general policy implication is that the maximum duration of benefits should be reduced. The unemployment benefit should be seen as a temporary source of protection and not as a permanent charge. In fact, the current profile of the unemployment benefit system, which brings together a high replacement rate and a perspective of a long maximum period for the granting of the benefit, may contribute to reducing the transitions from unemployment to employment and to increase unemployment duration[28].

Finally, this study seems to reveal that workers in firms at risk are able to accept wage moderation in order to avoid the firm's shutdown. This indication should be matched with the observation that wages have a positive impact on the failure rate and that firms with a higher incidence of minimum wage earners are more vulnerable to adverse demand shocks due to their inability to adjust wages downward (Carneiro and Portugal, 2003). In this context, policy measures such as the temporary lay-off may have an effect of reducing the high incidence of firm closings that characterize the Portuguese economy. In fact, the empirical evidence for Portugal suggests that due to higher adjustment costs (mainly the costs of firing workers) and in the face of unforeseen temporary shocks, it is preferable to employers, under certain circumstances, to close down instead of adjusting their level of employment by laying off workers (Blanchard and Portugal, 2001).

[28] It should be noted once again that only 45% of the displaced workers were reemployed until the end of the observation period.

APPENDIX A - Descriptive Statistics

Table A.1

Descriptive Statistics (1991-98)
Displaced Male Workers

	MEAN	STDV	MIN	MAX
Variables				
Age (in years)	35.8	10.7	16.0	64.0
Education (in years)	6.0	2.9	0.0	16.0
Tenure (in years)	5.6	7.1	0.0	50.0
Qualification Level (%)				
Manager and Highly Professional	0.020		0.0	1.0
Professional	0.019		0.0	1.0
Supervisors	0.053		0.0	1.0
Highly Skilled and Skilled	0.551		0.0	1.0
Semi-skilled and Unskilled	0.234		0.0	1.0
Apprentices	0.077		0.0	1.0
Non-defined	0.046		0.0	1.0
Average Real Hourly Earnings (in logs)	5.948	0.46	4.55	8.38
Firm Size (in logs)	3.379	1.69	0.00	9.99
Firm Sales per Worker (in logs)	8.483	1.19	0.25	15.51
Number of Observations	151489			
Number of Individuals	60827			

Table A.2

Descriptive Statistics (1991-98)
Non-displaced Male Workers

	MEAN	STDV	MIN	MAX
Variables				
Age (in years)	39.6	10.4	16.0	64.0
Education (in years)	6.3	3.2	0.0	16.0
Tenure (in years)	12.4	8.7	0.0	61.0
Qualification Level (%)				
Manager and Highly Professional	0.033		0.0	1.0
Professional	0.036		0.0	1.0
Supervisors	0.074		0.0	1.0
Highly Skilled and Skilled	0.591		0.0	1.0
Semi-skilled and Unskilled	0.203		0.0	1.0
Apprentices	0.032		0.0	1.0
Non-defined	0.032		0.0	1.0
Average Real Hourly Earnings (in logs)	6.257	0.55	4.92	8.45
Firm Size (in logs)	4.868	2.10	0.00	10.29
Firm Sales per Worker (in logs)	8.956	1.05	0.92	17.29
Number of Observations	621615			
Number of Individuals	105506			

Table A.3

Descriptive Statistics (1991-98)
Displaced Female Workers

	MEAN	STDV	MIN	MAX
Variables				
Age (in years)	32.9	9.5	16.0	64.0
Education (in years)	6.2	2.9	0.0	16.0
Tenure (in years)	5.3	6.7	0.0	45.0
Qualification Level (%)				
Manager and Higlily Professional	0.009		0.0	1.0
Professional	0.009		0.0	1.0
Supervisors	0.021		0.0	1.0
Highly Skilled and Skilled	0.446		0.0	1.0
Semi-skilled and Unskilled	0.352		0.0	1.0
Apprentices	0.125		0.0	1.0
Non-defined	0.038		0.0	1.0
Average Real Hourly Earnings (in logs)	5.719	0.39	4.62	8.39
Firm Size (in logs)	3.472	1.74	0.00	9.87
Firm Sales per Worker (in logs)	8.098	1.17	-0.09	16.13
Number of Observations	115753			
Number of Individuals	45722			

Table A.4

Descriptive Statistics (1991-98)
Non-displaced Female Workers

	MEAN	STDV	MIN	MAX
Variables				
Age (in years)	35.9	9.7	16.0	64.0
Education (in years)	6.5	3.1	0.0	16.0
Tenure (in years)	11.0	8.1	0.0	62.0
Qualification Level (%)				
Manager and Higlily Professional	0.014		0.0	1.0
Professional	0.018		0.0	1.0
Supervisors	0.028		0.0	1.0
Highly Skilled and Skilled	0.486		0.0	1.0
Semi-skilled and Unskilled	0.375		0.0	1.0
Apprentices	0.055		0.0	1.0
Non-defined	0.024		0.0	1.0
Average Real Hourly Earnings (in logs)	5.944	0.51	4.92	8.43
Firm Size (in logs)	4.587	1.98	0.00	10.29
Firm Sales per Worker (in logs)	8.557	1.08	0.92	15.49
Number of Observations	385741			
Number of Individuals	67118			

APPENDIX B - Regression Results: Men and Women

Table B.1

Fixed Effects Results
Dependent variable: log of Average Real Hourly Earnings

	Men		Women	
Independent Variables	Coef.	t-ratio	Coef.	t-ratio
Disp_3	0.010*	(1.8)	0.007*	(1.1)
Disp_2	-0.004*	(-0.7)	-0.003*	(-0.4)
Disp_1	-0.029	(-5.1)	-0.017	(-2.6)
$Disp_o$	-0.043	(-6.8)	-0.029	(-4.0)
$Disp_1$	-0.045	(-7.4)	-0.031	(-4.6)
$Disp_2$	-0.050	(-8.3)	-0.036	(-5.3)
$Disp_3$	-0.059	(-9.6)	-0.039	(-5.7)
$Disp_4$	-0.068	(-10.6)	-0.037	(-5.2)
Age Squared[a]	-0.049	(-82.6)	-0.036	(-46.3)
Education	0.004	(8.6)	0.006	(12.0)
Part-time Job	0.126	(65.7)	0.125	(62.8)
Early-leaver	0.024*	(1.7)	-0.013*	(-0.8)
$\bar{R}^2$	0.88		0.86	
Number of Observations	773104		501494	

Notes:

(i) a set of time dummies is included in the specification;

(ii) [a] variables divided by 100;

(iii) all estimates are significant at 1%, except those with an *.

Table B.2

WLS Results
Dependent variable: log of Average Real Hourly Earnings

	Men		Women	
Independent Variables	Coef.	t-ratio	Coef.	t-ratio
$Disp_{BD}$	-0.010*	(-2.3)	0.005*	(1.3)
$Disp_{AD}$	0.004*	(0.4)	0.057	(6.2)
Age	0.043	(144.4)	0.019	(57.6)
Age Squared[a]	-0.042	(-114.0)	-0.014	(-32.5)
Education	0.071	(488.6)	0.07	(417.4)
Part-time Job	0.092	(30.4)	0.11	(39.1)
Early-leaver	0.002*	(0.3)	-0.003*	(-0.4)
Size	0.057	(226.2)	0.054	(209.4)
Sales per Worker	0.114	(258.2)	0.095	(196.1)
OE (Out of Employment)	-0.065	(-5.3)	-0.039	(-3.1)
JC (Job Changes)	-0.012*	(-0.7)	-0.014*	(-0.8)
CIND (Changed Industry)	-0.067	(-5.3)	-0.133	(-10.0)
Tenure	0.006	(35.6)	0.013	(63.6)
Tenure Squared[a]	-0.001*	(-2.5)	-0.025	(-38.0)
Constant	3.443	(522.7)	3.85	(566.2)
$\overline{R}^2$	0.62		0.64	
Number of Observations	773104		501494	

Notes:

(i) a set of time, industry and regional dummies is included in the speciíication;

(ii) [a] variables divided by 100;

(iii) all estimates are significant at 1%, except those with an *.

APPENDIX C - Selectivity Test

Table C (cont.)

Sample Selection Model - Displaced Workers

	Reemployment Equation		Wage Equation		Wage Equation	
	(REEMP=1)		Selectivity Correction		No Selectivity Correction	
	(1)		(2)		(3)	
Independent Variables	Coef.	t-ratio	Coef.	t-ratio	Coef.	t-ratio
Age	0.011	(3.8)	0.03	(21.6)	0.03	-21.5
Age Squared[a]	-0.042	(-10.8)	-0.029	(-15.6)	-0.028	(-15.2)
Education	-0.005	(-2.7)	0.055	(77.6)	0.056	-78.6
Tenure	-0.015	(-7.6)	0.035	(10.7)	0.035	-10.7
Tenure Squared[a]	0.013*	(1.8)				
Female	-0.093	(-10.6)	-0.213	(-53.9)	-0.211	(-54.1)
Early-leaver	-0.481	(-43.2)				
Manufacturing	0.108	(11.9)	-0.001*	(-0.3)	-0.004*	(-0.8)
Log Hourly Earnings	-0.065	(-5.8)				
Log Unemployment Rate	-0.284	(-10.6)				
Part-time Job			0.085	(12.2)	0.086	-12.3
Size			0.04	(37.9)	0.04	-38
Sales per Worker			0.0006*	(1.1)	0.0007*	-1.5
$Disp_1$			-0.043	(-7.3)	-0.046	(-7.9)
$Disp_2$			-0.056	(-7.5)	-0.061	(-8.3)

Table C (cont.)

Sample Selection Model - Displaced Workers

	Reemployment Equation		Wage Equation		Wage Equation	
	(REEMP=1)		Selectivity Correction		No Selectivity Correction	
$Disp_3$			-0.048	(-4.8)	-0.056	(-5.7)
$Disp_4$			-0.065	(-4.6)	-0.076	(-5.6)
Constant	0.874	-10.3	4.692	-165.1	4.733	-181.7
Time Dummies	no		yes		yes	
Regional Dummies	yes		yes		yes	
λ			0.05	-3.5		
ρ			0.14			
Chi Squared; $\overline{R}^2$	5798.3		0.31		0.31	
Number of Observations	105507		33783		33783	

Notes:

(i) REEMP: dummy variable equal to 1 if reemployed (0 otherwise); Female: 1 if female (0 if male); Manufacturing: 1 if employed in manufacturing (0 otherwise); Unemployment rate: aggregate (source - INE).

(ii) [a]variables divided by 100; (iii) all estimates are significant at 1%, except those with an *;

References

Abraham, K. and H. Farber (1987), "Job Duration, Seniority, and Earnings", *American Económica Review*, Vol. 77, 278-297.

Abraham, K. and H. Farber (1988), "Returns to Seniority in Union and Nonunion Jobs", *Industrial and Labor Relations Review*, Vol. 42, 3-19.

Addison, J. and P. Portugal (1989), "Job Displacement, Relative Wage Changes, and Duration of Unemployment", *Journal of Labor Economics*, Vol. 7, 281-302.

Altonji, J. and R. Shakotko (1987), "Do Wages Rise with Job Seniority", *Review of Economic Studies*, Vol. 54, 437-459.

Arulampalam, W. (2001), "Is Unemployment Really Scarring? Ef-fects of Unemployment Experience on Wages", *Economic Journal*, Vol. 111, F585-F606.

Bender, S., C. Dustmann, D. Margolis and C. Meghir (1999), "Worker Displacement in France and Germany", Institute for Fiscal Studies Working Paper N° 99/14.

Blanchard, O. and P. Diamond (1994), "Ranking, Unemployment Duration and Wages", *Review of Economic Studies*, Vol. 61, 417-434.

Blanchard, O. and P. Portugal (2001), "What Hides Behind an Unemployment Rate: Comparing Portuguese and U.S. Labor Markets", *American Economic Review*, Vol. 91, 187-207.

Blanchflower, D., A. Oswald and M. Garrett (1990), "Insider Power in Wage Determination", *Económica*, Vol. 57, 143-170.

Bowlus, A. and L. Vilhuber (2002), "Displaced Workers, Early Leavers, and Re-employment Wages", *mimeo*.

Burda, M. and A. Mertens (2001), "Estimating Wage Losses of Displaced Workers in Germany", *Labour Economics*, Vol. 8,15-41.

Carneiro, A. and P. Portugal (2003), "Wages and the Risk of Displacement", Banco de Portugal Working Paper WP 10-03.

Carrington, W. (1993), "Wage Losses for Displaced Workers: Is It Really the Firm That Matters?", *Journal of Human Resources*, Vol. 28, 493-496.

Dustmann, C. and C. Meghir (1999), "Wages, Experience and Se-niority", Institute for Fiscal Studies Working Paper N° 01/01.

Fallick, B. (1996), "A Review of the Recent Empirical Literature on Displaced Workers", *Industrial and Labor Relations Review*, Vol. 50, 5-16.

Gregory, M. and R. Jukes (2001), "Unemployment and Subsequent Earnings: Estimating Scarring Among British Men 1984-94", *Eco-nomic Journal*, Vol. 111, F607-F625.

Hamermesh, D. (1989), "What Do We Know About Worker Dis-placement in the U.S.?", *Industrial Relations*, Vol. 28, 51-60.

Hildreth, A. and A. Oswald (1997), "Rent-sharing and wages: ev-idence from company and establishment panels", *Journal of Labor Economias*, Vol. 15, 318-337.

Jacobson, L., R. LaLonde and D. Sullivan (1993), "Earnings Losses of Displaced Workers", *American Economic Review*, Vol. 83, 685-709.

Jovanovic, B. (1979), "Job Matching and the Theory of Turnover", *Journal of Political Economy*, Vol.87, 972-990.

Kletzer, L. (1989), "Returns to Seniority After Permanent Job Loss", *American Economic Review*, Vol. 79, 536-543.

Kletzer, L. (1991), "Earnings After Job Displacement: Job Tenure, Industry, and Occupation" in J. Addison ed., *Job Displacement: Consequences and Implications for Policy*, Detroit: Wayne State University Press, 107-135.

Kletzer, L. (1998), "Job Displacement", *Journal of Economic Perspectives*, Vol. 12, 115-136.

Kouch, K. (2001), "Earnings Losses and Unemployment of Displaced Workers in Germany", *Industrial and Labor Relations Review*, Vol. 54, 559-572.

Kruse, D. (1988), "International Trade and the Labor Market Expe-rience of Displaced Workers", *Industrial and Labor Relations Review*, Vol. 37, 557-569.

Lefranc, A. (2002), "Labor Market Dynamics and Wage Losses of Displaced Workers in France and the United States", *mimeo*.

Lehmann, H., K. Phillips and J. Wadsworth (2002), "The Incidence and Cost of Job Loss in a Transition Economy: Displaced Workers in Estónia, 1989-1999", IZA Discussion Paper N° 567.

Lockwood, B. (1991), "Information Externalities in the Labour Market and the Duration of Unemployment", *Review of Economia Stud-ies*, Vol. 58, 733-753.

Machado, J. and J. Mata (1998), "Earning Functions in Portugal 1982-1994: Evidence from Quantile Regressions", Banco de Portugal Working Paper WP 2-98.

Margolis, D. (1999), "Part-Year Employment, Slow Reemployment, and Earnings Losses: The Case of Worker Displacement in France", in J. Haltiwanger *et al.* eds., *The Creation and Analysis of Employer-Employee Matched Data - Contributions to Economic Analysis*, North-Holland: Elsevier Science B.V., 375-416.

Nickell, S. and S. Wadhwani (1990), "Insider Forces and Wage De-termination", *Economic Journal*, Vol. 100, 496-509.

Phelps, E. (1972), *Inflation Policy and Unemployment Theory*, Lon-don: Macmillan.

Pissarides, C. (1992), "Loss of Skill during Unemployment and the Persistence of Employment Shocks", *Quarterly Journal of Economias*, Vol. 107, 1371-1391.

Podgursky, M. and P. Swaim (1987), "Job Displacement and Earnings Loss: Evidence from the Displaced Worker Survey", *Industrial and Labor Relations Review*, Vol. 41, 17-29.

Rosolia, A. (2002), "The Consequences of Job Displacement in Italy", *mimeo*.Ruhm, C. (1991), "Are Workers Permanently Scarred by Job Dis-placements?", *American Economic Review*, Vol. 81, 319-324.

Stevens, A. (1997), "Persistent Effects of Job Displacement: The Importance of Multiple Job Losses", *Journal of Labor Economias*, Vol. 15, 165-188.

Topel, R. (1991), "Specific Capital Mobility, and Wages: Wages Rise With Job Seniority", *Journal of Political Economy*, Vol. 99, 145-175.

HOW MUCH MOBILITY? CAREERS, PROMOTIONS, AND WAGES

Francisco Lima[*]

Instituto Superior Técnico, Universidade Técnica de Lisboa and CEG-IST

Abstract

The objective of this paper is to study the determinants of job mobility and the effect of mobility on wages, considering not only the workers' career between firms, but also within firms, using a longitudinal matching employer-employee data set. The results obtained show a negative relationship between tenure and the probability of exit and that the new jobs tend to end early. Moreover, the career advancement within the firm decreases the probability to exit. Concerning wages, job separations can have a positive impact on wage growth, especially for the younger workers. Movements to larger firms or another industry can also be associated with positive wage growth. This shows that the workers' movements between employers and industries are important to enhance their career prospects.

Keywords: Job mobility, Wages, Careers

JEL-code: J31, J63, M51

[*] Correspondence: Instituto Superior Técnico, Departamento de Engenharia e Gestão, Centre for Management Studies, Avenida Prof. Cavaco Silva, Taguspark, 2780-990 Porto Salvo Portugal. Tel.: +351 214233514 E-mail: francisco.lima@tagus.ist.utl.pt

The author thanks the helpful comments from Mário Centeno, Pedro Portugal, the Scientific Committee of the II Conferência do Banco de Portugal and the participants of the conference. The author also gratefully acknowledges the financial support from *Fundação para a Ciência e a Tecnologia* (grant POCTI/1999/ECO/33089) and the Bank of Portugal for support with the data. The usual disclaimer applies.

1. Introduction

The workers' mobility is rarely studied within the firm. When a firm hires a worker, he is assigned to a certain hierarchical level, and then what happens? Has the worker more chances of progressing within a firm or by moving to another? How important are employer and industry changes at the beginning of the working life? What are the differences of career progression by gender? These are the main questions that we try to answer in the case of Portugal.

The data set used is *Quadros de Pessoal* (QP) from the Portuguese Ministry of Employment. The survey allows tracking firms and workers. In addition to the typical individual information found in this type of survey, the QP has information about the workers' job assignment, thereby, providing a way to reconstruct the workers' careers. Furthermore, there is information about promotions, which enable identifying several career events, namely, horizontal promotions, vertical promotions, and demotions.[1]

This type of information makes it possible to depart further from previous research, which in any case is not abundant. Typically, the applied work rests on information provided by single firm data (Baker, Gibbs and Holmstrom, 1994a,b, stands as the best example), and individual based surveys like the NLSY (e.g., Pergamit and Veum, 1999) or the PSID (e.g., McCue, 1996). The QP has features from both types of data sets, since it covers the entire workforce of multiple firms. Therefore, issues like mobility in the labour market from the perspective of the internal working of firms, which has been left much uncovered by past research can be studied. Moreover, the study of gender differentials in job assignment and promotions (Winter-Ebmer and Zweimüller, 1997; Booth, Francesconi and Frank, 2003) can be extended to other dimensions, namely, to the job mobility.[2]

The main objective is to study workers' mobility concerning not only wages, but also the career advancement within and between firms, using a longitudinal matching employer-employee data set. This subject is particularly important in the context of the Portuguese labour market, given the prevalence of numerous low-wage jobs, low qualifications, and several restrictions to the functioning of the market.

What have we reached so far? With the objective of assessing the impact of job mobility, the study started to look for the incidence of long and short-term employment relationships. The probability of job separation was estimated and in order to relate the job separation with the worker's internal career path, the promotion event and time since the previous promotion were introduced. The effect of job mobility on wages was analysed, and the main result is that job separation can have a positive impact on wage growth, especially for the younger workers and for those who move to larger firms or changes industry. This shows that the workers' movements between employers and industries are important to enhance their career prospects.

---

1 See Cabral and Mata (2003) and Lima and Pereira (2003) for recent papers using QP.

2 The research on the internal working of firms has dealt with such issues as, for example: the effect of promotions on wages (Lazear, 1999); the relationship between careers and wages considering the effects of job assignment, human capital, and the learning about the employees' abilities (Lima and Pereira, 2003); the importance of the employer learning effect and the effect of education on wages (Bauer and Haisken-DeNew, 2001); and the fast-track effect (Ariga, Ohkusa and Brunello, 1999).

The paper is organized as follows. Section 2 presents the data set and characterizes several facts about job mobility, namely the importance of employment relationships with different durations. The third section provides the econometric results for the determinants of job mobility and wages. Section 4 discusses the main policy implications. The final section concludes the paper.

2. Data and Overview

The data set used in this paper comes from *Quadros de Pessoal* (QP). This survey is gathered annually by the Portuguese Ministry of Employment and covers virtually all firms in the economy. It is a longitudinal matched employer-employee data set, given that both firms and workers can be followed from one year to the other. The sampling procedure constitutes on selecting 80 percent of all firms in each year from 1991 to 1999, stratified by firm size (number of workers). Thus, there are a total of 1,339,627 firm-years distributed as presented in Table 1. The share of firms with less than 10 employees is very high (80%), and this share increased during the 1990s, justifying the highest number of firms found in the sample. The QP have information about all the workers employed in these firms which corresponds to more than fourteen million worker-years in the data set.

[insert Table 1 about here]

Several restrictions were needed in order to study job mobility. First, due to missing values and to the restriction of workers' age between 16 and 64 years old, the sample of workers is reduced to 12,698,312 worker-years (see Table 2). Second, in order to clearly identify the workers who leave the current employer, that is, the job separations, the firms observed in two consecutive years and the respective workforce were selected. Part of the analysis rests on this sub-sample of firms, which makes 80 percent of the original sample. Moreover, this procedure has the advantage of creating a suitable comparison group for those workers who terminate a job – the workers who remain in the former employer and the workers who are already in the new employer. Thirdly, the analysis is conducted separately for female and male workers.

[insert Table 2 about here]

There are eight hierarchical levels defined by law in the QP: 1) apprentices, internships, trainees; 2) non-skilled professionals; 3) semi-skilled professionals; 4) skilled professionals; 5) higher-skilled professionals; 6) supervisors, team leaders; 7) intermediary management; 8) top management.[3] Table 3 shows summary statistics across hierarchical levels for selected individual characteristics. Age and tenure are increasing in the hierarchical level. Entry is concentrated on the bottom levels of the hierarchy across the board, but there is no evidence of unique ports-of-entry, and the same is true for exit.

[insert Table 3 about here]

Concerning job mobility, there are three main facts that generally appear in applied work (Farber, 1999): (1) high incidence of long-term employment relationships; (2)

[3] As the hierarchical level increases, the task complexity, skill requirement, and the responsibility level also increase. See Lima and Pereira (2003) for a detailed description of the levels.

new jobs are short lived; (3) the probability of job ending declines with tenure. To what extent are these facts also true for Portugal?

Table 4 presents the incidence of long-term employment relationships for workers with more than ten years of tenure and aged 35-64 years. In 1991, they represented more than half of the same age group and more than one quarter of the entire sampled workforce, showing the high incidence of this type of employment relationship (15 percentage points higher than the comparable figure in the CPS for the US).

[insert Table 4 about here]

At the end of the decade, the incidence remains high, but there is a huge drop of more than 10 percentage points (pp.), which could represent a major increase in job instability (see discussion for the US, for example, in Farber, 1999; Jaeger and Steves, 1999; Neumark, Polsky, and Hansen, 1999; Bernhardt et al., 1999; and Gottschalk and Moffit, 1999). The incidence is higher for males than for females, and for workers with primary and secondary education.

The individuals working 20 years for the same employer and aged 45-64 are presented in Table 5. The figures confirm what was already apparent in Table 4, that is, the high frequency of long-term employment relationships. Even in 1991, they comprised 30 percent of the same age group. However, they only represent nine percent of all workers, a proportion that remain stable during the 1990s, and with a higher incidence for males.

[insert Table 5 about here]

The second fact is that new jobs end early. As Table 6 shows, the fraction of workers who leave the firm within one year is more than 40 percent. Though the numbers of job separations are overestimated due to missing data, they clearly show the importance of job mobility early in the individuals' working life.

[insert Table 6 about here]

The negative relationship between tenure and job separation is another fact usually observed in the functioning of the labour market. As mentioned above, we only observe workers who separate from firms that are present two consecutive years, and there are some missing data problems. Bearing this in mind, Figure 1 graphs the empirical exit hazard rate by years of tenure computed from 1991 to 1999. There is a sharp decline in the probability of job ending as tenure increases, confirming the negative relationship between these two variables.[4] Note also that the probability of job separation increases at higher tenures, probably due to the proximity of the workers' retirement age.[5] Furthermore, Table 7 shows that workers who change jobs are younger and less experienced, on average, than those who stay in the current job.

[insert Figure 1 about here]

[4] The monthly information on the date of hiring provided in the QP can be used to increase the frequency of the observation and thus capture the peak found previously (see Farber, 1999) in the hazard within the first year of tenure.

[5] The probability of job ending with age also shows the same pattern as the one presented by the relationship between job ending and tenure: a sharp decline early in life, and then rise again near the sixties.

[insert Table 7 about here]

The overview of the data shows that job mobility follows the same type of behaviour found elsewhere (e.g. Farber, 1999). The following section will try to study the main determinants of job mobility and their effect on earnings. Given the descriptive results presented above, there is no doubt that this phenomenon needs to be analysed, moreover, when thinking about the workers' career achievement and the labour market policies that enhance it (or not).

3. Econometric Results

The objective of this section is to study the determinants of job mobility and their effect on earnings. Given the longitudinal nature of the data set used in this paper, we can follow workers from 1991 to 1999, allowing us to reconstruct the individual career, even when the worker moves between firms.[6] Thus, there are several variables used here that are the result of the workers individual history.

3.1. The Determinants of Job Mobility

The worker during his career can be fired by his employer or can quit. In either case, the final outcome is a job separation and this event is the object of study in this subsection. The goal is to characterize its determinants, knowing in advance that there is no distinction between quits and layoffs in the data set, which can blur the true relationships to be found. Nevertheless, the main issue is that an expected bad quality match is terminated when a job separation occurs, whatever the reasons to the termination (see Jovanovic, 1979a,b). A probit model is applied and the tenure in the job terminated is introduced as an independent variable with an expected negative coefficient (a duration dependence).

The results of the estimation are presented in Table 8 and Table 9. As already mentioned, the sub-sample used corresponds to firms observed in two consecutive years. Given that the dependent variable is the job separation, we observe which workers exit the firms in year t and report that information in $t - 1$.

[insert Table 8 about here]

The first specification presented in Table 8 includes the human capital variables and dummies for the hierarchical levels. The coefficient on tenure is negative, as expected, as well as the coefficient on age. Given the positive signs of the respective coefficients on the squared terms, the negative effect is less pronounced later in (working) life, which can be related to Figure 1, where the hazard for mobility has a pronounced negative slope for the first years within a firm and then it becomes flatter.[7] More educated workers have a lower probability to move. In addition, the part-time status has a strong positive effect on the probability of leaving a firm. These effects are robust to the several specifications presented. The hierarchical levels are introduced to account for the workers' current position within the firm. The estimated results show that as the worker progresses inside the

[6] This section extends previous research on career determinants (Lima and Pereira, 2003; Lima and Centeno, 2003) by focusing on the importance of careers between firms, that is, the effect of job mobility on individual labour market outcomes.

[7] And even increases later in the working life as shown in Figure 1.

firm, the less probable he is to move to another one, but at the very top of the hierarchy there is a reverse in the tendency described. This reversion can be related with careers ending or more visible workers at the top, which males easier the search for a better match.

Concerning the firm characteristics, the results in Table 8 show that there are industries where the probability of job separation is higher (the comparison industry is Manufacturing, the most prevalent) like Agriculture, a declining sector, or Construction and Commerce, Hotel and Restaurants, two sectors with high turnover rates. Concerning firm location, all regions have lower job separations probabilities when compared with Lisbon, being the North the region with the more negative effect. The results also show that the largest the firm size is, the lowest the job separation probability is.

Knowing that the new jobs tend to end early (see previous section), the second specification for the exit probability considers two dummies for tenure less than one year and tenure equal to one year. In fact, the associated coefficients are positive, that is, if a female (male) worker is in her(his) first year within the firm, she (he) has a 15(16) pp. higher probability of leaving it (the highest direct marginal effect found across variables in all specifications), and 7(8) pp. if the female (male) worker has completed one year of tenure.

The first specification in Table 9 explicitly introduces the workers' career within the firm by considering the worker's promotion in the year just before separation. The effect is negative, that is, the promoted worker has a lower probability of moving. Apparently this effect is expected, but when the worker is promoted, he also becomes more visible to the external labour market as a high ability worker, which can lead other potential employers to bid him away from the current one (Bernhardt, 1995).

[insert Table 9 about here]

The second specification adds time since last promotion. The effect is positive, though modest, which means that the more time the worker remains without a promotion, the more prone he becomes to leave the firm. Either the employee leaves to find a better match, or the employer identified the worker as a low ability one and makes him exit the firm.

The promotions event is further separated between horizontal and vertical promotions across hierarchical levels in the third specification. Demotions are also considered. Both types of promotions decrease the probability of job separation, but demotions increase it by 8 pp. for both females and males, showing that a better career progression within a firm implies on average a lower probability of leaving the firm to pursue career in a different firm (or of moving to self-employment, unemployment or inactivity).

Summarizing, the main results for the determinants of job mobility are:

(i) The more years the worker spends in a firm, lower is the probability of job separation, but this negative effect fades away with time, and the same is true for age.

(ii) The new hires have a higher probability of leaving the current job, which conforms the result of new jobs end early and meaning that mobility is an important issue earlier in the working life.

(iii) The part-time status has a strong positive effect on the probability of leaving a firm.

(iv) Workers in larger firms have a lower probability of leaving it. Lisbon is the firm location where the workers have a higher probability of separation. Some industries – declining or with high turnover rates – also imply a higher separation probability.

(v) As the worker progresses within the firm, lower is the probability of leaving it.

(vi) Related with the previous result, if the worker is promoted, lower is his exit probability. In addition, the more time has passed since the last promotion, higher the exit probability is. The same is true for demotions. Thus, the better the career within the firm is, lower is the necessity to find a new career in another firm.

3.2. Mobility and Wages

The path of individual wages is probably the main variable of interest when analysing the workers' labour market outcomes. Thus, this section provides a first approximation to the effect of job mobility on individual wages. The first set of results presented concern the typical Mincerian wage equation extended to the workers' career path, not only within firms, but also between firms. At the end of the section, wage growth equations are also estimated. The sample is restricted to the years 1995-1999, because the change in the rules of the QP from 1994 to 1995 (change in the industry and education coding) can introduce a small amount of noise in the estimation. Furthermore, fewer years allowed lifting some processing restrictions imposed by the too high initial sample size.

Table 10 provides the results for wage levels. Several specifications are presented because there are internal and external paths to consider (Table A1 presents the basic specifications). The first regression presented is a typical wage equation with the hierarchical levels added. Tenure and age have a quadratic influence on wages. The coefficients on tenure are within the range usually found for other countries. Better-educated workers earn more. These set of results conform the expected positive impact of human capital on earnings. The two dummies for lower tenure are capturing the workers' entry and the associated coefficients are negative, showing that new hires generally receive lower wages. The set of dummies for hierarchical levels show that climbing the job ladder translates into higher wages. The workers' promotion is added in the second specification. In spite of this effect appearing significant, it must be interpreted carefully, given that the main effect of a promotion shows up in wage growth and not in wage levels (Lima and Pereira, 2003).

[insert Table 10 about here]

The worker's job separation is also introduced in Table 10. In order to capture what happens before and after separation, two dummies are introduced. The first dummy is for job separation in the previous year, which means that the associated coefficient captures the wage of a new worker. This coefficient is positive for both females and males and higher than the coefficient on entry (*Tenure = 0*), which means that the job move yields

positive returns.[8] The second dummy is for contemporaneous job change, that is, the negative coefficient separations imply that those workers who are about to leave the firm earn less on average than those who stay.

The main results concerning wage levels are:

(i) Tenure and age have the standard effect on wages.
(ii) Wages are increasing with the hierarchical level.
(iii) Promotions have a positive affect on wages
(iv) The workers who exit the firm earn lower wages on average before the job separation (old firm), but earn higher wages after the job separation (new firm).

The results presented in Table 10 for the wage level show the importance of the wage dynamics for the study of mobility. Therefore, the wage growth is estimated in Table 11 and Table 12 as a first-difference wage equation. In this way, not only it is possible to determine the wage premiums associated with the worker's career within the firm (Lima and Pereira, 2003), but also the worker's career between firms – the earnings effect of job mobility.

[insert Table 11 about here]

The first specification in Table 11 includes promotion and job separation. Concerning job separation, the timing of the variable is such that the wage growth is the wage after separation minus the wage before separation. The results of this first specification show positive wage premiums both for promotions and for job change, and especially for male workers, the effect for job change is close to the effect for promotion, three percent against four percent. The second specification considers the type of promotion and demotions and, as expected, the vertical promotion yields the highest effect, 8 percent for females and 10 percent for males. Demotions have a negative wage premium associated of one percent for both type of workers. The third and final specification of the wage growth regression in Table 11 adds the workers' education in order to capture possible wage-tenure profiles with respect to the level of education. In fact, more educated workers experience higher wage growth.

The next set of results considers several individual and firm attributes and its relationship with wage growth and job separation (Table 12). Column (1) and (2) accounts for the possible differences of the job separation effect across worker's age. For each age group there is a binary variable that interacts with the job separation variable. The results show a clear pattern: the gains from job change are higher for the younger workers; as the workers get older, the positive effect decreases; by the age of thirty-five it starts to be negative; and by the age of fifty it clearly becomes negative for male workers. These results show the importance of job movements earlier in the workers' career and the benefits of those movements (Topel and Ward, 1992).

The level of education upon job move is considered in the third specification (Bernhardt et al., 1999). The estimated coefficients on education show that the gains from moving are higher for the more educated workers and almost null for the workers with the

[8] The variable *Tenure = 0* captures entry of job movers and also entry of previous unemployed or inactive.

primary education, the lowest level. The gains are higher for males than females at the lower levels of education.

Finally, the change of firm characteristics is considered in the last two columns, namely the change in industry, location, and size. The job change that also implies a change in industry holds a positive premium, which is not expected provided that specific capital is considered.[9] This means that on average the workers do not loose from changing industry. The study of mobility between industries can clarify if there are career matches and employer matches, that is, if the young workers follow a job search strategy with two stages: first search for a career and second search for an employer (Neal, 1999). If there is a high proportion of industry changes, then it is more than merely searching for an employer within a given career path. The high wage premium associated with job change for the younger workers in the first specification of Table 12 seems to confirm this argument of career search before looking for an employer. The change in firm location holds on average a not significant effect meaning that workers do not suffer a negative effect for geographic mobility.[10]

[insert Table 12 about here]

The change in form size upon job change is separated in changes to larger firms and to smaller ones. As expected, given the usual wage-firm size positive relationship (see Table 10), the movements to larger firms have an associated positive wage premium, showing the alternative to a career within a firm is a change to another firm with a larger workforce, and therefore on average with higher potential gains from career progression. The reverse is apparently true: the move to smaller firms carries a negative premium for both females and males.

Concerning wage growth, the following main results were obtained:

(i) Promotions, and especially vertical promotions, have a positive effect on wage growth. The reverse is true for demotions.
(ii) Job separation yields a positive wage premium for both female and male workers.
(iii) The positive job change effect on wage growth is higher for the younger workers. As the workers get older, the positive effect decreases and may become negative.
(iv) The gains from job change are higher for more educated workers. The workers with the lowest level of education – primary – have no significant gains from changing employer.
(v) The change of industry upon job separation has a positive effect on wage growth. This result can be linked to the gains of job change for young

[9] The change in industry was calculated with the two digit codes and not considering possible differences in the similarity between industries, and this fact calls for a finer definition of the variable. The change of industry can be used to tackle different rewards individual attributes, namely education and labour market experience, that is, make wage comparisons of industry movers and stayers (McLaughlin and Bils, 2001). The above results seem to indicate that some movement hold positive gains, and these movements can be from industries that are declining to industries that are growing.

[10] As with industry changes, the region changes can be further detailed.

workers, indicating a search for a career across sectors of activity before searching for a specific employer.

(vi) Geographical mobility does not have a significant impact on wage growth.

(vii) Moving to larger firms holds a positive wage premium, but the reverse is true for movements to smaller ones.

4. Policy Implications

What are the policy implications? Firstly, the negative job separation-tenure relationship and the related negative job separation-age relationship shows the importance of having a more flexible labour market in order to allow workers and firms to find the better quality match. At the same time, given the prevalence of long-term employment relationships, it also shows the danger of an emergence of a "segmented" labour market: one where workers are protected and can be considered insiders; and another where workers at the beginning of the working life do not have the conditions to find a stable career, given that employers face an asymmetrical labour market legislation that leads them to use the unprotected side of the market as a buffer stock to respond to product market shocks. The strong positive effect associated with the part-time status seems to corroborate this view, showing a high job volatility associated with this stage.

The results on the relationship between wages and job separation show that job mobility can have a positive effect on wages, especially for the younger workers. This fact reinforces the need for a flexible labour market that allows the workers at the beginning of their working lives to search not only for a high quality employer-employee match, but also a high quality industry-employee match, meaning that the young worker can engage in a career search – sector of activity – before searching for best employer.

In addition, the results also show that job mobility can be a way of moving up in the wage scale and reducing the weight of dead-end low wage jobs.[11] However, the modest or no gains from mobility for the less educated workers call for a special concentration on policies concerning training and the associated accumulation of human capital. The analysis did not focus on the distinction between specific and general human capital (Topel, 1991; Farber, 1999), however the results on industry change seem to indicate that this is not the main issue concerning the Portuguese reality, but rather a lack of human capital, irrespective of its nature. Thus, if human capital portability is not the main issue, the policies promoting workers' human capital accumulation gain especial relevance.

The prevalence of small and medium size firms, in face of the results obtained, adds another field of possible action. The policies that promote flexibility and the accumulation of skills are not real effective if they do not take in account this reality, namely, that in smaller firms the human resources practices are incipient and highly dependent on external labour market performance, as opposed to its own internal labour market, which can be considered almost inexistent.

[11] Similar to the analysing the conditional probability of transiting up or down the earnings distribution given the worker's past history (Finnie and Gray, 2002).

5. Conclusion

This paper studies several issues related with job mobility for the case of Portugal. The data set used is largely representative of the national labour market and provides enough information to follow firms and workers, namely to reconstruct the worker career path within and between firms. Given the current discussion in the European Union about the functioning of the labour market and of the several institutions that affect it, the issues treated here are of special relevance.

The main objective was first to identify the determinants of job movements between firms. The results obtained show a negative relationship between tenure and the probability of job separation and that new job tend to end early. Moreover, the career advancement within the firm decreases the probability of job change. The second main objective was to analyse the impact of job change on wages. The impact was found to be positive, although higher for the younger workers and for the more educated. As the workers get older, the positive effect decreases and can even become negative. These results were obtained controlling for the workers' career progression within the firm.

Overall, the results show the importance of having a labour market that allows workers to move between firms in order to find a better match. Institutions that harm this aspect of the labour market functioning can have a negative consequence on the final individual outcomes.

Table 1

Distribution of Firm Size

Survey year	Number of Firms	Firm size						
		<10	10-19	20-49	50-99	100-199	200-499	>=500
1991	115,091	87,313	13,818	8,796	2,919	1,276	696	273
1992	123,225	94,469	14,513	9,102	2,897	1,282	700	262
1993	128,532	99,995	14,672	8,861	2,858	1,247	653	246
1994	142,359	113,541	15,111	8,949	2,747	1,174	618	219
1995	149,071	120,319	15,124	8,841	2,782	1,168	614	223
1996	153,458	124,819	15,172	8,765	2,699	1,186	577	240
1997	165,578	135,104	16,295	9,323	2,846	1,186	597	227
1998	174,371	143,025	16,777	9,564	2,886	1,284	604	231
1999	187,942	154,147	18,414	10,142	3,055	1,305	632	247
Total	1,339,627	1,072,732	139,896	82,343	25,689	11,108	5,691	2,168

The total number of firms corresponds to an 80 percent sample of all firms covered by the QP stratified by firm size (number of workers).

Table 2

Number of Workers By Gender, 1991-1999

Survey year	Gender		Total
	Male	Female	
1991	796,129	494,300	1,290,429
1992	821,160	525,972	1,347,132
1993	832,337	528,239	1,360,576
1994	803,385	534,067	1,337,452
1995	838,374	572,309	1,410,683
1996	818,390	568,475	1,386,865
1997	874,174	622,097	1,496,271
1998	869,976	642,215	1,512,191
1999	885,062	671,651	1,556,713
Total	7,538,987	5,159,325	12,698,312

Table 3

Workers' Characteristics and The Hierarchy, 1991-1999 (Mean And Std Dev)

	Hierarchical levels								
	1	2	3	4	5	6	7	8	Total
Age	37.519	23.106	36.903	36.055	36.394	38.895	42.825	40.416	41.006
	(11.043)	(6.048)	(12.390)	(11.221)	(10.723)	(10.073)	(10.041)	(10.305)	(10.016)
Tenure	2.025	4.895	8.744	8.309	11.311	13.267	11.373	9.609	7.977
	(2.454)	(6.492)	(8.743)	(8.492)	(9.822)	(10.556)	(10.415)	(9.435)	(8.650)
Tenure < 1	0.246	0.224	0.111	0.115	0.078	0.058	0.088	0.090	0.130
	(0.431)	(0.417)	(0.314)	(0.319)	(0.269)	(0.233)	(0.284)	(0.286)	(0.337)
Tenure = 1	0.289	0.184	0.110	0.114	0.084	0.065	0.096	0.104	0.132
	(0.453)	(0.387)	(0.313)	(0.318)	(0.277)	(0.246)	(0.294)	(0.305)	(0.338)
Primary Education	0.280	0.564	0.534	0.432	0.171	0.429	0.100	0.045	0.413
	(0.449)	(0.496)	(0.499)	(0.495)	(0.376)	(0.495)	(0.300)	(0.208)	(0.492)
Secondary Education	0.423	0.169	0.218	0.345	0.585	0.366	0.497	0.297	0.326
	(0.494)	(0.375)	(0.413)	(0.475)	(0.493)	(0.482)	(0.500)	(0.457)	(0.469)
Tertiary Education	0.010	0.001	0.003	0.014	0.128	0.042	0.317	0.598	0.051
	(0.098)	(0.032)	(0.050)	(0.118)	(0.334)	(0.200)	(0.465)	(0.490)	(0.221)
Job Separation	0.359	0.259	0.255	0.217	0.221	0.244	0.264	0.276	0.099
	(0.480)	(0.438)	(0.436)	(0.412)	(0.415)	(0.429)	(0.441)	(0.447)	(0.299)

Age and *Tenure* are defined in years. *Job separation* is identified with the firms that remain at least two consecutive years in the sample. *Job Separation, Tenure < 1, Tenure = 1*, and the variables for education are defined as binary variables.

Table 4

Long-Term Employment, More Than 10 Years of Tenure (Ages 35-64)

Survey year	Total	Female	Male	Primary Education	Secondary Education	Tertiary Education	N
Fraction of the age group 35-64							
1991	0.554	0.540	0.560	0.532	0.604	0.478	634.962
1992	0.548	0.532	0.556	0.528	0.598	0.486	671.984
1993	0.546	0.512	0.564	0.527	0.600	0.498	693.454
1994	0.510	0.472	0.530	0.479	0.605	0.485	663.993
1995	0.486	0.457	0.503	0.468	0.561	0.444	697.716
1996	0.473	0.441	0.491	0.453	0.556	0.442	690.798
1997	0.458	0.426	0.477	0.436	0.543	0.440	760.288
1998	0.435	0.402	0.456	0.423	0.502	0.428	777.042
1999	0.423	0.391	0.445	0.418	0.479	0.410	803.063
Total	0.490	0.457	0.508	0.473	0.560	0.452	6.393.300
Fraction of all workers (all age groups)							
1991	0.272	0.224	0.302	0.311	0.202	0.273	1,290,429
1992	0.273	0.228	0.302	0.322	0.202	0.270	1,347,132
1993	0.279	0.225	0.312	0.333	0.212	0.276	1,360,576
1994	0.253	0.205	0.285	0.304	0.248	0.250	1,337,452
1995	0.241	0.199	0.269	0.303	0.223	0.225	1,410,683
1996	0.235	0.195	0.264	0.301	0.222	0.213	1,386,865
1997	0.233	0.196	0.259	0.298	0.220	0.204	1,496,271
1998	0.224	0.189	0.249	0.301	0.198	0.193	1,512,191
1999	0.218	0.185	0.244	0.302	0.187	0.178	1,556,713
Total	0.246	0.204	0.276	0.309	0.211	0.222	12,698,312

Table 5

Long-Term Employment, More Than 20 Years of Tenure (Ages 45-64)

Survey year	Total	Female	Male	Primary Education	Secondary Education	Tertiary Education	N
Fraction of the age group 45-64							
1991	0.329	0.277	0.348	0.304	0.415	0.305	313,013
1992	0.340	0.304	0.354	0.319	0.419	0.283	334,086
1993	0.356	0.308	0.374	0.334	0.438	0.316	351,635
1994	0.362	0.304	0.387	0.332	0.464	0.346	330,015
1995	0.376	0.333	0.395	0.356	0.461	0.328	344,125
1996	0.372	0.328	0.393	0.348	0.466	0.333	344,421
1997	0.369	0.330	0.388	0.339	0.475	0.336	379,413
1998	0.343	0.307	0.363	0.318	0.443	0.319	388,435
1999	0.326	0.292	0.344	0.301	0.420	0.306	399,513
Total	0.352	0.310	0.372	0.328	0.443	0.320	3,184,656
Fraction of all workers (all age groups)							
1991	0.080	0.047	0.100	0.088	0.056	0.076	1,290,429
1992	0.084	0.054	0.104	0.099	0.058	0.067	1,347,132
1993	0.092	0.057	0.114	0.112	0.064	0.077	1,360,576
1994	0.089	0.056	0.112	0.112	0.078	0.078	1,337,452
1995	0.092	0.061	0.113	0.124	0.072	0.071	1,410,683
1996	0.092	0.062	0.114	0.128	0.076	0.069	1,386,865
1997	0.093	0.066	0.113	0.131	0.078	0.067	1,496,271
1998	0.088	0.064	0.106	0.131	0.069	0.062	1,512,191
1999	0.084	0.062	0.100	0.127	0.065	0.056	1,556,713
Total	0.088	0.059	0.108	0.116	0.067	0.068	12,698,312

Table 6

Fraction of New Jobs That End Within 1 Year, 1991-1999

Survey year	Total	Female	Male	Primary Education	Secondary Education	Tertiary Education	N (Tenure≤1)
1991	0.448	0.409	0.472	0.460	0.436	0.392	214,112
1992	0.454	0.431	0.468	0.462	0.451	0.384	210,612
1993	0.526	0.506	0.539	0.537	0.519	0.468	189,920
1994	0.419	0.399	0.432	0.440	0.402	0.340	234,474
1995	0.414	0.397	0.426	0.436	0.392	0.333	255,378
1996	0.407	0.392	0.417	0.430	0.390	0.347	262,601
1997	0.407	0.398	0.414	0.418	0.408	0.343	295,224
1998	0.389	0.389	0.389	0.392	0.399	0.328	324,791
Total	0.427	0.410	0.439	0.444	0.427	0.354	1,987,112

Data for job separations of workers with one year of tenure at most as a fraction of all workers with one year of tenure at most. Separations identified with firms observed two consecutive years in the sample. The year reported corresponds to the last year before separation.

Table 7

Mean (Std. Dev.) Age and Tenure For Job Stayers and Movers, 1991-1999

	Age			Tenure		
	Total	Female	Male	Total	Female	Male
Job Stayer	36.745	34.815	38.043	9.006	8.112	9.607
	(11.200)	(10.488)	(11.473)	(8.774)	(8.117)	(9.140)
Job Mover	34.755	33.436	35.607	6.179	5.692	6.493
	(11.911)	(11.058)	(12.356)	(8.247)	(7.551)	(8.653)

Data for firms observed two consecutive years in the sample.

Table 8 (cont.)

Job Separation – Individual and Firm Characteristics: Marginal Effects (Probit), 1991-1999

	(1) Female	(2) Male	(3) Female	(4) Male
Tenure < 1	--	--	0.1526***	0.1606***
	--	--	(0.0011)	(0.0009)
Tenure = 1	--	--	0.0702***	0.0756***
	--	--	(0.0010)	(0.0008)
Tenure	-0.0229***	-0.0228***	-0.0112***	-0.0122***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
$Tenure^2 \times 10^{-2}$	0.0644***	0.0585***	0.0319***	0.0321***
	(0.0004)	(0.0003)	(0.0005)	(0.0003)
Age	-0.0088***	-0.0133***	-0.0102***	-0.0141***
	(0.0002)	(0.0001)	(0.0002)	(0.0001)
$Age^2 \times 10^{-2}$	0.0108***	0.0163***	0.0125***	0.0173***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Secondary education	-0.0129***	-0.0150***	-0.0115***	-0.0129***
	(0.0006)	(0.0005)	(0.0006)	(0.0005)
Tertiary education	-0.0196***	-0.0195***	-0.0166***	-0.0151***
	(0.0015)	(0.0012)	(0.0014)	(0.0012)
Part-time	0.1221***	0.1430***	0.1207***	0.1393***
	(0.0008)	(0.0008)	(0.0008)	(0.0008)
Hierarchical Levels				
Level 1	0.0265***	-0.0002	0.0161***	-0.0047***
	(0.0011)	(0.0010)	(0.0010)	(0.0010)
Level 2	0.0425***	0.0454***	0.0387***	0.0411***
	(0.0010)	(0.0008)	(0.0010)	(0.0008)
Level 3	0.0101***	0.0146***	0.0076***	0.0141***
	(0.0007)	(0.0007)	(0.0007)	(0.0007)
Level 5	-0.0027**	-0.0083***	-0.0026**	-0.0073***
	(0.0013)	(0.0010)	(0.0013)	(0.0010)
Level 6	0.0094***	-0.0104***	0.0088***	-0.0082***
	(0.0020)	(0.0010)	(0.0019)	(0.0010)
Level 7	0.0248***	0.0125***	0.0261***	0.0150***
	(0.0021)	(0.0014)	(0.0021)	(0.0013)
Level 8	0.0352***	0.0391***	0.0382***	0.0443***
	(0.0023)	(0.0015)	(0.0023)	(0.0015)
Log number of workers	-0.0045***	-0.0078***	-0.0054***	-0.0093***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Industry				
Agriculture	0.0598***	0.0019	0.0496***	0.0012
	(0.0024)	(0.0017)	(0.0023)	(0.0017)
Extraction	0.0342***	0.0233***	0.0345***	0.0240***
	(0.0081)	(0.0023)	(0.0080)	(0.0023)
Electricity, gas, and water	-0.0896***	-0.1112***	-0.0954***	-0.1117***
	(0.0038)	(0.0016)	(0.0037)	(0.0015)
Construction	0.0513***	0.0362***	0.0488***	0.0320***
	(0.0024)	(0.0007)	(0.0023)	(0.0007)
Commerce, Hotels, and Restaurants	0.0388***	0.0199***	0.0352***	0.0189***

Table 8 (cont.)

Job Separation – Individual and Firm Characteristics: Marginal Effects (Probit), 1991-1999

	(1) Female	(2) Male	(3) Female	(4) Male
	(0.0008)	(0.0007)	(0.0008)	(0.0006)
Transports and Communications	0.0163***	-0.0045***	0.0147***	-0.0072***
	(0.0016)	(0.0009)	(0.0015)	(0.0008)
Banking and Insurance	0.0274***	-0.0001	0.0210***	-0.0053***
	(0.0011)	(0.0009)	(0.0011)	(0.0009)
Collective, Social, and Personal Services	-0.0068***	-0.0107***	-0.0075***	-0.0105***
	(0.0009)	(0.0012)	(0.0009)	(0.0011)
Region				
North	-0.0432***	-0.0479***	-0.0409***	-0.0448***
	(0.0007)	(0.0005)	(0.0007)	(0.0005)
Centre	-0.0345***	-0.0299***	-0.0325***	-0.0275***
	(0.0008)	(0.0007)	(0.0008)	(0.0007)
Alentejo	-0.0103***	-0.0234***	-0.0127***	-0.0237***
	(0.0018)	(0.0013)	(0.0017)	(0.0013)
Algarve	-0.0029*	-0.0124***	-0.0081***	-0.0150***
	(0.0017)	(0.0015)	(0.0016)	(0.0015)
Islands	-0.0168***	-0.0366***	-0.0183***	-0.0369***
	(0.0016)	(0.0012)	(0.0016)	(0.0011)
Observations	3,191,103	4,798,522	3,191,103	4,798,522
Wald chi2	157,119.25	297,774.63	187,203.78	346,710.20
Log Likelihood	-1,771,340.7	-2,666,817.2	-1,760,617.3	-2,649,009.9
Pseudo R-squared	0.0496	0.0611	0.0554	0.0674

Dependent variable is the job separation identified with firms observed in two consecutive years. *Tenure*, *Time since last promotion* and *Age* are measured in years. Hierarchical levels are defined as binary variables, from apprentices, internships, and trainees (Level 1) and non-skilled professionals (Level 2) to top management (Level 8). Manufacturing and Lisbon are the comparison group in the industry and region dummies, respectively. All specifications include dummies for year. Robust standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 9 (cont.)

Job Separation and Promotions: Marginal Effects (Probit), 1991-1999

	(1) Female	(2) Male	(3) Female	(4) Male	(5) Female	(6) Male
Tenure < 1	0.1607***	0.1714***	0.1610***	0.1724***	0.1469***	0.1551***
	(0.0012)	(0.0010)	(0.0012)	(0.0010)	(0.0011)	(0.0009)
Tenure = 1	0.0710***	0.0731***	0.0718***	0.0747***	0.0554***	0.0604***
	(0.0010)	(0.0009)	(0.0010)	(0.0009)	(0.0010)	(0.0008)
Tenure	-0.0112***	-0.0121***	-0.0124***	-0.0132***	-0.0106***	-0.0119***
	(0.0002)	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0001)
$Tenure^2 \times 10^{-2}$	0.0321***	0.0319***	0.0361***	0.0354***	0.0311***	0.0318***
	(0.0005)	(0.0003)	(0.0006)	(0.0004)	(0.0004)	(0.0003)
Age	-0.0103***	-0.0145***	-0.0103***	-0.0145***	-0.0103***	-0.0144***
	(0.0002)	(0.0001)	(0.0002)	(0.0001)	(0.0002)	(0.0001)
$Age^2 \times 10^{-2}$	0.0127***	0.0178***	0.0127***	0.0179***	0.0127***	0.0175***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Secondary education	-0.0114***	-0.0131***	-0.0113***	-0.0129***	-0.0119***	-0.0133***
	(0.0007)	(0.0006)	(0.0007)	(0.0006)	(0.0006)	(0.0005)
Tertiary education	-0.0180***	-0.0169***	-0.0178***	-0.0167***	-0.0176***	-0.0157***
	(0.0015)	(0.0013)	(0.0015)	(0.0013)	(0.0014)	(0.0012)
Part-time	0.1199***	0.1359***	0.1197***	0.1357***	0.1205***	0.1380***
	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0008)	(0.0008)
Promotion	-0.0160***	-0.0245***	-0.0121***	-0.0197***	--	--
	(0.0008)	(0.0006)	(0.0009)	(0.0007)	--	--
Vertical Promotion	--	--	--	--	-0.0267***	-0.0318***
	--	--	--	--	(0.0024)	(0.0020)
Horizontal Promotion	--	--	--	--	-0.0296***	-0.0384***
	--	--	--	--	(0.0014)	(0.0011)
Demotion	--	--	--	--	0.0830***	0.0831***
	--	--	--	--	(0.0006)	(0.0005)
Time since last promotion	--	--	0.0023***	0.0026***	--	--
	--	--	(0.0002)	(0.0001)	--	--
Time since last $promotion^2 \times 10^{-2}$	--	--	-0.0090***	-0.0094***	--	--
	--	--	(0.0007)	(0.0005)	--	--
Hierarchical Levels						
Level 1	0.0172***	-0.0028***	0.0170***	-0.0028***	0.0110***	-0.0089***
	(0.0011)	(0.0010)	(0.0011)	(0.0010)	(0.0010)	(0.0010)
Level 2	0.0383***	0.0411***	0.0379***	0.0407***	0.0327***	0.0351***
	(0.0010)	(0.0008)	(0.0010)	(0.0008)	(0.0009)	(0.0008)
Level 3	0.0076***	0.0134***	0.0076***	0.0134***	0.0042***	0.0109***
	(0.0008)	(0.0007)	(0.0008)	(0.0007)	(0.0007)	(0.0007)
Level 5	-0.0045***	-0.0067***	-0.0046***	-0.0064***	-0.0035***	-0.0085***
	(0.0014)	(0.0011)	(0.0014)	(0.0011)	(0.0013)	(0.0010)
Level 6	0.0112***	-0.0085***	0.0111***	-0.0086***	0.0064***	-0.0090***
	(0.0021)	(0.0011)	(0.0021)	(0.0011)	(0.0019)	(0.0010)
Level 7	0.0270***	0.0158***	0.0266***	0.0156***	0.0210***	0.0103***
	(0.0022)	(0.0014)	(0.0022)	(0.0014)	(0.0020)	(0.0013)
Level 8	0.0390***	0.0470***	0.0389***	0.0469***	0.0360***	0.0418***
	(0.0024)	(0.0015)	(0.0024)	(0.0015)	(0.0022)	(0.0014)
Log number of workers	-0.0044***	-0.0080***	-0.0043***	-0.0079***	-0.0043***	-0.0079***

Table 9 (cont.)

Job Separation and Promotions: Marginal Effects (Probit), 1991-1999

	(1) Female	(2) Male	(3) Female	(4) Male	(5) Female	(6) Male
	(0.0002)	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Observations	2,831,157	4,205,729	2,831,157	4,205,729	3,191,103	4,798,522
Wald chi2	171,656.15	303,937.91	171,887.85	304,582.15	216,424.45	391,241.61
Log Likelihood	-1,556,603.6	-2,317,785.9	-1,556,513.6	-2,317,569.9	-1,748,442.3	-263,0701.6
Pseudo R-squared	0.0569	0.0668	0.0570	0.0669	0.0619	0.0738

Dependent variable is the job separation identified with firms observed in two consecutive years. *Tenure*, *Time since last promotion* and *Age* are measured in years. Hierarchical levels are defined as binary variables, from apprentices, internships, and trainees (Level 1) and non-skilled professionals (Level 2) to top management (Level 8). All specifications include dummies for region, industry and year. Robust standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 10 (cont.)

Wages, Job Separation, and Promotions: Wage Regression, 1995-1999

	(1) Female	(2) Male	(3) Female	(4) Male
Tenure < 1	-0.0061***	-0.0176***	-0.0061***	-0.0175***
	(0.0011)	(0.0011)	(0.0011)	(0.0011)
Tenure = 1	-0.0069***	-0.0201***	-0.0127***	-0.0207***
	(0.0020)	(0.0019)	(0.0021)	(0.0019)
Tenure	0.0142***	0.0092***	0.0143***	0.0092***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
$Tenure^2 \times 10^{-2}$	-0.0238***	-0.0133***	-0.0239***	-0.0133***
	(0.0006)	(0.0005)	(0.0006)	(0.0005)
Age	0.0164***	0.0300***	0.0165***	0.0301***
	(0.0003)	(0.0002)	(0.0003)	(0.0002)
$Age^2 \times 10^{-2}$	-0.0159***	-0.0307***	-0.0161***	-0.0307***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Secondary education	0.2347***	0.1950***	0.2345***	0.1950***
	(0.0010)	(0.0010)	(0.0010)	(0.0010)
Tertiary education	0.5534***	0.5106***	0.5532***	0.5106***
	(0.0028)	(0.0027)	(0.0028)	(0.0027)
Part-time	0.1423***	0.1820***	0.1425***	0.1820***
	(0.0022)	(0.0025)	(0.0022)	(0.0025)
Job Separation (t-1)	0.0159***	0.0266***	0.0158***	0.0266***
	(0.0021)	(0.0019)	(0.0021)	(0.0019)
Job Separation (t)	-0.0135***	-0.0092***	-0.0135***	-0.0092***
	(0.0011)	(0.0010)	(0.0011)	(0.0010)
Promotion	--	--	0.0183***	0.0017**
	--	--	(0.0010)	(0.0009)
Hierarchical Levels				
Level 1	-0.1343***	-0.1711***	-0.1349***	-0.1712***
	(0.0013)	(0.0014)	(0.0013)	(0.0014)
Level 2	-0.2078***	-0.2430***	-0.2073***	-0.2429***
	(0.0013)	(0.0012)	(0.0013)	(0.0012)
Level 3	-0.1165***	-0.1134***	-0.1163***	-0.1133***
	(0.0009)	(0.0010)	(0.0009)	(0.0010)
Level 5	0.2990***	0.2375***	0.2992***	0.2375***
	(0.0021)	(0.0017)	(0.0021)	(0.0017)
Level 6	0.2658***	0.2438***	0.2659***	0.2438***
	(0.0031)	(0.0017)	(0.0031)	(0.0017)
Level 7	0.4686***	0.4828***	0.4689***	0.4829***
	(0.0035)	(0.0024)	(0.0035)	(0.0024)
Level 8	0.6194***	0.6753***	0.6193***	0.6753***

Table 10 (cont.)

Wages, Job Separation, and Promotions: Wage Regression, 1995-1999

	(1) Female	(2) Male	(3) Female	(4) Male
	(0.0041)	(0.0030)	(0.0041)	(0.0030)
Log number of workers	0.0588***	0.0661***	0.0585***	0.0661***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Constant	0.2992***	0.2211***	0.2950***	0.2207***
	(0.0049)	(0.0046)	(0.0049)	(0.0046)
Observations	1,257,565	1,828,897	1,257,565	1,828,897
R-squared	0.65	0.64	0.65	0.64
F(df)	F(36;662900) =28840.70	F(36;947067) =48357.82	F(37;662900) =28097.89	F(37;947067) =47104.58

Dependent variable is the log real wage (Constant 1999 Euros). *Tenure* and *Age* are measured in years. *Job Separation (t-1)* and *Job Separation (t)* are binary variables equal to one in the moment of separation, respectively, in the year before (worker already in the new employer) and in the year of the separation (worker still in the previous employer). Hierarchical levels are defined as binary variables, from apprentices, internships, and trainees (Level 1) and non-skilled professionals (Level 2) to top management (Level 8). All specifications include dummies for region, industry and year. Robust standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 11

Wage Growth, Job Separation, and Promotions: First-Difference Wage Regressions, 1995-1999

	(1) Female	(2) Male	(3) Female	(4) Male	(5) Female	(6) Male
Promotion	0.0356***	0.0409***	--	--	--	--
	(0.0007)	(0.0006)	--	--	--	--
Job Separation (t-1)	0.0197***	0.0303***	0.0222***	0.0312***	0.0202***	0.0300***
	(0.0019)	(0.0017)	(0.0019)	(0.0017)	(0.0019)	(0.0017)
Vertical Promotion	--	--	0.0835***	0.0963***	0.0818***	0.0931***
	--	--	(0.0017)	(0.0017)	(0.0017)	(0.0017)
Horizontal Promotion	--	--	0.0263***	0.0322***	0.0248***	0.0308***
	--	--	(0.0008)	(0.0007)	(0.0008)	(0.0007)
Demotion	--	--	-0.0125***	-0.0130***	-0.0152***	-0.0158***
	--	--	(0.0012)	(0.0011)	(0.0012)	(0.0011)
Secondary education	--	--	--	--	0.0086***	0.0097***
	--	--	--	--	(0.0004)	(0.0004)
Tertiary education	--	--	--	--	0.0269***	0.0268***
	--	--	--	--	(0.0009)	(0.0008)
Constant	0.0563***	0.0487***	0.0426***	0.0451***	0.0526***	0.0457***
	(0.0004)	(0.0004)	(0.0002)	(0.0002)	(0.0005)	(0.0004)
Observations	1,215,124	1,788,735	1,215,124	1,788,735	1,215,124	1,788,735
R-squared	0.01	0.01	0.01	0.01	0.01	0.01
F(df)	F(5;644354) = 1136.58	F(5;929140) = 2060.20	F(4;644354) = 965.98	F(4;929140) = 1628.96	F(9;644354) = 874.55	F(9;929140) = 1504.32

Dependent variable is the log real wage in year *t* minus log real wage in year *t* – 1 (Constant 1999 Euros). All specifications include dummies for year. Robust standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 12

Wage Growth, Job Separation, Individual Attributes, and Firm Characteristics: First-Difference Wage Regressions, 1995-1999

	Job Separation and Age		Job Separation and Education		Job Separation and Firm Characteristics	
	(1) Female	(2) Male	(3) Female	(4) Male	(5) Female	(6) Male
Job Separation (t-1)	0.0436***	0.0843***	--	--	0.0440***	0.0196
	(0.0116)	(0.0090)	--	--	(0.0135)	(0.0122)
Vertical Promotion	0.0809***	0.0912***	0.0814***	0.0926***	0.0832***	0.0940***
	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0016)
Horizontal Promotion	0.0247***	0.0308***	0.0251***	0.0312***	0.0258***	0.0321***
	(0.0008)	(0.0007)	(0.0008)	(0.0007)	(0.0008)	(0.0007)
Demotion	-0.0154***	-0.0160***	-0.0150***	-0.0155***	-0.0138***	-0.0144***
	(0.0012)	(0.0011)	(0.0012)	(0.0011)	(0.0012)	(0.0011)
Secondary education	0.0085***	0.0094***	0.0077***	0.0080***	--	--
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	--	--
Tertiary education	0.0269***	0.0270***	0.0221***	0.0218***	--	--
	(0.0009)	(0.0008)	(0.0009)	(0.0008)	--	--
16 ≤ Age < 20 x Separation	0.0567***	0.0454***	--	--	--	--
	(0.0070)	(0.0052)	--	--	--	--
20 ≤ Age < 30 x Separation	-0.0135	-0.0323***	--	--	--	--
	(0.0120)	(0.0093)	--	--	--	--
30 ≤ Age < 35 x Separation	-0.0197	-0.0620***	--	--	--	--
	(0.0125)	(0.0098)	--	--	--	--
35 ≤ Age < 40 x Separation	-0.0520***	-0.0742***	--	--	--	--
	(0.0129)	(0.0101)	--	--	--	--
40 ≤ Age < 50 x Separation	-0.0373***	-0.0754***	--	--	--	--
	(0.0126)	(0.0099)	--	--	--	--
50 ≤ Age < 60 x Separation	-0.0474***	-0.0956***	--	--	--	--
	(0.0144)	(0.0107)	--	--	--	--
60 ≤ Age < 65 x Separation	-0.0392	-0.1177***	--	--	--	--
	(0.0270)	(0.0182)	--	--	--	--
Primary Ed. x Separation	--	--	0.0010	0.0089***	--	--
	--	--	(0.0024)	(0.0021)	--	--
Secondary Ed. x Separation	--	--	0.0283***	0.0503***	--	--
	--	--	(0.0032)	(0.0029)	--	--
Tertiary Ed. x Separation	--	--	0.1136***	0.1130***	--	--
	--	--	(0.0080)	(0.0070)	--	--
Change Industry x Separation	--	--	--	--	0.0065*	0.0247***
	--	--	--	--	(0.0039)	(0.0033)
Change Region x Separation	--	--	--	--	-0.0086	-0.0002
	--	--	--	--	(0.0054)	(0.0043)
Move to:						
Larger Firm	--	--	--	--	0.0319**	0.0670***
	--	--	--	--	(0.0137)	(0.0125)
Smaller Firm	--	--	--	--	-0.0800***	-0.0619***
	--	--	--	--	(0.0137)	(0.0124)
Constant	0.0525***	0.0457***	0.0532***	0.0465***	0.0569***	0.0498***
	(0.0005)	(0.0004)	(0.0005)	(0.0004)	(0.0004)	(0.0004)
Observations	1,215,124	1,788,735	1,215,124	1,788,735	1,215,124	1,788,735
R-squared	0.01	0.01	0.01	0.01	0.01	0.01
F(df)	F(16;644354) = 500.92	F(16;929140) = 869.90	F(11;644354) = 718.96	F(11;929140) = 1233.59	F(11;644354) = 692.75	F(11;929140) = 1208.30

Dependent variable is the log real wage in year *t* minus log real wage in year *t* – 1 (Constant 1999 Euros). Move to larger or smaller firm are dummy variables computed using the number of employees. All specifications include dummies for year. Robust standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A1

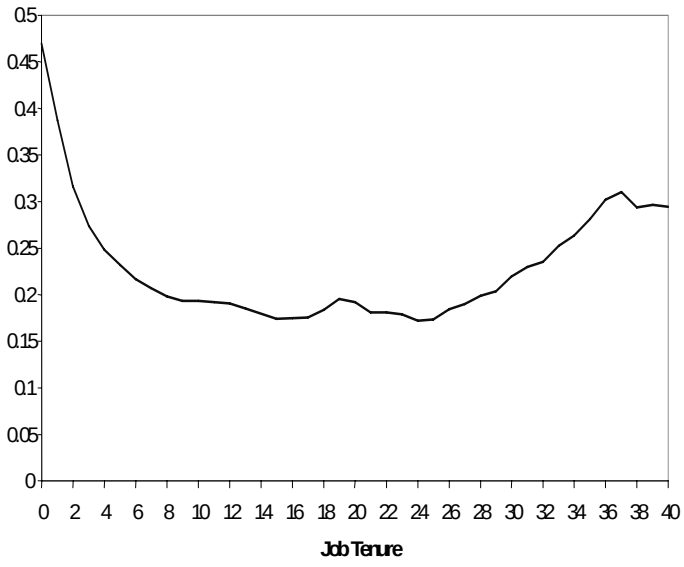
Basic Wage Regressions, 1995-1999

.	(1) Female	(2) Male	(3) Female	(4) Male
Tenure < 1	--	--	-0.0107***	-0.0251***
	--	--	(0.0008)	(0.0008)
Tenure = 1	--	--	-0.0322***	-0.0473***
	--	--	(0.0010)	(0.0009)
Tenure	0.0171***	0.0129***	0.0144***	0.0092***
	(0.0001)	(0.0001)	(0.0002)	(0.0001)
$Tenure^2 \times 10^{-2}$	-0.0310***	-0.0216***	-0.0236***	-0.0121***
	(0.0004)	(0.0003)	(0.0005)	(0.0004)
Age	0.0178***	0.0301***	0.0181***	0.0303***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
$Age^2 \times 10^{-2}$	-0.0183***	-0.0314***	-0.0186***	-0.0316***
	(0.0003)	(0.0002)	(0.0003)	(0.0002)
Secondary education	0.2224***	0.1879***	0.2224***	0.1876***
	(0.0008)	(0.0008)	(0.0008)	(0.0008)
Tertiary education	0.5477***	0.5199***	0.5472***	0.5187***
	(0.0022)	(0.0021)	(0.0022)	(0.0021)
Part-time	0.1463***	0.1729***	0.1474***	0.1760***
	(0.0015)	(0.0017)	(0.0015)	(0.0017)
Hierarchical Levels				
Level 1	-0.1429***	-0.1888***	-0.1407***	-0.1867***
	(0.0009)	(0.0010)	(0.0009)	(0.0010)
Level 2	-0.2022***	-0.2407***	-0.2013***	-0.2392***
	(0.0009)	(0.0009)	(0.0009)	(0.0009)
Level 3	-0.1184***	-0.1185***	-0.1179***	-0.1185***
	(0.0007)	(0.0007)	(0.0007)	(0.0007)
Level 5	0.3072***	0.2382***	0.3072***	0.2381***
	(0.0017)	(0.0014)	(0.0017)	(0.0014)
Level 6	0.2714***	0.2428***	0.2717***	0.2424***
	(0.0026)	(0.0014)	(0.0026)	(0.0014)
Level 7	0.4806***	0.4793***	0.4803***	0.4785***
	(0.0028)	(0.0019)	(0.0027)	(0.0019)
Level 8	0.5913***	0.6405***	0.5909***	0.6394***
	(0.0032)	(0.0024)	(0.0032)	(0.0024)
Log number of workers	0.0577***	0.0658***	0.0580***	0.0663***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Constant	0.2605***	0.1968***	0.2717***	0.2175***
	(0.0035)	(0.0033)	(0.0035)	(0.0033)
Observations	2,498,050	3,589,965	2,498,050	3,589,965
R-squared	0.62	0.62	0.62	0.62
F (df)	F(33;1089573) =45808.79	F(33;1519621) =78750.03	F(35;1089573) =43274.53	F(35;1519621) =74444.52

Dependent variable is the log real wage (Constant 1999 Euros). *Tenure* and *Age* are measured in years. Hierarchical levels are defined as binary variables, from apprentices, internships, and trainees (Level 1) and non-skilled professionals (Level 2) to top management (Level 8). All specifications include dummies for region, industry and year. Robust standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Figure1

Empirical Exit Hazard for Job Separation
By Years of Tenure, 1991-1999 (N = 7,989,649)



Referências

Ariga, Kenn, Yasushi Ohkusa, and Giorgio Brunello. 1999. Fast Track: Is it in the Genes? The Promotion Policy of a Large Japanese Firm. *Journal of Economic Behavior and Organization* 38: 385-402.

Baker, George, Michael Gibbs, and Bengt Holmstrom. 1994a. The Internal Economics of the Firm: Evidence from Personnel Data. *Quarterly Journal of Economics* 109: 881- 919.

_____. 1994b. The Wage Policy of a Firm. *Quarterly Journal of Economics* 109: 921-955.

Bauer, Thomas K. and John P. Haisken-DeNew. 2001. Employer Learning and the Return to Schooling. *Labour Economics* 8: 161-180.

Bernhardt, Annete, Marina Morris, Mark S. Handcock, and Marc A. Scott. 1999. Trends in Job Instability and Wages for Young Adult Men. *Journal of Labor Economics* 17: 65-90.

Bernhardt, D. 1995. Strategic promotion and compensation. *Review of Economic Studies* 62: 315-339.

Both, Alison L., Marco Francesconi, and Jeff Frank. 2003. A Sticky Floors Model of Promotion, Pay, and Gender. *European Economic Review* 47: 295-322.

Cabral, Luís M. B. and José Mata. 2003. On the Evolution of the Firm Size Distribution: Facts and Theory. *American Economic Review* 93: 1075-1090.

Farber, Henry S. 1999. Mobility and Stability: The Dynamics of Job Change in Labor Markets. In Ashenfelter, O. and Card, D. (Eds), *Handbook of Labor Economics*, 2439-2483, Elsevier, Amsterdam.

Finnie, Ross and David Gray. 2002. Earnings Dynamics in Canada: An Econometric Analysis. *Labour Economics* 9: 763-800.

Gottschalk, Peter and Robert Moffit. 1999. Changes in Job Instability and Insecurity Using Monthly Survey Data. *Journal of Labor Economics* 17: 199-236

Jaeger, David A. and Ann Huff Steves. 1999. Is the Job Stability in the United States Falling? Reconciling Trends in the Current Population Survey and Panel Study of income Dynamics. *Journal of Labor Economics* 17: 1-28.

Jovanovic, Boyan. 1979a. Job Matching and the Theory of Turnover. *Journal of Political Economy* 87: 972-990.

Jovanovic, Boyan. 1979b. Firm-specific Capital and Turnover. *Journal of Political Economy* 87: 1246-1260.

Lazear, Edward. 1999. Personnel economics: past lessons and future directions. *Journal of Labor Economics* 17: 199-236.

Lima, Francisco and Mário Centeno M. 2003. The Careers of Top Managers and Firm Openness: Internal versus External Labour Markets. *Working Paper* 15-03, Banco de Portugal.

Lima, Francisco and Pedro Telhado Pereira. 2003. Careers and Wages within Large Firms: evidence from a matched employer-employee data set. *International Journal of Manpower* 24: 812-835.

McCue, Kristin. 1996. Promotions and Wage Growth. *Journal of Labor Economics* 14: 175-209.

McLaughlin, Kenneth J. and Mark Bils. 2001. Interindustry Mobility and the Cyclical Upgrading of Labor. *Journal of Labor Economics* 19: 94-135.

Neal, Derek. 1999. The Complexity of Job Mobility among Young Men. *Journal of Labor Economics* 17: 237-161.

Neumark, David, Daniel Polsky, and Daniel Hansen. 1999. Has Job Stability Declined Yet? New evidence for the 1990s. *Journal of Labor Economics* 17: 29-64.

Pergamit, Michael R. and Jonathan R. Veum. 1999. What is a Promotion? *Industrial and Labor Relations Review* 52: 581-601.

Topel, R.H. and Ward, M.P. 1992. Job mobility and the careers of young men", *Quarterly Journal of Economics* 107: 439-479.

Topel, R. 1991. Specific capital, mobility, and wages: wages rise with job seniority", *Journal of Political Economy* 99: 145-176

Winter-Ebmer, Rudolf and Josef Zweimüller. 1997.Unequal Assignment and Unequal Promotion in Job Ladders. *Journal of Labor Economics* 15: 43-71.

Sessão 3 - Avaliação de Programas Públicos
Session 3 - Evaluation of Social Programmes

THE EMPIRICAL EVALUATION OF SOCIAL PROGRAMMES

Mónica Costa Dias[*]
Banco de Portugal

January, 2004

Abstract

This study discusses the different approaches to the evaluation of social policies currently used in the literature. It separates between the structural and the non-structural methodologies. The relative advantages and disadvantages of each approach are discussed, including data requirements, the meaning and importance of their output as well as its validity. It is argued that much is gained by using the two approaches together. A recent Portuguese educational policy is used to illustrate the discussion. Several solutions to the evaluation problem are discussed in the context of this policy, drawing attention to the trade-off between data and policy design.

[*] Acknowledgements: The author is sincerely grateful to Isabel Lopes da Silva and Teresa Otero from the Ministry for Education for invaluable help and support in the collection and interpretation of information on the educational policy in Portugal. Sincere thanks are also due to Diogo Lucena for discussing this paper and to Isabel Horta Correia, José Ferreira Machado and Pedro Portugal for helpful discussions. The usual disclaimer applies.

1. Introduction

A recent feature of policy interventions is the call for independent evaluation. This is mainly for two reasons. First, it is now recognised that different groups in the society extract different payoffs either directly or indirectly from a policy intervention. Some policies may produce quite significant payoffs, making their size and distribution an important input in the designing process of new policies. And Second, such interventions often require considerable financial and human resources to be implemented and operated.

The evaluation process, however, has not proved an easy task. The main problem can be regarded as one of *missing data*. Individuals are either affected or not-affected by the policy intervention at each moment in time, never both simultaneously. At the economy at large, either the policy has been implemented or not, but again the two cases are not simultaneously observed. Self-selection along with heterogeneous effects, different markets with specific economic conditions, economy-wide effects through price's adjustments and scarcity of good detailed data all contribute to the problem of finding the adequate counterfactual in evaluation studies.

Social experiments could, in principle, solve part of the problem with randomisation, but they are rarely available especially in Europe where some of the largest interventions take place. The advantages and disadvantages of experimental data have been widely discussed.[1] For all drawbacks that can be named, good quality experimental data can provide more reliable results than non-experimental one and be very useful to assess the adequacy of different empirical methods in evaluation.

This study discusses the two main approaches to the evaluation of social programmes, the structural and the non-structural approaches. To do so in an integrated form, the evaluation problem is formalised in a way that is both sufficiently general to embrace a wide range of policies and evaluation exercises, and sufficiently specific to explicitly reveal the problems and strengths of each approach. The discussion draws attention to the nature and adequability of the assumptions underlying each methodology, the data requirements and the quality and usefulness of the identifiable information. A specific policy recently released in Portugal is used to illustrate the arguments, the "*Programa de Desenvolvimento e Expansão da Educação Pré-Escolar*".

The policy effects can be assessed on two main grounds: the individual and the aggregate levels. At the individual level, the performance of the policy in changing individuals' outcomes has received most of the attention. At the aggregate level, research has focused on the measurement of the relative size of social costs and benefits from interventions. In this study I discuss how the structural and the non-structural approaches handle the different policy questions at the individual and the aggregate levels. It will be argued that there are gains from taking the two approaches simultaneously as they complement each other in terms of the analysis and inform the policy-maker of different aspects of the impact of interventions. When applied to a well-posed problem, the non-structural or *ex-post* approach provides a simple setting to answer a precisely defined question. The question is usually formulated as "What is the short-run, direct impact of the policy intervention on the outcome x for a specific group?", which is convincingly answered by a number or set of numbers extracted from the actual observation of the post-

[1] See Bassi (83 and 84), Hausman and Wise (85) and, more recently, Heckman, Lalonde and Smith, 98.

policy-reform behaviour of the agents in that group. On the contrary, the structural or *ex-ante* approach uses a considerable amount of structure to provide an overall framework that guides the economic reasoning about the potential effects of social interventions. It answers the broader question of explaining the economic behaviour of individuals causing the measured effects. Simultaneously, it justifies the choice of the parameters of interest in the *ex-post* analysis by establishing its economic meaning and sheds light on the potential drawbacks of that analysis.

In what follows, I use the jargon commonly used in the evaluation literature applied to welfare and active labour market programmes, perhaps its most dynamic field presently. Though this terminology is quite intuitive in that particular field, it is not so much so when applied to other policy areas. Thus, I briefly present it here in a general way. A *programme* is used to denote the new policy regime after an intervention and *treatment* refers to what in the new policy may directly affect some agents. In this context, a *direct effect* of a policy intervention on a certain variable of interest is one that arises when prices and/or other aggregate economic conditions are kept unaltered apart from the policy specific features. By opposition, an *indirect effect* is one that arises due to the adjustments in prices and/or aggregate economic conditions in response to the new behaviour patterns of agents directly affected by the policy. Thus, treatment can be defined by a new tax rate, a new subsidy or a new educational programme, to name only a few. Notice that, given a particular policy, treatment can generally be defined in more than one way, possibly linked to the goal of the evaluation exercise. For example, suppose a new subsidy is introduced, targeted at agents of a given educational level that enrol in school to acquire more education. The treatment can be defined as the subsidy itself when one is willing to study its impact on the odds of enrolling in school for a particular group. Alternatively, one may be willing to define treatment as the additional schooling experience if the aim is to study its impact on the human capital accumulation or labour market performance among agents that receive such treatment. A *participant* or a *treated* is an agent that receives treatment, thus potentially being directly affected by the policy. Clearly, the meaning of participant depends on the adopted definition for treatment. In the above example, participants can be understood as the agents eligible for the subsidised educational period or the agents that have actually experienced a period of subsidised education, respectively. A *non-participant* or a *non-treated* is, quite obviously, an agent that is not a participant. An (*non-*)*eligible* agent is one that has been defined by the policy design to be (un)able to participate in the programme. Finally, the *outcome* is defined as the variable over which the researcher wants to measure the impact of the treatment.

The rest of the paper is organised as follows. It starts by discussing the evaluation problem in more detail in section 2. This is done by developing a formal setup, where the evaluation problem and parameters can be rigorously defined. Section 3 discusses the two approaches to the evaluation problem. I do not attempt to discuss the specific methods that can be used. This has been recently done elsewhere (see Heckman, LaLonde and Smith, 98, Angrist and Krueger, 99 and Blundell and Costa Dias, 00 and 01). The goal here is to compare the information that can be extracted from the structural and the non-structural approaches and to discuss its usefulness to the policy maker. Finally, section 4 discusses the possible evaluation of the "*Programa de Desenvolvimento e Expansão da Educação*

Pré-Escolar" using the two approaches presented before.[2] I discuss how the design of the policy and the available data determine the evaluation process and the chances of extracting useful information from it. Finally, section 5 contains some concluding remarks.

2. The Evaluation Problem

The evaluation problem is always one of missing data. One cannot observe an agent that has simultaneously been and not been treated or an economy that has both undergone and not undergone a policy change. The ultimate goal of the evaluation exercise is to understand the outcomes from treatment, how they arise and how they depend on the characteristics of the agents, groups or economy. A major problem, of course, is that such outcome is not observable. Possibly even more serious for the purpose of optimal design, only a very limited number of policies can ever be implemented or experimented. Thus, absolutely no data exist on the much larger range of possible but not implemented policies.

The missing counterfactual problem is particularly problematic when selection occurs.[3] Selection means that agents, or someone in their name, select whether to participate in the programme and become treated. The selection process itself implies that treated and non-treated agents are different as they have revealed to be by deciding differently about participation. Thus, the evaluation exercise aims at separating empirically the differences in behaviour of treated and non-treated that are due to the composition of the groups from those caused by the treatment. The latter is called the *causal effect from treatment*.

In addition to selection, there is strong empirical evidence on the importance of heterogeneity on the effects of policy interventions (see the recent discussion by Heckman, 01). The existence of heterogeneity rises new difficulties to the evaluation problem. To see why, consider first the homogeneous case. Under homogeneous treatment effects, the gains from treatment are the same for all agents by definition. This means that such gains are not the relevant information for the agent deciding about participation. Instead, he/she may use other factors to decide. This does not mean that the knowledge about who participates is not important: it is to the extent that it matters who collects the returns from the intervention.

However, under heterogeneous treatment effects there are numerous parameters of interest to be defined. These are inexorably linked to the data being used in an empirical study. In particular, the coverage and richness of the data are key issues for the definition of what is identified. The former defines the subpopulation for which the estimates are valid. The latter determines the definition of the treatment to be used in the analysis by establishing the treatment and control groups. The consequent diversity of possible parameters complicates the interpretation process considerably when results from different studies are being compared, all based on different data, policies and underlying assumptions. A major advantage of structural models is, precisely, that the estimated effects are clearly interpretable as the underlying economic assumptions are openly discussed.

---

2 There are few examples in the literature of policies that have been simultaneously studied under the two approaches. Two such cases are the tax credit scheme (Blundell, Costa Dias and Meghir, 03, Blundell, Costa Dias, Meghir and van Reenen, 04 and Lise, Seitz and Smith, 03) and the educational subsidy (Attanasio, Meghir and Santiago, 03, Behrman, Sengupta and Todd, 02 and Todd and Wolpin, 03).

3 This excludes the case where no counterfactual exists due to all agents being treated.

Simultaneously, heterogeneous impacts also make the interpretation of the parameters more important to the policy-maker. This is because different agents are differently affected by the policy and value interventions differently. Thus, the information on who gains and loses from the intervention along with the distribution of gains becomes politically relevant.

Other issues, like the dimension of the programme, may still complicate the evaluation procedure. Large programmes with wide implementation may affect a considerable number of agents if effective treatment is provided. In such case, not only the treated but also the non-treated are expected to be affected by the existence of the programme. The effect spreads to the economy at large given its impact on prices and market conditions. By affecting both participants and non-participants, indirect effects contaminate the counterfactual information. The plague of the evaluation problem under such circumstances is that non-participants that closely resemble participants are more likely to be affected by the indirect effects of the programme, while non-participants of different characteristics are usually not in the same market. However, and although indirect effects are expected in the large majority of public interventions, they are rarely considered in the empirical evaluation literature. The main problem is that their existence generally invalidates the use of non-structural methods.

In what follows, I formalise the problem in a rigorous way and then define several possible evaluation parameters of interest, differently suited to answer to different policy questions.

2.1. Formalising the Problem

Let t represent calendar time and consider a policy change or intervention to occur at time $t=k$. At this stage, such intervention is deliberately left unspecified as the model is general enough to apply to a variety of cases.[4] To keep generality, I formalise public policy to be characterised by a number of parameters, θ. These may include information related to tax rates, subsidies, active labour market programmes, publicly managed insurance mechanisms, among others. Before time k the policy parameters are $\theta=\theta_0$. The public intervention to be considered here consists of altering such parameters to $\theta=\theta_1$ at time k and keeping them at such level for $t>k$. A combination of policies can be considered within the present context though I will not discuss such possibility here. It is important to notice, however, that the impact of a policy may depend on what other policy parameters are since these affect the aggregate economic conditions at the start and may determine other options available to the agents.

I denote by y the outcome of interest in the evaluation of the policy under consideration. y_{it} is the particular realisation of y for agent i at time t. Moreover, let (X_{it}, Z_{it}, W_t) be the set of observable endogenous and exogenous variables that characterise the state space faced by agent i at time t. X and Z are individual-level variables that determine y and the participation decision, respectively. They may include common variables or not. W contains aggregate information other than the policy parameters. It

[4] Just to name a few, potential examples include welfare programmes aimed at improving the labour market performance of the least skilled or the incentives to work of lone parents; educational policies aimed at ensuring equal opportunities to all at the start of life; health policies determining how health centres/hospitals are financed/managed; fiscal policies altering the type, level or progressivity of the tax system.

includes prices, technology and market conditions. Though W is allowed to change with time, it is assumed to be unaffected by the policy intervention.

Suppose one wishes to measure the impact of the policy on the outcome y at time t for an agent i deciding about participation at time τ. Let's assume that the outcome of interest follows the rule,

$$y_{it} = f\left(X_{it}, d_{i\tau}, W_t, \theta, \delta_{it}\right) \tag{1}$$

which allows for heterogeneity in the policy's effects both across agents and economic environments. $d_{i\tau}$ represents the enrolment decision of agent i at time τ and the error term, δ, represents the unpredictable part of the outcome.[5]

The policy intervention may impact directly on y (the effect of a tax change on income) or indirectly through its effect on x (the impact of training on future wages) and W (the impact training on the composition of labour market groups and, through that, on the prices for labour). A typical rule for the evolution of x could be as follows,

$$X_{i,t+1} = \begin{cases} h_X^1\left(X_{it}, d_{i\tau}, \nu_{it} \middle| W_t, \theta_1, t-\tau\right) & \text{for} \quad t > k \\ h_X^0\left(X_{it}, \nu_{it} \middle| W_t, \theta_0\right) & \text{for} \quad t < k \end{cases} \tag{2}$$

where h_X^0 and h_X^1 represent the laws of motion of X when the policy intervention has and has not occurred, respectively. τ is the calendar time of the treatment for agent i and is only relevant for treated agents. k is the time of the policy intervention. Both W and X may include lagged information about their own if that is required to establish the law of motion for X. Finally, ν stands for the unpredictable innovation on the evolution of the state variables.

Similarly, the law of motion for W can be represented as

$$W_{t+1} = \begin{cases} h_W^1\left(W_t, \eta_t \middle| \theta_1, t-k\right) & \text{for} \quad t > k \\ h_W^0\left(W_t, \eta_t \middle| \theta_0\right) & \text{for} \quad t < k \end{cases} \tag{3}$$

where h_W^0 and h_W^1 represent the laws of motion of W when the policy intervention has and has not occurred, respectively.

I choose to formalise the individual problem as that of deciding about participation at time τ given the agent's expectations about the impact of treatment. Thus,

5 Some form of autocorrelation is allowed for but left unspecified.

$$d_{i\tau} = \begin{cases} 1 & \text{if } g\left(Z_{i\tau}, \varepsilon_{i\tau} \middle| W_\tau, \theta_1\right) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where $d_{i\tau}$ denotes the agent participation decision and ε represents its non-deterministic component.[6] Depending on the nature of the programme, participation may either be endogenously influenced by the agent or not. The latter case corresponds to only variables defining eligibility being included in Z. In the more frequent case where the agent may exert some control over becoming eligible/participant, individual characteristics explaining such choice should also be included in Z.

2.2. Types of Policy Effects

There are a number of criteria that define the different policy parameters. One of the main factors determining the definition of the parameters is *the timing of the information*. We can use *ex-ante* or *ex-post* data to estimate the impact of treatment. Let's start from the latter. *Ex-post* information is used to observe the actual impact of the policy. It is conditional on the particular state of the world that happened to exist at the time of the intervention. In this context, the individual-specific impact of treatment at time t as measured *ex-post* is,

$$\alpha_{it} = y_{it}^1 - y_{it}^0 \tag{5}$$

where the superscripts ^{1} and ^{0} denote the treatment and non-treatment status. Thus, y_{it}^1 is the outcome for agent i at time t if he/she had been treated and y_{it}^0 is what the outcome would have been in the non-treatment case.

On its turn, *ex-ante* data are used to predict the impact of treatment. The ex-ante counterpart of α_{it} as measured from time $\iota < k$ is,

$$\gamma_{it} = E_\iota\left[y_{it}^1 - y_{it}^0 \middle| X_{it}, W_\iota, \delta_{it}, \theta_0, \theta_1 \right] \tag{6}$$

This parameter is most important for the policy maker deciding about whether or not to implement a programme. *Ex-ante* data is typically used together with a set of economic assumptions to predict how the policy will work. Thus, its use is usually associated with structural analysis and the parameters identified are directly linked to the economic behaviour of the agents. Contrary to *ex-post* data, the *ex-ante* data is not conditional on the specific state of the world observed at the programme introduction or on the selective behaviour of the treated agents. Thus, these data can be used to inform about

---

6 For example, in the random experiment case, ε represents whether the eligible agent has been randomly drawn or not.

the optimal individual decisions and the source of the potential outcomes, that is, the structure of the individual problem.

However, the meaning of α and γ is not entirely clear yet. One of the main factors determining its meaning is the concept of treatment on wishes to assess. Natural definitions for treatment include, among others, the existence of the programme, being eligible to participate, or actually participate. Take the example of a training programme targeted at the long-term unemployed (this is the example studied in Blundell, Costa Dias, Meghir and Van Reenen, 04). The training itself may impact on the outcomes of trainees by changing their human capital. This determines the effect of training conditional on enrolling into it. However, on the top of this effect the sole possibility of enrolling into training may lead agents to change their employment decisions before reaching the long-term unemployment depending on their preferences for training. Moreover, it may also change the employment decisions of agents that could never become eligible if they expect the market conditions to change as a consequence of the programme.

A second important determinant of the meaning of the treatment effect is the definition of the counterfactual. For instance, the counterfactual may represent what the agent would have got in the non-treatment status in the treated world or in the non-treated world. The former is the important information for an agent deciding about whether to participate and taking the existence of the programme as given. It corresponds to $y_{it}^0 = f\left(X_{it}^{10}, W_t^1, \theta_t, \delta_{it}\right)$ where X^{10} represents the observable individual characteristics in the treated world had he/she not been treated, and W^1 is the economy information in the treated world. The latter is the important information for the policy maker deciding about the introduction of a programme the programme as given. It corresponds to $y_{it}^0 = f\left(X_{it}^{00}, W_t^0, \theta_0, \delta_{it}\right)$ where X^{00} and W^0 are the individual- and economy-level variables in the non-treated world.

The above discussion is only important if indirect effects from treatment exist. These occur if the sole existence of the programme impacts on both participants and non-participants. Take, for instance, the example of an educational policy. To the extent that it affects a large number of children, it will change the composition of the labour market in the future and affect the prices for labour. Independently of whether the educational decisions of a particular child have changed as a consequence of the policy, he/she will be affected by its existence. However, the indirect effect from the policy (arising through changes in prices) may affect participants and non-participants differently. In such case, taking $y_{it}^0 = f\left(X_{it}^{10}, W_t^1, \theta_1, \delta_{it}\right)$ as counterfactual does not allow to separate the impact on the participants from that on the non-participants.

Under heterogeneity, the composition of the group being studied is extremely important. It is not the same to study the impact of treatment on agents that opt in or out of the programme. Similarly, it is not the same to study the impact of treatment on the treated or some subgroup of the treated. For the policy maker deciding about the extension of a policy to some non-treated population, it is often much more informative to know about the impact of the programme at some margin of the treated that is close to such population than to know about the average impact of treatment on the treated.

3. Different Approaches to the Evaluation Problem

The approaches to evaluation depend crucially on the data available. The two major categories are the *ex-ante* and the *ex-post* approaches. As discussed below, they deal

differently with the problem of estimating the evaluation parameters of interest, providing information on different, but potentially complementary, policy questions.

Let's start by the *ex-ante* or structural approach. It aims at embedding the policy in an economic model in order to define and estimate economically meaningful parameters and to fully understand the nature of the effects. It focuses on broad questions regarding the impact of the policy on the different groups, the relative advantages/disadvantages of that specific policy when compared to other policies and the optimality of the policy. The main problem faced by this approach is that of defining the correct structure for the economic problem at hand. Its goal is to find the correct specifications for the structural relationships, f, h_W, h_X and g based on the before policy intervention data. After this, all the information about the policy can be revealed using simulation methods, including the nature and size of its impacts, its distribution and the economic mechanisms at their origin. Moreover, different types of effects can be computed and comparisons with alternative policies are possible. These are the main advantages of this approach.

The structural approach is, perhaps, the most natural in economics. A representative example is Auerbach and Kotlikoff (83 and 87) and their use of the overlapping generations model to study the impact of social security (see also Krueger and Kubler, 03). It has also been frequently used in other areas of welfare economics (see Hanusheck, Leung and Yilmaz, 01, Bovenberg and Jacobs, 01, for recent examples on the impact of tuition subsidies and wage subsidies). However, this literature is rarely based on the empirical micro evidence, casting doubts about the appropriateness of its results (see Browning, Hansen and Heckman, 99, for a discussion; an exception is Heckman, Lochner and Taber, 98). This suggests the main drawback from such approach: its dependence on the structural assumptions at the basis of the whole analysis. This is not to say that the non-structural approach does not depend on a set of assumptions, which it does, but they are stronger and more determinant here.

Despite the relative richness of the structural methods, most evaluation literature uses a non-structural, also known as *ex-post*, approach. It focuses on very specific, narrowly defined questions about *the* particular policy that has been implemented, trying to measure its impact on some group. It uses post-policy data and formulates the missing data problem as one of constructing the correct counterfactual for the treated. Data, along with a number of assumptions, allow the policy effects to be estimated without knowledge of the structural relationships, f, h_W, h_X and g. Thus, any conclusions to be drawn from these methods cannot be generalised to other groups, other economies or other policies because the nature of the decision process, at the origin of the observed impact, is not understood. It is conceivable that other agents or other economic environments would produce different reactions and, therefore, different policy effects.

The main challenge of non-structural methods is to construct the correct counterfactual without clearly modelling what motivates individual behaviour. Thus, data is the most important ingredient here, providing the required information on both participants and non-participants after (and preferably also before) the policy intervention to assess its effect, usually on the treated.

There are two pieces of data always missing for these types of studies: those describing what the agents and the economy would have been had no policy been changed. The latter is usually assumed to be well represented by the aggregate characteristics observed under the policy case. That is, the existence of indirect effects from the

programme is ruled out.[7] Out-ruling indirect effects ensures that non-treated agents are not affected by the existence of the programme, establishing them as a potentially suitable group to draw from in the process of constructing the counterfactual. Still, this is generally not enough to identify the correct counterfactual. An additional assumption is required to establish that a non-treated actually reproduces what the treated would have been in the event of not being treated. Different methods postulate different identification conditions and may identify different parameters as has been extensively discussed in the literature (Heckman and Robb, 85 and 86, Heckman, LaLonde and Smith, 99, Angrist and Krueger, 99, Blundell et al., 00 and 01).

These two approaches are mutually informative about the adequacy of each other. On the one hand, the definition of the evaluation parameters is established by the structural approach as it lays the economic foundations of the problem. On the other hand, the adequacy of such assumptions to the data is ultimately assessed by the results of the non-structural exercise and their comparison with simulated policy effects produced by the structural analysis.

4. An Example in Portugal: Subsidising the Pre-School Education

In 1997 a new educational policy was launched in Portugal, called "*Programa de Desenvolvimento e Expansão da Educação Pré-Escolar*", from now on designated by PDEPE. The PDEPE is an educational programme targeted at children in their pre-school years, aged between 3 and 5. Its stated aim was to create the conditions for all children of this age to have access to pre-school education, at a first stage giving priority to the older children and progressively extending it to the younger. The programme was not compulsory, meaning that it was not intended to force all children to attend pre-school. On the contrary, the goal was to give the parents the option to have their children in pre-school if they wish to do so.

Two main arguments were used to justify this intervention. First, the perception that pre-school education has a strong positive impact on the child's cognitive, social and emotional development that affects, in particular, her/his future academic performance. And second, the apparent lack of supply of pre-school education, unevenly distributed throughout the country and especially affecting the more densely populated areas.

In this context, two main areas of action were defined. First, subsidies to the creation of new or the re-structuring of existing pre-school premises were implemented. Subsidies were made available to all types of institutions though its size depended on the legal status of the school. More concretely, the subsidies were established to cover 25-100 per cent, 25-75 per cent or 15-25 per cent of the total cost when requested by public pre-schools managed by the local council, charities and other non-profit-oriented institutions, or private profit-oriented organisations, respectively. Additional criteria were used to determine the actual proportion of the cost being subsidised, the most important being the local available capacity, thus introducing differences across regions. To this end, regions were classified as "very deprived", "deprived" and "not deprived" depending on whether the available capacity corresponds to 0-25 per cent, 25-50 per cent or 50-90 per cent of the

[7] This is because finding an adequate counterfactual for an economy in the data is unlikely to succeed: the closest (geographically or in definition) markets are likely to be more similar to the treated one, but also more vulnerable to indirect effects when these occur.

target population (children aged 3 to 5). Other criteria privileged heavily populated urban areas and areas characterised by poor academic performance.

Second, a new educational subsidy (partially) covering the tuition costs of pre-school education was also implemented. This subsidy replaced an old existing one, changing its rules and generosity. Two components of the pre-school services were identified, corresponding to the educational and the caring roles. The subsidy was mainly directed at the educational costs though further support was offered to families with low income, contributing towards the caring costs. In public, charitable and other non-profit-oriented schools, the subsidy covers the full cost of education and is handed to the school. Private schools, on the other hand, were eligible for a tuition subsidy directed at families with low income. Such subsidy is first set to amount to up to 52 per cent of the tuition costs of education though the actual subsidised proportion of the costs depends on the total tuition fees of the school. Its generosity is revised yearly. The subsidy is deducted by the school on the tuition fees of subsidised children and is progressively phased out as the income rises.[8]

4.1. Potential Effects from the Policy

This type of policies can affect a number of dimensions in the economy, as well as a wide range of agents.[9] Contemporaneously, it may impact on the parents' availability for working by reducing the value of staying at home. This may induce some parents into the labour market and others to work more hours, increasing labour supply. Such effect should be stronger among low/medium-skilled workers as the fixed and variable costs of working are expected to be relatively more important for them. If the number of agents entering the labour market is large enough, indirect effects may also arise as the supply of labour increases. Under the assumption that prices change freely, a drop in the wage rate of unskilled labour is expected. However, labour market institutions like the minimum wage may prevent this to happen when the price restriction is binding. If this is the case, then some substitution may occur, with some among the otherwise inactive parents taking the jobs that other agents willing to work would have got in the no-policy regime.

On the other hand, the academic and labour market performance of treated children may be affected by the intervention in the medium- to long-run. There is large evidence in the literature that early childhood interventions do improve academic performance as measured by the rate of grade retention. Heckman and Forum (01) describe results from several experiments ran in the USA. Performance in school as measured both by the rate of retention and by test scores is strongly positively affected by early childhood interventions. A number of studies that follow children into the labour market also report higher earnings for the treated ones, overcoming the cost of investment by far. The returns when measured in terms of earnings seem to be very large, at least among the most disadvantaged children and when high quality treatment is available together with family support in the form of home visits by childhood specialists. Another remarkable result

---

8 Caring costs, if any, are partially subsidised in public and non-profit-oriented schools for families with low levels of income only. Again, the subsidy is phased out progressively as the income rises.

9 The identification and discussion of the potential policy effects is only possible if having a structural model of human capital and labour market participation in mind. That is, the use of a structural approach is needed in first place to define the parameters of interest in a non-structural analysis.

concerns to the incidence of criminal activities. Crime seems to be greatly reduced by early childhood interventions targeted at the most disadvantaged, high risk children. The estimated effect is so large to more than compensate for the investment only in terms of savings on incarceration expenses (see Donohue and Siegelman, 98).

As before, all these long-term effects, if strong and widespread enough, are expected to have other impacts in the economy through indirect effects. In terms of the labour market, these originate from the change in the supply of labour and skills, affecting the wage rates and the skill-premium.[10] In terms of the economy at large, more integrated and participative agents reduce the need for expensive public services like the ones related with security and social support, alleviating the public budget and creating the conditions for tax reductions.

The PDEPE is, to a certain extent, different from the interventions studied in the USA. One of the main differences relates to its extension. Contrary to programmes in the USA, that are typically focused on narrow, homogeneous groups, the PDEPE is a programme of global implementation. Theoretically, any child is eligible for most of its subsidies when opting for public schools.[11] But then, a large deadweight loss is expected as many children that would have enrolled in pre-school education in the absence of the programme become subsidised students when the programme is implemented. If the programme does not affect the economic environment, the advantage that these children have from participation is to receive an extra subsidy, not to receive an extra educational treatment that they would have got anyway. Thus, even if the effects of the pre-school experience are homogeneous across agents, the effects of the subsidy will be heterogeneous, depending on the agents optimal choice in the non-policy world. This should be taken into account when trying to produce a global evaluation of the programme like a cost-benefit analysis: deadweight loss is often responsible for the large dimension and heavy financial requirements of social interventions, simultaneously limiting their potential beneficial impact.

On the other hand, this programme is not of compulsory participation. Parents decide on their children's participation depending on the valuation of their own time and expected benefits for children. Families of low income are usually the ones that also value labour market participation by less. Thus, everything else constant, these families are the ones less likely to have their children participating in the programme.[12] The programme, however, introduces an extra subsidy for children from deprived families, contributing towards the food and caring expenses.[13] Such subsidy, however, is only available to very low income families and is relatively unimportant when compared with the educational subsidy. Thus, a selection behaviour may contribute towards keeping a disproportionately large share of deprived children out of the programme. Such process can be strengthened or

10 This is abstracting from technological changes that may be induced by the relative abundance of skills.

11 Any divergence to what is empirically observed results from the fact that the system does not guarantee a place for every children that requires one, at least when the programme is implemented. The selection criteria during the first years of the programme is defined in terms of age: older children are given priority.

12 This is not contradictory with the previous observation that low/medium-skilled parents may react more to the programme given the relative importance of fixed costs from working for them. However, this programme is necessarily characterised by a large deadweight loss as every children in the public/non-profit sector is eligible. Thus, although the incremental impact is expected to be stronger among the children of low/medium skilled agents, the total participation is probably lower among families of low income.

13 This is on top of already existent services targeted on deprived children and ministered by charities.

weakened if present income is, respectively, positively or negatively related with the families' valuation of the children's expected returns from participation.

Selection is all more important if the impact of education is heterogeneous. In such case, children of different characteristics and different background benefit differently from treatment. If parents value their children's benefits then, taking everything else constant, participants are those with higher expected returns. But even if all parents are altruistic, some family characteristics may be correlated with how important children's benefits are in the family utility. Families that value children's returns less are less likely to both decide optimally with respect to participation from the children's point of view and provide good quality environments at home. Admitting that pre-school treatment has a positive payoff, selected out children benefit more or less from treatment depending on whether pre-school services are substitutes or complementary to the home treatment, respectively.

An adequate evaluation of a programme should, therefore, take selection into account: the impact of the intervention depends on who decides to become treated, which is likely to be determined by the expected returns from treatment. Self-selected treated and non-treated are, by definition, different. If, in particular, the aim of the programme is to widespread treatment, guaranteeing access to all agents, its success depends on how able the programme is to provide the right incentives to different agents.

4.2. How Can the PDEPE Be Evaluated EX-POST?

In what follows, I discuss the simplest possible evaluation of the impact of the PDEPE by using the *rate of pre-school take-up* as the outcome of interest. It disregards as unimportant considerations about "who takes up education"' that would have been important to outcomes such as the individual returns from education or the labour market opportunities of parents. Moreover, indirect effects are ruled-out as their existence invalidates this type of approach.

The design of the PDEPE makes it particularly difficult to evaluate *ex-post*. A major problem is that it includes a number of subsidies targeting different components of the market. The two most important examples, and the ones that will be discussed here, are the subsidies to the creation/restructuring of pre-school buildings and to the cost of education. The former creates incentives to increase the supply of pre-school places. The latter subsidises the demand for pre-school services. Both subsidies are available to the private and public/non-profit-oriented sectors but they are much more generous in the second case.

It is difficult to disentangle the effects caused by the different parts of the policy. However, such difficulties can be overcomed with an appropriate source of exogenous variation and adequate data. Apart from the time dimension, I identify two possible sources of identification in the present policy.[14] First, the regional variation. And second, the age cut-off. In what follows it will be argued that the former cannot be used appropriately given the design of the programme. The second one, however, could provide the required identification conditions but is rather demanding in what concerns to data.

---

[14] As in most evaluation exercises, the time dimension is only helpful if used together with a second source of randomisation. This is because, apart from the policy introduction, there are other events occurring over time that may affect the outcome of interest.

Let's start by exploring the regional features of this policy. Although the programme is implemented national-wide, there is some regional variability related with the local available capacity before the policy intervention. This only applies to the subsidy to the creation and restructuring of educational premises, not to the subsidy to the cost of education. Deprived areas are heavily subsidised to develop new premises, the subsidy covering up to 100 per cent of the total cost. The subsidy to non-deprived areas is much less generous, amounting to around 25 per cent of the total cost. Thus, different areas are expected to be differently affected by each of the newly available subsidies. On the one hand, the impact of the subsidy to the creation/restructuring of new premises should dominate in deprived areas not only because of its generosity but also as a consequence of the likely excess demand occurring in these areas. On the contrary, the impact of the subsidy on the costs of education should dominate in non-deprived areas by changing the parents' preferences for this type of education.

So, can one compare regions of different characteristics to assess the impact of the policy? This comparison is unlikely to be valid. If it is performed in one moment in time after the policy is released it should be meaningless given that the realities in the two sets of regions are a priori different by construction. If one compares the evolution from before to after the policy is introduced in the two sets of regions such a priori differences may be controlled for if they would have been maintained in the no-policy world. But this is unlikely to hold given the strong differences between the two sets of regions. Changes in the number of pre-school children in deprived areas are likely to be driven by the supply while changes in demand should be more important in non-deprived areas. To the extent that supply and demand respond to different factors, there are reasons to believe that the differences between the two types of region would have been kept unchanged in the absence of treatment.

The solution would be to find two subsets of regions at opposite sides of the "deprived" frontier but sufficiently close to be comparable. However, the very small number of regions limits considerably the practicality of this solution. Nevertheless, if successfully implemented, it would identify the differential impact of a more generous subsidy to the construction/restructuring of premises at the defined margin. Such parameter is possibly not what the policy maker is looking for as it measures the effect of introducing a more generous against a less generous benefit. How generous it ends up being depends to a large extent on the local agents' initiative to expand the net of pre-schools. Moreover, no information is revealed about the impact of the subsidy to the costs of education.

On its turn, the age cut-off rule establishes that the policy is only available for children between 3 and 5 years old. Neither younger nor older children are eligible. Such eligibility rule can be used as a source of identification if one admits that the school take-up rate or its evolution for ineligible children of close age groups reproduces that of eligible groups had they not been treated. The slightly older children, however, are not a suitable comparison has school enrolment becomes compulsory at 6 years of age. On the contrary, the slightly younger children, of 2 years of age, could be used to provide the missing counterfactual at that margin of the age variable, for the 3 years old. If one believes that the parents of 2 and 3 years old operate in the same market for education, being equally affected by eventual shocks and trends that affect its conditions, the comparison of the evolution of the educational take-up from before to after the programme release could reveal its impact on the youngest eligible children.

To make such comparison valid one should take the two following points into account. First, comparisons can only be made within region in order to guarantee that the same market conditions apply to both age groups. This is particularly important given the strong variation in the supply of educational places across region (and, possibly, the equally strong variation in the taste for education across regions). Second, detailed information about the family composition and social-economic status is required to ensure an equal composition of families across the two groups. Under heterogeneous preferences for education determining the selection process, such a procedure is fundamental to ensure comparability between the two groups.

Although much can be gained from the use of the age dimension for as long as detailed family-level data is available both before and after the programme is introduced, the impact of the two subsidies cannot be disentangled. This is because comparisons are only valid within region and there is no variation on the subsidy for the creation/restructuring of premises at this level.

4.2.1. The Role of the Policy Design

It is possible to argue against the 2 years old as a possible counterfactual. This is because the pre-school age starts only at the age of 3 and there is a large number of institutions, perhaps the majority, that work only with children of at least 3 years of age. If this is the case, the evaluation process is seriously compromised as there are no further "clean" sources of identification allowed for by the design of the programme.

Though good quality data is always necessary to produce good quality evaluation studies, the design of the policy can diminish considerable the need for very detailed information. One extreme case is that of social experiments. If randomisation is strictly followed at all stages of the selection process, the simple comparison of the average outcome of randomised-in and -out agents is a consistent estimator for the impact of treatment on the experimental group.

But experiments are not the only solution to alleviate the heavy burden typically carried by the data in this type of studies for as long as the design of the policy takes evaluation into account. Similar sources of variation to the ones identified above can establish a good start for a successful evaluation. Take, for example, the priorities established in the PDEPE with respect to age groups. At a first stage, the older children were given preference in the access to pre-school places. However, it is not clear how this was ensured or the actual timing of the priority rules. Instead, the subsidy on the cost of education could have been attributed exclusively to older children in a first stage and then extended to younger children afterwards. This phase-in of the second group could provide the required source of identification if, as discussed before, children at the margin of eligibility could be assumed to operate in the same market, being similarly affected by its fluctuations. To compare two closely aged subgroups of children between 3 and 5 years old is probably more convincing than to compare 2 to 3 years old.

Another possible source of identification consists of piloting the policy first in a random set of regions and extend it afterwards to the whole economy. Together with the age rule, this further instrument would allow the researcher to separate the impact of the two subsidies: the comparison between two sets of similar regions identifies the overall impact of the policy; the comparison between age groups identifies the impact of the

subsidy to the cost of education; thus, the difference identifies the impact of the subsidy to the creation/restructuring of pre-school premises.

4.3. The Structural Evaluation

The structural approach deals easily with complex policy designs. This is because all the analysis is based on the structural behavioural rules of the agents and markets. Any results are obtained using simulation methods. The structural model can be interpreted as a laboratory where all experiments (simulations) can, in theory, be tried out. Such experiments fully control for any factors that may influence the results, thus being able to separate the impact of different components of the programmes and to disentangle each effect on its basic parts (see the decomposition of the impact of a wage subsidy and a tuition subsidy in Blundell et al., 03).

Probably the most complex procedure required by structural methods is the estimation of the structural parameters. This approach does not aim at estimate the impact of treatment directly. When applied to the PDEPE, the estimation stage would probably involve defining the family preferences for education by approximating not only the parameters on the tastes for education but also those defining the tastes for working and consumption of parents and the rules of human capital accumulation. To make things clearer, suppose the problem is that of deciding whether to have children enrolled in pre-school education. The problem of the family is modelled, including the costs and benefits from this type of education: costs with tuition, material, transportation and utility, and benefits from the free time to work (wages), returns to education and utility. This is a specific version of the economic setup presented in section 2.

Given that investment in education cannot be understood in a static framework, human capital models always involve a dynamic feature. But this implies that its estimation should make use of longitudinal data, observing the same agents over time. A particularly rich source of information that is typically used in the structural estimation of human capital models is that provided by cohort studies. Cohort studies include detailed longitudinal information on a particular group of individuals born during a typically short time interval. The sampled agents are usually followed throughout infancy into childhood and adulthood. The information respects to the same individuals at all times as the panel is not a rotating one. All sorts of information are usually collected, including demographic, family environment, socio-economic status of the parents, health, ability, development, school performance and labour market performance.

Using these data, the researcher may acquire a detailed knowledge about the surveyed agents and observe their life-path conditional on the initial conditions. Under some conditions, the family educational options can be characterised in terms of what the expected returns to education are, both for children and parents. Thus, these sort of data is frequently used to provide empirical clues about the structure of the agents' decision process. [15]

---

[15] Simulation methods are a popular option for this type of estimation: start with some values for the structural parameters, solve the family optimisation problem, simulate what each family in the data would have done if following the decision rules just established, compare with the decision actually taken by the family, adjust the parameters to get closer to the observed decision and repeat the procedure until an optimality criteria is fulfilled. This procedure has been applied in Keane and Wolpin,97. See also Keane and Wolpin, 94, for a more detailed discussion of the estimation method.

Once the model is estimated, the structural estimates of the effects of a policy can be obtained from the simulation of the structural model. As discussed before, the introduction of a policy corresponds to changes in some environmental variables. A large range of policies can be tried out, experimenting, for instance, with different levels of subsidy, types of subsidy or target groups. Disentangling the effects of different subsidies being simultaneously introduced is not a problem when this approach is used as simulations can follow in stages. Moreover, disentangling the different basic components of the impact, as those arriving directly from the policy change, the composition of the treated (selection) or indirect changes in prices, is also possible.

A major drawback of the structural approach concerns to the validation of the model. Is frequently difficult, or even impossible, to assess how closely the model resembles the reality it aims at represent. A recent improvement on this direction involved the combination of cohort studies information with a social experiment. This was done in the context of the PROGRESA programme, an educational programme aimed, among other things, at retaining children in school in rural Mexico.[16] The combination of long longitudinal information with an experimental design has proved extremely useful to validate the structural approach. Todd and Wolpin (03) use the experimental data collected before the start of the treatment to estimate a structural model of schooling (and fertility) decisions and then simulate the model to produce structural estimates of the impact of the intervention. At the same time, they use the experimental feature of the data to estimate the impact of the intervention using the observed after treatment outcomes. The comparison of the two estimates provides valuable information about the adequacy of the structural assumptions being used.

5. Conclusions

This study discusses the different approaches to evaluation currently used in the literature, distinguishing between the purely empirical and the structural methods. This is done in the context of a pre-school educational programme recently released in Portugal, the PDEPE. The identification of the treatment effects is discussed under the conditions of the specific policy being studied.

The need for an exogenous source of variation was established if the aim is to produce an *ex-post* evaluation of the policy. It may either be provided by rich data or by the design of the policy. However, it is not clear that it exists in the specific case of the PDEPE, even for the least demanding outcome of interest possible. The identified sources of variation are not enough to separate the impact from the different components of the policy given its complex design. Possible solutions are advanced, which involve both data requirements and policy re-design.

An alternative is to use structural methods. The complexity of the policy design is generally not relevant here. Moreover, a detailed description of the impact of the policy on a wide range of outcomes of interest is possible. Comparisons between the effects of different policies can also be performed, independently of whether the policy has actually ever been released. However, the estimation of structural models is difficult and requires

[16] The data used in this study is similar in shape to a cohort study as detailed family and individual characteristics are collected and families are followed for an extensive period of time. However, it was collected in the context of an experiment, where access to treatment was randomised.

rich information. If not adequately estimated, the model is likely to produce unreliable information about the potential effects of policies. The estimation process becomes all more important given that validation is often difficult.

References

Abadie, Alberto, Joshua Angrist and Guido Imbens (1998), "Instrumental Variables Estimation of Quantile Treatment Effects", NBER *working paper* 229, forthcoming in Econometrica.

Angrist, Joshua and Alan Krueger (1999), Empirical Strategies in Labor Economics, *in* Ashenfelter and Card (eds), *Handbook of Labor Economics*, vol 3A, Amsterdam: North Holland, pp 1277-1366.

Attanasio, Orazio, Costas Meghir and Ana Santiago (2002), "Educational Choices in Mexico", mimeo, University College London.

Auerbach, Alan and Laurence Kotlikoff (1983), "An examination of empirical tests of social security and savings", *in* Helpman and Sadka (eds.) *Social Policy Evaluation: An Economic Perspective*, New York: Academic Press.

Auerbach, Alan and Laurence Kotlikoff (1987). Dynamic Fiscal Policy, Cambridge: Cambridge University Press.

Bassi, Laurie (1983), "The Effect of CETA on the Post-Program Earnings of Participants", *Journal of Human Resources*, 18(Fall), 539-556.

Bassi, Laurie (1984), "Estimating the Effects of Training Programs with Nonrandom Selection", *Review of Economics and Statistics*, 66(1), 36-43.

Behrman, Jere, Piyali Sungupta and Petra Todd (2002), "Progressing through PROGRESA: An impact assessment of a school subsidy experiment", mimeo, University of Pennsylvania.

Blundell, Richard and Monica Costa Dias (2000), "Evaluation Methods for Non-Experimental Data", *Fiscal Studies*, 21(4), 427-68.

Blundell, Richard and Monica Costa Dias (2001), "Alternative Approaches to Evaluation in Empirical Microeconomics", *Portuguese Economic Journal.*

Blundell, Richard, Monica Costa Dias and Costas Meghir (2003), "The impact of human capital policies: a general equilibrium approach", mimeo, Institute for Fiscal Studies.

Blundell, Richard, Monica Costa Dias, Costas Meghir and John Van Reenen (2004), "Evaluating the Employment Impact of a Mandatory Job Search Assistance Program", *Journal of the European Economic Association*, forthcoming.

Bovenberg, Ary and Bas Jacobs (2001), "Redistribution and education subsidies are siamese twins", CEPR *working paper* 3309.

Browning, Martin, Lars Hansen and James Heckman (1999), "Micro Data and General Equilibrium Models", *in* Taylor and Woodford (eds.), *Handbook of Macroeconomics*, volume 1, Amsterdam: Elsevier Science.

Donohue, John and Peter Siegelman (1998), "Allocating Resources among Prisons and Social Programs in the Battle Against Crime", *Journal of Legal Studies*, XXVII(1): 1-43.

Hausman, Jerry, and David Wise (1985), Social Experimentation, NBER, Chicago: University of Chicago Press.

Hanushek, Eric, Charles Leung and Kuzey Yilmaz (2003), "Redistribution through education and other transfer mechanisms", *Journal of Monetary Economics*, 50(8): 1719-50.

Heckman, James (2001), "Accounting for heterogeneity, diversity and general equilibrium in evaluating social programs", *Economic Journal*.

Heckman, James and Aaron Forum (2001), "Policies to Foster Human Capital"', Harris School working paper.

Heckman, James, Robert LaLonde and Jeffrey Smith (1998), "The Economics and Econometrics of Active Labor Market Programs", *in* Ashenfelter and Card (eds.), *Handbook of Labor Economics*, Volume III, Amsterdam: Elsevier Science.

Heckman, James, Lance Lochner and Christopher Taber (1998), "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents", *Review of Economic Dynamics*, 1, 1-58.

Heckman, James and Richard Robb (1985), "Alternative Methods for Evaluating the Impact of Interventions", *in* Longitudinal Analysis of Labour Market Data, New York: Wiley.

Heckman, James and Richard Robb (1986), "Alternative Methods for Solving the Problem of Selection Bias in Evaluating the Impact of Treatments on Outcomes", *in* Howard Wainer (ed.), Drawing Inferences from Self-Selected Samples, Berlin: Springer Verlag.

Keane, Michael and Kenneth Wolpin (1994), "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation: Monte Carlo Evidence", *Review of Economics and Statistics*, 76(4): 648-72.

Keane, Michael and Kenneth Wolpin (1997), "The Career Decisions of Young Men", *Journal of Political Economy*, 105(3): 473-522.

Krueger, Dirk and Felix Kubler (2003), "Computing equilibrium in OLG models with stochastic production", *Journal of Economic Dynamics and Control*, forthcoming.

Lise, Jeremy, Shannon Seitz and Jeffrey Smith (2003), "Equilibrium policy experiments and the evaluation of social programs", IZA *discussion paper* 758.

Todd, Petra and Kenneth Wolpin (2003), "Using a social experiment to validate a dynamic behavioral model of child schooling and fertility: assessing the impact of a school subsidy program in Mexico", mimeo, University of Pennsylvania.

WAGE DIFFERENTIALS AND WAGE SPILLOVERS OF FOREIGN FIRMS: EVIDENCE FROM DIFFERENT ESTIMATORS[*]

Pedro Silva Martins[#]
Queen Mary, University of London

16 July 2004

Abstract:

We evaluate the impact of Foreign Direct Investment on the Portuguese labour market using a matched employer-employee panel covering the manufacturing sector between 1991-99. Wage differentials are found to be negligible when considering methods such as propensity score matching and difference-in-differences. Using OLS, the multinational premium also falls substantially when firm and/or worker characteristics are added and does not vary monotonically with foreign control. Spillover effects (from foreign presence to wages paid by domestic firms) are significantly positive, even when one accounts for unobserved heterogeneity, endogeneity and the domestic industry size. Foreign-presence elasticities of domestic wages range between 2% and 3%.

JEL codes: C23, F23, J31.

Keywords: FDI, Spillovers, Wages, Matched Employer-Employee Data, Differences-in-Difference, Propensity Score Matching, Portugal.

[*] I thank Davide Castellani, Arnaud Chevalier, Sourafel Girma, Diogo Lucena, Robin Naylor, Ian Walker, seminar participants at HWWA (Hamburg) and GEP (Nottingham) and the Scientific Committee of the Conference for useful comments and suggestions. Financial support from "Fundação para a Ciência e a Tecnologia" (POCTI/ECO/33089/99 and SFRH/BD/934/2000) and the British Council and logistical support from the Bank of Portugal and ETLA (Helsinki) are gratefully acknowledged. Some of this work was carried out when I was affiliated with the University of St Andrews. The standard disclaimer applies.

[#] Address: Centre for Business Management, University of London, Mile End Road, London E1 4NS, United Kingdom. Email: p.martins@qmul.ac.uk. Web: http://alpha.qmul.ac.uk/~bsw019. [Other versions of this chapter and related papers can be found at this website.]

Sumário Executivo [Executive Summary, in Portuguese]

Os fluxos de Investimento Directo Estrangeiro (IDE) para Portugal tornaram-se uma componente importante da economia portuguesa desde a adesão à União Europeia. Não se poderá, no entanto, ignorar o custo destas entradas de capitais, na medida em que elas foram influenciadas por dispendiosas políticas públicas para a atracção de multinacionais para Portugal. Assim, num contexto mundial em que a captação de IDE se torna cada vez mais competitiva, e num contexto nacional em que se discute a melhor orientação para a economia portuguesa, torna-se particularmente importante avaliar qual o benefício para a economia portuguesa destes investimentos.

Este trabalho contribui para este objectivo, examinando o papel do IDE no mercado de trabalho português e, em particular, nos salários. Esta abordagem, complementar às que estudam o impacto do IDE na produtividade das empresas e no emprego, justifica-se pela importância dos salários para o bem-estar da população. De facto, se o IDE não tiver implicações salariais, então o seu impacto económico para a generalidade das pessoas será provavelmente limitado.

Neste sentido, consideramos que os benefícios do IDE são mais importantes se: 1) os trabalhadores das empresas estrangeiras em Portugal forem melhor pagos que os seus equivalentes nas empresas nacionais; e 2) os trabalhadores das empresas nacionais forem melhor pagos como consequência da presença estrangeira no seu sector de actividade (i.e., se existirem "spillovers" salariais). São estas duas relações entre as empresas estrangeiras e as empresas nacionais que são examinadas neste trabalho, cujos resultados se baseiam na análise da base de dados dos "Quadros de Pessoal" para o sector secundário, no período 1991-1999.

Em relação ao primeiro aspecto, o dos diferenciais de salários entre trabalhadores em empresas nacionais e estrangeiras, uma primeira análise (usando métodos econométricos convencionais) sugere que as empresas estrangeiras pagam cerca de 32% mais que as empresas nacionais. No entanto, este prémio reduz-se substancialmente até cerca de 11% quando se têm em conta algumas das diferentes características dos trabalhadores e, sobretudo, das empresas, e que também influenciam os salários (educação, sector de actividade, dimensão da empresa, etc). Por outro lado, não se encontra evidência de uma relação clara entre o grau de controlo de uma empresa por investidores estrangeiros e o prémio salarial pago, ao contrário do que se esperaria caso o grau de controlo por parte de empresas estrangeiras tivesse um impacto causal nos salários.

Outras análises, usando métodos mais sofisticados e precisos de comparação de trabalhadores em empresas nacionais e estrangeiras, sugerem um prémio salarial ainda mais baixo: cerca de 1%. Finalmente, quando se compara a evolução dos salários dos trabalhadores de empresas nacionais adquiridas por investidores estrangeiros com a mesma evolução para trabalhadores de empresas nacionais que se mantêm na posse de investidores nacionais, o prémio salarial passa mesmo a incluir valores negativos.

O segundo aspecto examinado neste trabalho refere-se aos "spillovers" da presença de empresas estrangeiras em cada indústria nos salários pagos pelas empresas nacionais nessas indústrias. Aqui, ao contrário da situação anterior, constata-se que, pelo menos quando se utilizam métodos mais robustos, baseados na relação entre a variação da presença estrangeira e a variação dos salários das empresas nacionais, existe uma influência positiva importante da primeira na segunda. De acordo com estes resultados, se se duplicar o peso das empresas estrangeiras num dado sector, mantendo constante o tamanho

doméstico do sector, os salários pagos pelas empresas nacionais aumentam, em média, entre 1,9% e 2,7%. Estes valores implicam aumentos salariais anuais para os trabalhadores de empresas nacionais entre aproximadamente 730€ e 1.070€ por cada novo trabalhador de uma empresa estrangeira.

Como conclusões, os nossos resultados indicam que não é tanto a nacionalidade das empresas que explica as diferenças de salários entre empresas portuguesas e estrangeiras, mas sim as diferentes características de cada conjunto de empresas. No entanto, é importante ter em conta que há relativamente poucas empresas portuguesas com características semelhantes às das empresas estrangeiras, características essas (como o tamanho) que estão associadas a maiores salários. Encontrou-se ainda evidência de "spillovers" salariais das empresas estrangeiras para as empresas nacionais, pelo menos dentro do mesmo sector de actividade. Este último resultado sugere que há diferenças entre o benefício privado e o benefício social da presença de empresas estrangeiras, diferenças essas que podem justificar intervenções públicas que aproximem os dois tipos de benefícios, por exemplo através de subsídios.

Assim, em termos de implicações de política económica, os resultados obtidos constituem evidência empírica que apoia a defesa do IDE como mecanismo de desenvolvimento económico em Portugal. Por um lado, as empresas estrangeiras têm características associadas a melhores remunerações; por outro lado, a presença de empresas estrangeiras numa dada indústria leva a melhores remunerações nas empresas nacionais nessa mesma indústria. Por último, este trabalho sublinha a importância de se caracterizar rigorosamente os impactos – salariais e outros – do IDE, de modo a se maximizar os benefícios líquidos para Portugal da política de atracção de investimento estrangeiro.

Introduction

Large foreign direct investment (FDI) inflows have become an important feature of the Portuguese economy following EU accession.[1] These massive flows were not however without a cost, as expensive government policies were developed in order to attract multinationals to Portugal.[2] Moreover, this attraction process will probably become more expensive, as globalisation (and EU enlargement in particular) implies that multinationals have increasingly more geographical alternatives for their investments.

This empirical paper contributes to the evaluation of the usefulness of inward FDI for the Portuguese economy. Such an evaluation is particularly important given the current debate about the direction of the economy and the subsequent optimal design of public policies. We thus contribute for an answer to the question: *To what extent should taxpayers' money be used for attracting multinationals?*

Using data from about five million employees over the 1991-1999 period, the perspective adopted here is to examine the impacts of foreign multinationals in the Portuguese labour market and, in particular, in the wages paid to workers. This is pursued via considering two types of impacts. The first impact is about the differences in pay between workers in foreign firms and those in domestic firms. The second impact concerns the spillovers upon pay of workers of domestic firms that are related to the presence of multinationals in the relevant industry.[3]

The structure of the paper (and a short summary of the results) is as follows. Section 2 briefly surveys some of the key contributions in the literature on FDI and labour markets. One stylised fact from this literature is that workers in foreign firms are typically better paid than their counterparts in domestic firms. However, most of the evidence on spillovers suggests that, at best, these latter indirect benefits of FDI are negligible.

Section 3 presents the data set used, a large matched employer-employee panel, and some descriptive statistics. It is found that, while workers in foreign firms are indeed better paid, they also present some characteristics that are typically associated with higher productivity (e.g. higher schooling attainment). Foreign firms are also found to have some very different characteristics (e.g. bigger size) from their domestic counterparts.

Section 4 presents the results on the wage differentials: as in other studies and countries, foreign firms are found, in a first analysis, to pay higher wages. However, this premium falls by almost two thirds, from 30% to 11%, when taking into account worker and firm characteristics. This pay gap is found to fall even further and eventually to become insignificant when we use more sophisticated methods, that control for compositional effects and unobserved heterogeneity, such as differences-in-differences and propensity score matching.

---

1 In 2000, FDI outflows (disinvestments) reached a record level of 16% of GDP, leading to a net inflow of 6% of GDP. See also Ministério das Finanças (2003) for a recent and detailed analysis of inward and outward FDI in Portugal and other countries, covering the period 1996-2001.

2 A recent example (2001) is an investment of €37 million in the car components industry made by a French multinational in north-eastern Portugal, in which the Portuguese government awarded benefits with a present value of about €4 million or about 11% of the amount invested (author's calculations based on the investment contract and information from Ministério das Finanças, 2003, p. 105).

3 Work in progress, not included in this paper, includes the analysis of a third impact, that related to the mobility of workers from foreign to domestic firms and its implications on wages.

Section 5 is about the wage spillovers related to foreign presence. Unlike before, here more sophisticated methods produce a positive impact. The elasticity of wages paid by domestic firms with respect to the employment of foreign firms is found to range between 1.9% and 2.7%. This is important evidence supporting the case for the attraction of FDI.

Finally, section 6 summarises the key findings, concludes and provides some policy-oriented comments.

1. Literature

There is a large literature, both theoretical and empirical, on the impact of FDI on host economies. From the theoretical contributions, this impact is ambiguous, as it depends on different factors of opposite effects. Such factors include the size of wage and/or productivity differentials, spillovers and market-stealing effects and, on the cost side, the level of transfers from the host countries to multinationals.

With respect to wage differentials, this line of research typically finds that workers of multinationals are paid more than their observably similar counterparts in domestic firms. Some examples are Aitken, Harrison and Lipsey (1996) for Mexico, Venezuela and the USA, Feliciano and Lipsey (1999) for the USA, and Lipsey and Sjoholm (2001) for Indonesia. (Other papers that examine this issue by drawing on the acquisition of domestic firms are reviewed in Section 4.4.)

The first paper finds evidence of higher wages in foreign firms in the three countries covered using establishment cross-sections. Feliciano and Lipsey (1999) focus on the US case using industry-state-ownership cells and again find evidence of a positive wage differential. Lipsey and Sjoholm (2001) finds a premium of 12% for blue-collar workers and 22% for white-collar workers, drawing on a 1996 cross-section of plants and regressing average wages on plant characteristics, including its domestic or foreign ownership.

Given these results, it is important to ask why is it that foreign firms apparently pay higher wages. A first explanation is that foreign firms, as they enter the domestic country, may increase the demand of labour and thus wages for the marginal workers hired by the multinational. However, under a competitive model of the labour market, these differentials would not be sustained for long. Indeed, in this case the rents of workers in foreign firms would be competed away via worker mobility from domestic firms.[4]

A second explanation is that multinationals may pay a wage premium to dissuade workers from moving to other firms. This is consistent with the Ownership-Location-Internalisation model of Dunning (1977), in particular with the internalisation motive.[5] Under this model, foreign firms own special assets (e.g. new products or technologies) that are more profitably exploited abroad via FDI than, say, via exporting. Foreign firms would then pay higher wages in order to prevent the costly turnover which would occur when quitters provided insights about the multinationals' new methods and/or products to competing firms.

---

4 A competitive model is however unlikely to apply in the Portuguese case, given the many constraints to flexibility. See Martins (2003b), for instance, for evidence of substantial levels of rent sharing in the Portuguese labour market. Blanchard and Portugal (2001) highlight the large burden to economic efficiency brought by restrictive firing regulations.

5 Other prominent models of FDI, such as the vertical, horizontal and capital-knowledge models, are more difficult to be used in the context of models of pay determination.

Another explanation is about compensating differentials: multinationals may provide less interesting job amenities than domestic firms. The little evidence available suggests the opposite however, including Sutherland (2002) and Brown et al (2003) for developed and developing countries, respectively.

Finally, foreign firms may select more skilled workers in dimensions unobservable to the researcher. For instance, given their larger size, multinationals may benefit from scale economies in setting up sophisticated human resource departments. This would allow these firms to screen and hire the best applicants for their vacancies, also along dimensions of skill not observable by the researcher. Under this case, the wage differential attributed to multinationals would be spurious as no causal impact of foreign ownership upon wages could be inferred. This also means that evidence of wage differentials should not necessarily support the case for government intervention.

As to productivity spillovers, they have been categorised into three types: demonstration, competition, and labour mobility effects – see Blomstrom and Kokko (1998) for a survey. Demonstration effects occur when national firms learn new practices and technologies from observing multinationals. Competition effects result from national firms being forced to become more efficient by the presence of foreign firms in their markets. Finally, labour mobility effects are present when workers acquire new skills in foreign firms that may then benefit domestic firms when workers leave the multinationals.

While the theoretical results are ambiguous, the empirical evidence on spillovers is clearer. Gorg and Strobl (2001), in a meta-analysis of the empirical literature of FDI spillovers, find that most recent studies, in particular those that use panel data, fail to detect any improvements in domestic productivity arising from foreign direct investment.

One influential example is the study of Aitken and Harrison (1999), that documents negative spillovers The authors look at a panel of Venezuelan firms and regress plant output on the share of foreign ownership at the industry (weighted by employment) and on a number of controls, including time, industry and region dummies. The results differ between foreign and domestic firms: for the former, a small positive impact of foreign presence in the industry is found; for the latter, a negative effect is uncovered. This finding is consistent with their theoretical model, which predicts an ambiguous result, depending on whether positive spillovers or market-stealing effects dominate: the former shift the average cost curve downward whereas the latter increase average cost, due to lower output.[6]

The stylised result of no positive spillovers from FDI is of great importance, as it weakens the main economic argument for governments to spend resources in attracting FDI. If there is no gap between the private benefits of investment (from the point of view of the multinational) and the social benefits of that investment in the host country, then there are no economic reasons for the large hand-outs of public money observed in many years and in many countries.

However, one of the few papers we know of that looks at wage spillovers of FDI, Aitken et al. (1996), finds evidence supporting their existence. (The same applies to Driffield and Girma, 2003). In Aitken et al, The authors use plant data aggregated at the 4-

[6] Haddad and Harrison (1993) also find negative spillovers in the productivity growth of domestic firms in Morroco. However, while a recent study focusing on the UK, using panel data, Haskel et al (2002), finds positive spillovers, Harris and Robinson (2003), using the same data, finds negative effects. See also Girma et al (2001).

digit industry level and regress the logarithm of average wages on several control and the share of FDI at each cell, defined by industry, region and skill level, either for all firms or for the subset of domestic firms. Using panel data for Mexico and Venezuela and cross-section data for the USA, the authors find that higher levels of foreign ownership are associated with higher wages across all countries, although this magnitude is smaller for the USA.

With respect to the specific case of Portugal, studies find evidence of FDI spillovers in terms of productivity, unlike in most of the literature summarised above. This result for Portugal holds either for the full set of domestic firms (Mata and Farinha, 1996) or for a subset of firms which present intermediate technological gaps with respect to their foreign counterparts (Flôres et al., 2000 and Proença et al., 2002).[7]

However, these studies do not examine the wage implications of FDI. (The only exception we know of is Almeida, 2003, which is discussed in Section 4.3.) They also draw on data at the firm or industry level only, thus disregarding individual-level compositional changes that may be involved in FDI. This study fills in these gaps, providing at the same time a more encompassing analysis of the range of implications that FDI may have upon the Portuguese labour market.

2. Data

Our empirical analysis draws on a large matched employer-employee panel data set, "Quadros de Pessoal". This is a survey that has covered the Portuguese universe of employees since at least 1982 and in which all firms with at least one employee are required to fill in a survey about some of their characteristics and those of their workforce. The former characteristics include variables such as geographical location, industry, sales and, crucially from the point of view of this paper, the share of the firm's equity owned by foreign parties. The set of worker characteristics includes variables such as gender, highest level of education attainment, tenure, wages, hours worked, etc.

The sub-sample considered here is that of an 80% representative sample of manufacturing sector firms, for each year between 1991 and 1999. This amounts to an annual average of about 678,000 workers per year and 29,481 firms (and a total of 71,240 different firms) – the yearly data are available on Tables 2 (workers) and 3 (firms). From this initial sample, missing observations and observations with inconsistent information in key variables were dropped. Also only firms based in the continent are considered. This gave rise to a data set with an annual average of about 496,000 workers and 16,935 firms (and a total of 39,783 different firms).

[insert Tables 2a, 2b, 3a and 3b about here]

(see Appendix Tables A1a, A1b and A2)

As mentioned before, the foreign/domestic ownership of the firms is derived from the share of equity held by foreign nationals. In particular, we define as foreign firms those for which at least 50% of equity is owned by foreign parties, although below we also

[7] Other papers that look at different aspects of FDI in Portugal include Cabral (1996), Barry (1999), Guimarães et al. (2000), Mata and Portugal (2000, 2002 and 2004), Nogueira Leite et al (2001), and Barbosa and Louri (2002).

consider the 10% threshold. According to this criterion, we find an annual average of 16,557 domestic and of 378 foreign firms. These correspond to about 434,000 workers in domestic firms and about 62,000 workers in foreign firms, which implies that, over the 1990's, an average of about 13% of employees worked in foreign-owned firms. This share also follows an increasing trend, consistently with the net inflows of FDI and the increasing share of services (to the detriment of manufacturing) in the Portuguese economy.

The descriptive statistics reveal some important differences between foreign and domestic firms. For instance, in all measures of earnings (monthly or hourly), workers in foreign firms earn more than their counterparts. In 1991, for example, the average worker in a foreign firm earned €691 per month (2000 prices, values gross of taxes), whilst the average worker in a domestic firm earned €528, which implies a pay premium of 30.9%. In the end of the decade, the corresponding values were €863 and €628 (and a pay premium of 37.4%).

These differences in earnings can be partly explained by differences in schooling attainment: workers in foreign firms are found to have an average of 6.9 years of schooling across the nine years covered, whereas the same figure for workers in domestic firms is only 5.8. Affecting the results in the opposite direction, workers in domestic firms are found to be more experienced. This is partly due to their lower schooling, as we measure experience here as (age–education–6). Tenure (a proxy for firm-specific skills) is also found to be slightly higher for workers in domestic firms, which would again go against an explanation of higher wages for workers in foreign firms due to human capital differences. Similarly, there is a larger share of women working for foreign firms.

As to the distribution of workers by industries, there are again clear differences between domestic and foreign firms. A particularly striking contrast is the over-representation of foreign firms in the metallic and transport industry (43% against 19% of workers). In an opposite direction stands the textiles, clothing and leather industry (29% of foreign workers against 38% of domestic workers) and the wood and cork industry (1% against 8%). These differences have also a geographical dimension, as workers of foreign firm are relatively more prevalent in the Lisbon region (40% versus 23%) and less prevalent in the North region (40% versus 52%).

A complementary way to establish the contrast between domestic and foreign ownership can be pursued by looking at firm (rather than worker) statistics. This amounts at looking at the same data as that of Table 2, but without weighting firm characteristics by firm size (number of workers). Instead, here we average firm characteristics and then look at simple averages of such firm averages. This exercise – see Table 3 – reveals even stronger pay differences between foreign and domestic firms. For instance, in 1999, the average multinational paid an average of €889 to an average worker, whereas the average domestic firm paid €502 to an average worker. This amount to a pay premium of 77.1%, more than twice the one documented for same year in Table 2.

An important explanation for this difference between the results for firms and workers lies on the fact that foreign firms are, on average, much larger (in terms of number of workers) than domestic firms. The former have an average of 190 workers whereas the latter have an average of only 32. This helps in understanding the very large pay difference as it is well known from the labour economics literature that smaller firms pay lower wages (see Brown and Medoff, 1989, and Oi and Idson, 1999). No other sizeable differences are found in terms of the industry or geographical distributions of domestic and foreign firms with respect to the previous table.

For the benefit of robustness and comparability, we also consider a threshold of 10% for the definition of foreign firms. This is also the criterion adopted by the IMF to distinguish between portfolio and direct investments. We have opted for the 50% level because we focus on firm ownership (and not investments in firms). However, one should be aware that this level is neither a sufficient nor a necessary condition for foreign firm ownership, given the information that we have available. For instance, a foreign investor with only 25% of equity may control a firm if the remaining equity is spread over a large number of stakeholders. But we may also be before a case in which two separate foreign investors with 30% each do not control a firm if, for instance, a third, domestic party holds 40%.

In the appendix we present descriptive statistics for workers over the nine-year period covered under the 10% classification. We find a relatively small increase in the percentage of workers in foreign firms (from 13% to 16%), indicating that most firms with a positive level of foreign ownership are above the 50% threshold, a result which is consistent with Barbosa and Louri (2002). From the comparison of domestic and foreign firms in this case, few differences are found with respect to the previous classification.

3. Wage Differentials

This section examines the robustness of the substantial pay premia for workers in foreign firms documented in the descriptive statistics examined above. A first step, producing benchmark results, involves an OLS estimator, allowing for like-for-like comparisons, considering progressively more extended sets of control variables. In fact, it could be the case that the pay premium is fully explained away by the characteristics of the workers (e.g. education) and/or their firms (e.g. size).

Secondly, we study whether the pay premia depend on the level of control of the foreign party, as measured by the share of equity held by the latter. The motivation for this is to examine whether the impact of foreign ownership upon wages underpins some causal process, which would probably imply that a greater level of control of the firm would be translated into higher wages.

Thirdly, the propensity score method is used. Under this approach, national- and foreign-firm workers are matched according to their many observable characteristics. From these comparisons, a measure of the wage difference between strictly comparable groups of workers is obtained.

It is also of great interest to know the wage impacts of acquisitions of domestic firms by foreign parties. The motivation for this is that foreign firms may hire workers that are more skilled along unobservable dimensions. This possibility – unobserved heterogeneity – would explain the premium documented via the OLS estimator. However, one can circumvent this problem by using a differences-in-differences estimator. In this case, this amounts to comparing wage differences (i.e. wage growth) for workers whose domestic firms are acquired by foreign parties and wage differences for workers whose domestic firms remain domestic.

Finally, the differences-in-differences estimator and the propensity score method are combined. Some research, namely Heckman et al (1997), has suggested that this approach may be the most appropriate for evaluation studies. The intuition behind this result is that one can simultaneously difference out the unobserved heterogeneity variables

and still tackle the problems of lack of common support and different distribution of characteristics allowed for by the propensity score matching method.

3.1. OLS

In this section, different versions of the following wage equation are considered, separately for each year:

$$y_i = X_i'\beta_1 + F_{j(i)}'\beta_2 + \beta_3 Foreign_{j(i)} + \varepsilon_i \,, \quad (1)$$

where y_i denotes the logarithm of real hourly wages. X_i denotes a set of human capital characteristics (six dummies for educational attainment, a quartic in experience, a quadratic in tenure – measured in months, and a female dummy). $F_{j(i)}$ denotes a set of characteristics of the firm of worker i (four regional dummies, log number of workers, a dummy for public firms and eight industry dummies[8]). $Foreign_{j(i)}$ is a dummy variable taking value one if the firm of worker *i* is a foreign multinational and value zero otherwise. Our attention in the first part of the paper is focused on the β_3 coefficient.[9] ε_i is an error term following the standard assumptions.

The first version of equation (1) includes only the foreign dummy. In this case, the wage premium for workers in foreign firms is found to average 32% and range between 27% and 37%, depending on the year – see Table 4. However, as suggested before, these wage differences may be attributed to different levels of human capital. Indeed, it is found that, when one controls for these variables, the average premium falls to 27% (ranging between 23% and 32%). The adjusted R^2 statistic also increases markedly, from about 3% to about 45%.

[insert Table 4 about here]

Finally, when firm characteristics are also considered, the premium falls further and more substantially than when human capital variables were added. The average premium is now only 11%, ranging between 8% and 13% across the different years. The adjusted R^2 statistic now increases more modestly, to about 53%.

As to time trends, there is some evidence, in the first and second specifications, of increasing foreign multinational wage premia between 1991 and 1995, when they peak, and then declining premia from then until 1999.[10] This pattern is however not present in specification that includes both human capital and firm controls.

Overall, these results suggest that about two thirds of the average difference in pay between domestic and foreign firms is determined by the different human capital of their

---

8 The industry coding was changed in 1995. We adopt the code used until that year ("CAE-Rev.1") in the entire period. Given that some firms are not available in our sample in the period 1991-1994, their industry code was obtained from extrapolating from the changes observed in such codes for firms present in both periods (1991-94 and 1995-99).

9 This coefficient is discussed in terms of percentage differences in wages, after taking into account the logarithmic transformation of the dependent variable.

10 A similar pattern is found for returns to education – see Pereira and Martins (2002).

workers and, more importantly, by the different characteristics of foreign and domestic firms. However, at 11%, the pay gap between domestic and foreign firms is still economically significant and may suggest an important direct effect of multinationals in the labour market. Similar findings are obtained when the 10% threshold is considered (see Table A2), except that the foreign premium is slightly bigger in all specifications and years.

These results assume a constant foreign pay premium for different industries. However, there are many reasons for one to expect such wage differentials to depend on the industry examined. For instance, and given the theoretical considerations before, in skill-intensive industries multinationals may have to pay large premia so to prevent the dissemination of their new methods to competitors via workers' mobility. Conversely, in low-skill industries, multinationals may be just happy to pay standard market wages, especially if these are much lower than in their own domestic labour markets.[11]

We therefore allow the multinational premium to depend on the industry. Table 5 presents the results: the first eight main rows denote the average industry premia, the following row refers to the average foreign firm premium and the last eight rows denote the industry-specific premium. The comparison group (dropped dummy) is the Clothing, Textiles and Leather industry, the biggest one in our data and also one of the industries that pays the lowest premia.[12] In contrast, the four industries that exhibit higher average premia are Wood, Paper, Chemicals, Non-Metallic Minerals and Metallurgic sectors.

[insert Table 5 about here]

As to the differences in the multinational premium, the industries where this is larger are Food, Wood, Paper and Chemicals, where the average additional premia range between 26% (food) and 17% (chemicals). In the remaining industries, Clothing, Non-Metallic Minerals, Metallurgic, Metallic and Others, the average additional premia range between 8% (Non-Metallic Minerals) and 0% (the dropped industry, Clothing).

These findings may represent some support to the hypothesis described above about different incentives across different industries to pay above-market wage rates. For instance, Chemicals pay higher foreign premia and are probably a good example of an industry where patents and other inputs that generate rents are prevalent. On the other hand, Clothing, which pays the lowest wages, may again be an example of an industry that largely uses standard technical processes. These results may, however, be affected by the impact of wage spillovers. For instance, it may be no coincidence that the foreign differential is lower in industries that exhibit an important foreign presence, such as Textiles and Metallic products.

3.2. Different Levels of Control

In this section, we allow the wage impact of foreign firms to differ depending on the degree of control of the firm by the multinational, as implicit in the share of the firm's equity held by the latter. This is done by creating different dummy variables for firms with different share of foreign control: 1%-9%, 10-19%, …, 90%-99% and 100%. Within the

---

11 There is also a large literature on inter-industry wage differentials, documenting in general substantial differences in pay across industries. See Vieira et al (2000) for evidence for Portugal.

12 See Martins (2003a) for a study of wage determination in this industry.

group of firms which are considered as foreign-owned, the most prevalent category is that of full foreign control (which accounts for 60% of workers) and then that of 90%-99% (20% of workers). Within the group of firms that have less than 50% foreign shares, they correspond to between .5% and 2% of the workers in domestic firms.

We find – see Table 6a – positive premia for all workers whose firms have some positive share of equity held by foreign parties, regardless of the size of that share. For instance, firms with a share of equity of between 1% and 9% pay their workers, on average across the decade, 12% more than similar workers in firms without any share of equity held by foreigners.

Although average premia are higher in the 50%-100% range than in the 1%-49% range, there is no clear evidence of a monotonic relationship between premia and the share of foreign ownership. This result weakens the case for a strong causal relationship between the degree of control and the wage differential. However, one must bear in mind the relatively low number of workers in firms that exhibit low but positive levels of foreign ownership. This may explain the considerable volatility in the estimates for lower levels of foreign ownership as different years are considered.

As a further check on this link between the degree of foreign control and the wage premium, we also investigated the wage difference between domestic and foreign firms when the latter are defined at different thresholds. Here we consider ownership starting alternatively at the 10%, 20%, 30%, 40% and 50% levels and compute the premia accordingly. Taking into account the hypothesis described above that greater control translates into higher premia, we expect that the greater the ownership threshold, the higher will the premium be.

The results – see Table 6b – indicate that, not only is there no monotonically positive relationship between the premium and the level of control, as in most years the premium is found to fall at higher levels of control. These results are further evidence against the existence of a causal link between foreign ownership and wages.

[insert Tables 6a and 6b about here]

3.3. Propensity score matching

One concern with standard OLS estimators is that their "like-for-like" comparisons disregard the possibility of a different support (i.e. range of values of the regressors) between observations with and without some characteristic whose impact is of interest. Moreover, the distribution of characteristics over that region of common support may also be very different between the two groups.

Propensity score matching is a non-parametric method that allows one to tackle these issues directly, by restricting the estimation of the impact of interest over the region of common support only and making such estimation take into account the distribution of the variables. Heckman et al (1997) suggest that these two biases (non-overlapping support and different distributions of covariates within groups) are more important than that related to selection on unobservables, which has received much more attention in the literature on programme evaluation. In their words, "the simple balancing of observables goes a long way towards effective evaluation" (p. 607).

The relevance of this method in this paper stems from the large differences documented in terms of the observable characteristics of the two types of workers. These differences are particularly pronounced on the educational attainment of workers, and on firm size and industry distribution. Failure to take this into account may thus have biased our results.

The new findings are obtained using the same annual data sets described before. The variables considered for the estimation of the propensity score are: six education dummies, experience, tenure, gender, log firm size, three region dummies (north, centre and Lisbon) and four industry dummies (textiles, wood, chemicals and metallic). It is this propensity score that is used to match workers in foreign firms (the "treatment group") to those in domestic firms. Observations from both groups are then paired via one-to-one matching. Standard errors are obtained with bootstrapping (50 repetitions).

The first nine rows of Table 7 present the results. The estimates of the foreign firm premium range between –5% (1994) and 6% (1999), averaging 1.06% over the nine years studied. Some coefficients are significant while others are not. Overall, the small magnitude of the coefficients and their insignificancy (in some cases) stands in clear contrast with the initial OLS findings in this paper and in the literature on the impacts of foreign ownership on wages. However, this is not at odds with the finding in section 3.1 that a more detailed comparison of workers in domestic and foreign firms reduces the estimates of the foreign-firm wage premium and the finding in section 3.2 of no monotonic relationship between the extent of ownership and the size of the premium.

[insert Table 7 about here]

3.4. Acquisitions – Difference-in-differences

The method followed in this subsection involves contrasting the change in a variable of interest in a group of observations that have undergone some treatment with the change in the same variable in a similar group of observations but which have not undergone treatment. The advantage of this approach is that one is able to control indirectly for variables that may influence the parameter of interest but which are not available, provided that such variables are time-invariant and the assignment to treatment is random. Here, we apply this method for the case of domestic firms that are acquired by multinationals: this acquisition is the treatment whose impacts are studied vis-à-vis the control group of domestic firms that are not acquired over the period considered.

A similar approach is used in Conyon et al (2002) and in Almeida (2003).[13] The first paper uses UK firm level panel data and focus on productivity and wage impacts of acquisition, contrasting the case of acquisitions originated by foreign multinationals and other domestic firms. The authors find evidence of higher wage growth when domestic firms are acquired by multinationals than when domestic firms are not acquired by any firm. However, those domestic firms that are acquired already exhibit higher levels of wage

[13] Other papers include Girma and Gorg (2003) for the UK and Lipsey and Sjoholm (2003) for Indonesia, both drawing on firm-level data. Both papers document higher wage growth for acquired firms, although the first finds that this result applies only to some industries and when the acquirers are of specific nationalities. A third paper, Harris and Robinson (2002), finds decreased productivity in acquired firms in the UK.

growth before acquisition than those firms that are not acquired, suggesting some role for "cherry picking". In any case, the authors find that, after controlling for productivity, the wage differential is eliminated.[14]

Almeida (2003) also studies the wage impacts of foreign acquisition using the "Quadros de Pessoal" data set. However, some important differences in relation to what is done in this subsection are that Almeida (2003) considers both the manufacturing and non-manufacturing sector, data aggregated at the firm level and foreign ownership or acquisition are defined at the 10% share of equity threshold. As to the results, and focusing on the case of manufacturing firms, the estimates of the foreign firm wage premium are found to fall from 0.16 in the OLS cross-section estimator to an insignificant coefficient of 0.02 in the differences-in-differences results. As with Conyon et al (2002), important observational differences between acquired and always-domestic firms are also documented, even before acquisition.

Our contribution to this literature is twofold. Firstly, we take the analysis to worker-level data, in order to minimise the aggregation bias implicit in firm-level information. Secondly, we focus on the same workers before and after acquisition. This should substantially decrease the scope of unobserved heterogeneity to bias our estimates of the foreign ownership effect. Moreover, this approach also means that we do not suffer from the biases that may occur when, upon acquisition, firms change their workforces, particularly when such changes occur along unobserved dimensions.

We construct the data set used as follows. The control group includes all firms which are never foreign owned over the years in which they are available in our sample. Only two years are considered for our analysis: the "before" year is randomly selected in the 1991-1998 range, while the "after" year is the subsequent year available in the sample for that firm. (This is not necessarily the "before" year plus one as the panel is unbalanced.) A total of 23,991 firms (and about 250,000 workers) were found with this method.

The treatment group (domestic firms acquired by foreign parties) was defined as the set of firms whose ownership is initially domestic (i.e. share of equity owned by foreign parties below 50%) and which in some subsequent year become foreign-owned. Firms following this criterion but in which the foreign acquisition process is reversed at a later stage were dropped. A total of 231 firms (and about 18,000 workers) were found.

For the benefit of robustness, a further treatment group was obtained: the subset of firms in the former target group that exhibit a positive level of foreign ownership before acquisition: 26 firms (and 2,578 workers) satisfy this condition. The motivation is to control further for the possible selectivity involved in the process of foreign acquisition. In particular, foreign firms may target domestic firms whose workers have unobservable characteristics that make them different from the control group of not-acquired domestic firms. To the extent that this subset of firms has already been targeted by foreign firms, we may be more likely to pick up with our differences-in-difference estimator only the acquisition effect, rather than a combination between the selection and the acquisition effects.

These three groups were found to be very different, as far as their observable characteristics are concerned – see Table 8a. Focusing on the larger treatment group, the

[14] Similar results are obtained when the acquisition dummy is instrumented. However, it is not clear whether the instruments used in the paper pass the validity tests, i.e. whether they can be excluded from the main equation and whether they contribute to the explanatory power of the auxiliary regression.

treatment workers are more educated than their control group counterparts. The former are also over-represented in bigger firms and in the metallic industry while they are under-represented in the clothing industry, for instance. The distribution of acquisitions per year, in terms of workers, is also less balanced as that of the control group, particularly in 1992 (25% vs 12%) and in 1995 (6% vs 13%). Importantly, there is also a sizeable difference in wage growth for each group of workers: the treatment group has an average wage growth of .6% whereas for the control group that figure is 5.1%.[15]

These differences in wage growth may however be due to the different characteristics of each type of workers, not to the acquisition effect. The impact of the treatment was therefore obtained from running a regression in which the dependent variable is wage growth between the "before" and "after" period, i.e. differences of the log hourly wage, as in the equation below:

$$\Delta y_{it,t-1} = X_{it-1}'\beta_1 + F_{j(i),t-1}'\beta_2 + \beta_3 Foreign_{j(i),t} + \beta_4 \Delta t_{it} + \varepsilon_{it}, \qquad (2)$$

The regressors are the characteristics of the workers and their firms in the "before" period (t-1) and on a dummy variable (Foreign) taking value one for workers in firms that become foreign-owned and value zero for workers in firms that are still domestically-owned in the second period. A control for the difference in years between the two periods is also included.

The results indicate that, after controlling for the differences in worker characteristics, the impact of foreign acquisition on wages is significantly negative (–3.1 p.p.) for the larger treatment group. However, the coefficient becomes significantly positive (2 p.p.) for the subset of firms with a positive level of foreign ownership in the first period.[16]

We also find that the first result is due to the fact that the domestic firms that are subject to an acquisition (in particular those whose equity is entirely owned by domestic agents) are paying wages much above the market rates (conditional of firm and worker characteristics) than their counterparts that will not be subject to an acquisition. Taking the latter group as the benchmark, pooled cross-section OLS results (not reported) find that the domestic firms that will be acquired are paying 11.7% more. After acquisition, the acquired firms are paying a 7.1% premium while those firms that remain domestically owned are paying 0.8% more.

As a further check on the robustness of these results, we considered a different control group: those firms that are always foreign-owned. The motivation for this choice is that the latter firms (and their workers) are, from the observable point of view, more similar to the treatment group under study, particularly the larger one. To that extent they stand as a better control group. It turns out, however, that wage growth is still lower for acquired firms than for always-foreign-owned firms, controlling or not for worker and firm

[15] Another concern is that the subset of treated firms that have some foreign ownership in the before period present even more different characteristics than the entire group of treated firms with respect to the control group.

[16] Similar results were obtained when one disregards the panel nature of the data (i.e. if one runs a pooled cross-section regression and focuses on the interaction between the before/after and the foreign dummies).

characteristics. From the regression results – see Table 8b –, the coefficient on the treatment dummy is –0.042 (–4.1p.p.).

The findings of negative or small positive wage premia are surprising, given the opposite stylised fact in the literature. However, this result is not unexpected from the point of view of the literature on takeovers. Here, it has been hypothesised that a key motivation for the acquisition of firms is for new employers to renege on implicit contracts and to appropriate a larger portion of the surplus produced by employees (Schleifer and Summers, 1988). Other authors claim that mergers and acquisitions stem from a process of improved matches between firms and managers, whereby firms that are being badly run are more likely to be taken over (Lichtenberg and Siegel, 1990).

Two other objections should be taken into account. The first is that one may not extrapolate from the wage impact of foreign acquisitions to the wage differences between "standard" multinationals and domestic firms. Possibly a better control group would be domestic firms acquired by different domestic owners but, unfortunately, there is not information about this in the data set. This criticism is, however, attenuated by the stylised fact that most of the current levels of FDI in developing countries are derived from firm acquisitions and not greenfield investment.[17] Another objection, probably more important, is that "cherry picking" may affect the comparison between acquired and non-acquired domestic firms. This is examined in the next section.

[insert Tables 8a and 8b about here]

3.5. Acquisitions – Difference-in-differences and Propensity Score Matching

In this section, we combine the methods of the two previous sections. As mentioned before, this is motivated by Heckman et al (1997) that found that a difference-in-differences extension of the method of matching is particularly effective in eliminating bias, especially that due to time-invariant unobserved variables. In this case, that of domestic firms acquired by multinationals, there may be an important dimension of "cherry picking" over both observed and unobserved variables. However, the direction of the bias is not clear, as some authors argue that the asymmetry of information inherent in the acquisition of firms may be particularly acute for foreign firms, leading them to pick "lemons" rather than "cherries" (see Gioia and Thomsen, 2003).

To the extent that the unobserved variables are time-invariant, the combined method used here will provide unbiased estimates of the impact of foreign ownership upon wages. To the more likely extent that some of those variables are not time-invariant but are picked up by the observable variables used here, our approach will reduce the amount of bias.

For the benefit of robustness, we consider as before the same two control groups (always-domestic and always-foreign firms) and two treatment groups (domestic firms that become foreign-owned and the subset of firms in which foreign parties have a positive level of ownership in the first period).

---

[17] In Portugal, over the period 1996-2001, acquisitions (and equity increases) took 39% of net inward FDI, while greenfield investment (establishment of new firms) represented only 2% (Ministério das Finanças (2003, p. 17). The remaining share concerns credits and reinvested profits.

In the first two cases, with the wider treatment group, we find negative estimates (see Table 9): –1.5 p.p. and –7.8 p.p. for the always domestic and the always foreign control groups, respectively. We find, however, a positive estimate of 3.9 p.p. in the third case, with the restricted treatment group and the control group of always-domestic firms but we do not attach too much relevance to this result given the small and particularly unrepresentative sample used

Overall, we conclude that the wage premium of foreign multinationals is, in the case of Portugal, most likely due to compositional effects and other biases. After one takes these into account, the premium is of a negligible magnitude and, in some cases, even negative. Our evidence goes against the case that there is a causal link between foreign ownership and wages.

In current work (Martins, 2003c and 2004), we examine this causal link from additional approaches – rent sharing and instrumental variables – and still find similar, i.e. insignificant, results. In particular, we do not find any differences in the extent of rent sharing between domestic and foreign firms. Moreover, when instrumenting the foreign-ownership dummy in a wage equation (using agglomeration effects), we also find that the foreign-firm wage premium becomes insignificant with respect to the OLS case.

[insert Table 9 about here]

4. Wage Spillovers

In this section, we examine the importance of wage externalities that may be exerted upon domestic firms by the presence of foreign firms. In practical terms, we first aggregate the wage equations described in Section 4.1 into the firm level and restrict the analysis to domestic firms. Finally, these wage equations are augmented with the employment levels of foreign and domestic firms in the three- or four-digit industry of each domestic firm. The coefficient of the variable measuring the size of foreign firms can then be interpreted as capturing the level of the externality associated with foreign firms.

One important problem, neglected in some of the literature, is that the distribution of multinational presence across industries may be influenced by unobserved factors that also influence overall pay determination, i.e. multinational presence may be endogenous. Many examples would be consistent with this hypothesis. For instance, foreign firms may prefer industries where host-country firms already pay higher wages because, for instance, such industries are undergoing a period of expansion and face upward-sloped labour supply curves. Such industries may also draw disproportionately more on workers that are more skilled along unobservable dimensions and which demand higher premia to be attracted away from their previous employers. The industries targeted by foreign multinationals may also benefit from rents (due to economies of scale, for instance). If rent-sharing is a common practice in the labour market, then those industries will also pay higher wages even without foreign presence.

In all these cases, the externality coefficient would be spuriously inflated should one not control for endogeneity. However, the direction of the bias is unclear: foreign investment may also target sectors that pay lower wages because, for instance, they are undergoing difficult economic conditions and offer good opportunity for acquisitions. This

would generate a negative correlation between the foreign participation variable and the error term that would bias downward the spillover effect.

This matter is looked at, in a first instance, through the inclusion of fixed effects for firms. In this way, the role of foreign presence is estimated from the relationship between changes in foreign presence and changes in wages. This method allows one to control for unobserved time-invariant factors that may simultaneously influence wage determination and foreign presence. Given the previous examples, such factors would be whether an industry is expanding or not, the unobservable skills of the work force or the existence of rents, respectively.

As a further control for this endogeneity problem, current foreign presence in a given industry cell is also instrumented. Indeed, as suggested before, changes in foreign presence may be related to changes in unobservables that drive wages. The instruments used here are the (first) lagged values of current foreign shares. The motivation for this choice is that one may argue that while lagged foreign shares are correlated with current foreign shares, lagged shares should not enter directly the wage equation as current shares are the ones that matter as far as wage determination is concerned. Under this view, lagged shares satisfy the requirements of a valid instrument.

4.1. Methodology

Unlike in Section 3, here we use worker data aggregated at the firm level. This is due to the larger size of the data set, which is now examined as a panel, and to the inherent computational constraints. This new data set amounts to more than 140,000 domestic firms-year. Moreover, we compute industry-year cells with the level of employment of the foreign firms available in the data set and also with the domestic level of employment in each industry-year combination. Two definitions of industries are also taken into account, at a level of aggregation of three- and four-digits, both of which are based on the same two-digit, nine-industry classification as before.

The wage equation considered corresponds to the one presented in (1), but now aggregated at the firm level and augmented with controls for industry employment. The equation is as follows:

$$y_{it} = X_{it}'\beta_1 + F_{it}'\beta_2 + \beta_3 DomesticEmpl_{ind(i)} + \beta_4 ForeignEmpl_{ind(i)} + \lambda_i + \sigma_t + \varepsilon_{it}, \quad (3)$$

where y_{it} denotes the logarithm of the average hourly wage paid by firm i in year t, X_{it} denotes the average characteristics of the workforce of firm i in year t, F_{it} denotes the average characteristics of the firm i itself in year t, $DomesticEmpl_{ind(i)}$ denotes the logarithm of the domestic employment level of the industry of firm i (defined at either the three- or four-digit level), $ForeignEmpl_{ind(i)}$ denotes the logarithm of the foreign employment level of the industry of firm i, λ_i is a firm fixed effect, σ_t are the year fixed effects and ε_{it} is the standard random error term.

Our motivation for controlling simultaneously for domestic and foreign employment in a given industry is based on Castellani and Zanfei (2002, 2003). From the perspective of the productivity spillovers literature, these authors show that it may be too restrictive to impose no impact from the growth of a given industry and only from the *share*

of foreign employment in total employment, the variable typically used in previous research.

Special attention is placed on the β_4 coefficient, which measures the foreign firm spillover on the wages of workers of domestic firms. For reasons explained above, this equation is estimated with firm fixed effects, first disregarding the endogeneity of the foreign employment variable and then instrumenting it, using foreign employment lagged values.

Another issue is that because we draw on a sample of 80% of all firms, there is probably extra variability in the industry size variables that is not related to the normal entry and exit of firms to each industry. The measurement error problem that this will bring (which is compounded with panel data – see Griliches and Hausman, 1986) may attenuate the estimates of the spillover effects. We try to correct this using new measures of the domestic and foreign employment per industry, obtained by interpolating the employment level of firms that are not sampled in intermediate years. For instance, if a firm is available in our data set in 1993 and 1995, but not in 1994, then we assign to the latter year the average employment level of 1993 and 1995.

4.2. Results

The first four columns of Table 10 present the first set of results, in which firm fixed effects are considered but the possible endogeneity of foreign presence is ignored. The first two columns consider three-digit industries, and the not-adjusted and adjusted industry size cells, respectively. The foreign-employment spillover coefficients are insignificant if foreign employment is not adjusted but become significant (marginally at the 10% and at the 5% levels, respectively) at the 3- and 4-digit level of aggregation, when foreign employment is adjusted for measurement error. However, even in the last two cases, the point estimates are very low, at .4% or .2%.

As mentioned before, we attach greater importance to the estimation method that instruments for foreign shares, as this allows one to draw on exogenous variation to identify the spillover effect. In this case, presented in the last four columns of Table 10, we find a more consistent picture, in which all foreign-share coefficients are positive and highly significant, ranging between 1.9% and 2.7%, or about 10 to 30 times bigger than those without instrumentation.[18] Finally, we also find higher elasticities for more disaggregated industries. This suggests that the impact of foreign presence fades away as we consider a progressively broader set of firms.[19]

Taking these estimates at face value, one can work out back-of-the-envelope monetary impacts of the spillovers. These figures may have some policy relevance as they can be regarded as a benchmark for how much public authorities may spend in attracting multinationals. Considering the case of three-digit industries, not-adjusted measurement, and an average gross monthly wage of €628 for workers in domestic firms, we obtain a wage increase of €12 per worker per month. We then multiply this figure by 14 months and

---

[18] This increase after instrumentation is consistent with measurement error and/or the case in which foreign firms locate in specific industries that, because of their own unobserved characteristics, pay lower wages.

[19] In the models that deal with endogeneity, we also find negative elasticities of wages with respect to domestic employment in the industry. This is consistent with a process of rent sharing, whereby as the industry expands and becomes more competitive, rents decrease, forcing wages to decrease (or to grow by less).

by the average number of domestic workers per industry (19,459) and divide it by the average number of workers in foreign firms per industry (4,443). This results in a spillover of €728 per new job created by a foreign multinational. Considering the case of four-digit industries, the same spillover per job is of €1,002. (There is an average of 6,882 workers in domestic firms and 1,492 workers in foreign firms in each sector defined at this level of aggregation.) In the case of adjusted measurement, we obtain benchmark values of €832 and €1,073, respectively.

[insert Table 10 about here]

5. Summary and Conclusions

Large foreign direct investment (FDI) inflows have become an important feature of the Portuguese economy, representing up to 23% of GDP in 2000. These massive flows were not however without a cost, as expensive government policies were developed in order to attract multinationals to Portugal. In a global context in which the attraction of such flows is likely to become even more difficult, and in a national context in which the future direction of the Portuguese economy is under debate, it is important to assess the benefits to the Portuguese economy of FDI.

This paper provides some evidence on this topic from the perspective of the Portuguese labour market and of wages in particular. This approach contrasts with that taken in most related research, which focuses on productivity. Our motivation for this alternative method lies on the importance of wages to people's welfare. In fact, if, for some reason, FDI-related increases in productivity do not trickle down to a large share of the population via higher wages, one can argue, at least on equity grounds, that the contribution of FDI to a country's economy is not that important.

To this extent, we consider that the role of foreign firms is more relevant if 1) workers of such firms are better paid than their equivalent counterparts in domestic firms and/or 2) workers of domestic firms are better paid as a consequence of foreign presence in their industry. These are the two specific impacts of foreign firms on the Portuguese labour market that are examined in this paper.

With respect to the first aspect, that of the wage differentials between foreign and domestic firms, we conclude that they are difficult to reconcile with a causal relationship. When using standard methods (OLS), we find, in line with most of the international literature, that foreign firms pay more – an average premium of about 32%. However, even when considering this simple method, the addition of further control variables – those related to worker characteristics and, in particular, firm characteristics – substantially decreases the pay premium to an average of about 11%. Moreover, no evidence is found of a positive relationship between the degree of foreign control of a firm and the wage premium of such firm. However, should foreign control have a positive causal effect upon pay, a positive relationship between such control and pay would be expected.

Further evidence on this was obtained with the Propensity Score Matching (PSM) method. This is a non-parametric estimator that assumes no selection on unobservables and which works by performing a detailed matching of observations from the two groups being compared: workers in domestic and foreign firms, in our case. Although the assumption on unobservables is probably too strong, PSM improves on OLS by allowing for better like-

for-like comparisons. This possibility is especially important in our data, since they provide a rich set of matching variables, both at the worker and at the firm level, and many observable differences between domestic and foreign firms are documented. Using PSM we find, on average, a foreign wage premium of only 1%.

Another complementary approach used is a difference-in-differences (DID) estimator based on the foreign acquisition of domestic firms. This allows one to control for unobservable effects that may vary between workers in domestic and foreign firms and which would therefore undermine a causal interpretation of the wage differentials reported before. Under this approach, the wage differential becomes in some specifications even negative, at –4 and –3 percentage points (p.p.), with an upper bound of 2 p.p. in the less reliable case. An explanation for these differences is that domestic firms that undergo this "treatment" may be very different from domestic firms that are not subject to a foreign acquisition. This matter was looked at by combining DID and PSM, thus allowing for a better comparison of treated and not treated firms. However, we still find negative figures in the main cases, which overall range between –8 p.p. and 4 p.p.

The second major aspect examined in this paper concerns the size of spillovers from foreign to domestic firms in terms of the wages paid to workers of the latter group. In fixed effects specifications, in which the impact of foreign presence on domestic firms' wages is derived from the time differences of the variables, combined with the instrumentation of foreign presence, one finds statistically significant and economically large elasticities.

These elasticities range between 1.9% and 2.7%, increasing as we narrow the definition of industries, which suggests that foreign presence has a greater impact in domestic firms that produce similar goods. Moreover, these estimates allow for back-of-the-envelope calculations of the monetary impacts of these spillovers, which may be taken into account by policy-makers as benchmarks when deciding on how much to spend when attracting multinationals. These estimates range between €730 and €1,070 per new job created by a foreign multinational.

Overall, our best evidence on wage differentials and on wage spillovers of foreign firms suggests that while differentials are explainable by the different characteristics of foreign firms, spillovers are of a significant magnitude. In the first case, the lack of a wage premium when domestic and foreign firms are better matched suggests that other firm characteristics than "foreigness" drive the differential. In other words, workers would benefit from more firms with the characteristics of multinationals, regardless of their nationality. In the second case, the significant wage spillovers suggest that workers in domestic firms would benefit from a larger presence of foreign firms in their industries.

In terms of policy implications, the findings presented in this paper provide encouraging support for the role that foreign direct investment can play in the development of the Portuguese economy and in the welfare of its citizens. However, and as emphasised before, one should underline that the impacts of FDI are multi-dimensional, while in this analysis only the wage aspect was addressed. Further studies can complement the one presented here. In any case, the results in this paper do mean that the case for FDI as a mechanism for economic development in Portugal can draw on supporting empirical evidence.

Table 2a (cont.)

Descriptive Statistics, Workers in Foreign and Domestic Firms, 1991-1999

		1991		1992		1993		1994		1995	
	Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Foreign	Monthly Earnings	690.672	565.65	730.441	610.22	801.833	707.12	767.412	676.25	801.801	710.91
Firms	Hourly Earnings	4.057	4.40	4.410	10.53	4.655	4.93	4.564	4.75	4.731	5.01
	Log Hourly Earnings	1.194	0.57	1.263	0.58	1.307	0.61	1.294	0.59	1.342	0.58
	Schooling Years	6.206	3.10	6.350	3.19	6.417	3.29	6.779	3.18	6.958	3.19
	Experience	21.713	11.77	21.628	12.05	22.450	11.66	20.537	11.43	20.725	11.31
	Tenure (Months)	115.027	104.89	115.957	108.70	121.226	106.47	99.074	100.01	103.560	101.68
	Female	0.561		0.550		0.538		0.576		0.568	
	Food, Beverages	0.061		0.044		0.068		0.111		0.093	
	Textiles, Clothing, Leather	0.351		0.331		0.354		0.291		0.268	
	Wood, Cork	0.013		0.012		0.022		0.016		0.008	
	Paper, Graphical Arts, Edition	0.038		0.035		0.032		0.022		0.024	
	Chemicals	0.152		0.157		0.130		0.111		0.082	
	Non-Metalic Minerals	0.020		0.040		0.035		0.032		0.042	
	Metalurgic	0.009		0.009		0.007		0.008		0.012	
	Metalic, Transport	0.347		0.369		0.351		0.407		0.469	
	Others	0.008		0.003		0.002		0.002		0.002	
	North	0.447		0.507		0.472		0.350		0.364	
	Centre	0.100		0.096		0.109		0.241		0.193	
	Lisboa and T. V.	0.434		0.390		0.416		0.404		0.438	
	Observations	52.485		41.671		53.432		57.843		73.795	
Domestic	Monthly Earnings	528.463	419.95	566.039	495.42	582.574	537.62	575.465	1322.57	562.403	477.85
Firms	Hourly Earnings	3.118	4.12	3.309	4.50	3.399	4.42	3.438	11.71	3.342	4.30

Table 2a (cont.)

Descriptive Statistics, Workers in Foreign and Domestic Firms, 1991-1999

	1991		1992		1993		1994		1995	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Log Hourly Earnings	0.954	0.52	1.005	0.53	1.023	0.54	1.024	0.54	1.026	0.52
Schooling Years	5.260	2.67	5.353	2.71	5.481	2.76	5.735	2.78	5.758	2.74
Experience	23.510	12.78	23.467	12.81	23.901	12.72	23.235	12.46	23.496	12.38
Tenure (Months)	116.348	108.02	116.343	107.76	120.108	108.50	110.819	104.92	114.901	108.29
Female	0.423		0.430		0.417		0.431		0.434	
Food, Beverages	0.102		0.105		0.109		0.113		0.107	
Textiles, Clothing, Leather	0.401		0.399		0.375		0.384		0.391	
Wood, Cork	0.073		0.076		0.080		0.084		0.082	
Paper, Graphical Arts, Edition	0.053		0.055		0.060		0.060		0.056	
Chemicals	0.070		0.057		0.065		0.063		0.059	
Non-Metalic Minerals	0.087		0.080		0.080		0.077		0.083	
Metalurgic	0.028		0.025		0.025		0.016		0.016	
Metalic, Transport	0.175		0.193		0.197		0.192		0.196	
Others	0.010		0.009		0.010		0.010		0.010	
North	0.511		0.513		0.499		0.513		0.540	
Centre	0.215		0.207		0.224		0.227		0.223	
Lisboa and T. V.	0.253		0.257		0.254		0.232		0.211	
Observations	471.745		475.336		441.029		404.563		427.461	

Notes:

Monetary values are in real terms (2000 prices) and in euros.

Alentejo and Algarve regional dummies are omitted (their share ranges between .0% and 3.8%).

Table 2b (cont.)

Descriptive Statistics, Workers in Foreign and Domestic Firms, 1991-1999

		1996		1997		1998		1999		Annual
	Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Average
Foreign	Monthly Earnings	806.128	763.80	841.352	721.85	841.777	713.32	863.091	756.85	793.8
Firms	Hourly Earnings	4.728	7.80	4.837	4.42	5.001	4.58	5.147	5.56	4.7
	Log Hourly Earnings	1.340	0.57	1.386	0.55	1.426	0.54	1.446	0.55	1.3
	Schooling Years	7.196	3.31	7.255	3.32	7.221	3.27	7.367	3.36	6.9
	Experience	20.486	11.51	20.635	11.45	21.348	11.62	21.290	11.58	21.2
	Tenure (Months)	105.067	102.62	106.347	103.00	111.682	104.88	106.749	105.78	109.4
	Female	0.534		0.516		0.548		0.552		0.55
	Food, Beverages	0.105		0.075		0.077		0.066		0.08
	Textiles, Clothing, Leather	0.258		0.262		0.267		0.234		0.29
	Wood, Cork	0.006		0.008		0.016		0.023		0.01
	Paper, Graphical Arts, Edition	0.021		0.011		0.024		0.024		0.03
	Chemicals	0.070		0.073		0.116		0.085		0.11
	Non-Metalic Minerals	0.045		0.036		0.039		0.036		0.04
	Metalurgic	0.016		0.023		0.018		0.011		0.01
	Metalic, Transport	0.476		0.510		0.441		0.521		0.43
	Others	0.002		0.002		0.002		0.001		0.00
	North	0.360		0.325		0.382		0.398		0.40
	Centre	0.209		0.229		0.254		0.264		0.19
	Lisboa and T. V.	0.423		0.434		0.324		0.296		0.40
	Observations	72.561		74.804		67.971		68.837		62.600
Domestic	Monthly Earnings	603.105	526.74	593.721	509.92	610.191	506.39	627.803	520.11	583.3
Firms	Hourly Earnings	3.519	4.96	3.536	5.00	3.708	3.78	3.854	4.13	3.5

Table 2b (cont.)

Descriptive Statistics, Workers in Foreign and Domestic Firms, 1991-1999

	1996		1997		1998		1999		Annual
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Average
Log Hourly Earnings	1.066	0.53	1.091	0.50	1.150	0.49	1.188	0.49	1.1
Schooling Years	5.922	2.87	5.958	2.84	6.124	2.95	6.206	2.98	5.8
Experience	23.936	12.34	23.884	12.39	24.328	12.45	24.528	12.39	23.8
Tenure (Months)	120.348	110.83	116.409	110.41	117.825	112.59	118.600	113.08	116.9
Female	0.435		0.449		0.436		0.442		0.43
Food, Beverages	0.112		0.101		0.106		0.110		0.11
Textiles, Clothing, Leather	0.391		0.389		0.363		0.362		0.38
Wood, Cork	0.075		0.092		0.089		0.088		0.08
Paper, Graphical Arts, Edition	0.061		0.066		0.063		0.062		0.06
Chemicals	0.064		0.055		0.062		0.065		0.06
Non-Metalic Minerals	0.082		0.082		0.090		0.085		0.08
Metalurgic	0.017		0.018		0.016		0.017		0.02
Metalic, Transport	0.188		0.186		0.202		0.203		0.19
Others	0.010		0.010		0.009		0.008		0.01
North	0.537		0.548		0.524		0.533		0.52
Centre	0.222		0.225		0.236		0.225		0.22
Lisboa and T. V.	0.214		0.201		0.211		0.216		0.23
Observations	412.961		428.839		409.721		434.319		433.997

Notes:

Monetary values are in real terms (2000 prices) and in euros.

Alentejo and Algarve regional dummies are omitted (their share ranges between .0% and 3.8%).

Table 3a (cont.)

Descriptive Statistics, Foreign and Domestic Firms, 1991-1999

		1991		1992		1993		1994		1995	
	Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Foreign	Monthly Earnings	684.276	422.88	736.970	373.94	776.020	457.39	807.893	474.84	837.143	512.65
Firms	Hourly Earnings	4.046	2.58	4.374	2.36	4.573	2.75	4.801	2.96	4.922	3.08
	Log Hourly Earnings	1.167	0.45	1.246	0.46	1.262	0.48	1.312	0.47	1.337	0.48
	Number of Workers	181.324	305.52	180.021	304.82	181.115	341.77	169.094	397.01	200.867	450.71
	Food, Beverages	0.080		0.088		0.089		0.101		0.111	
	Textiles, Clothing, Leather	0.291		0.271		0.268		0.251		0.259	
	Wood, Cork	0.050		0.039		0.061		0.054		0.047	
	Paper, Graphical Arts, Edition	0.050		0.046		0.042		0.054		0.049	
	Chemicals	0.180		0.180		0.182		0.170		0.143	
	Non-Metalic Minerals	0.047		0.063		0.059		0.049		0.057	
	Metalurgic	0.017		0.018		0.020		0.022		0.022	
	Metalic, Transport	0.269		0.282		0.268		0.288		0.301	
	Others	0.017		0.014		0.011		0.010		0.010	
	North	0.374		0.377		0.369		0.367		0.383	
	Centre	0.161		0.165		0.193		0.200		0.193	
	Lisboa and T. V.	0.438		0.426		0.425		0.411		0.402	
	Observations	361		284		358		406		405	
Domestic	Monthly Earnings	413.133	177.60	438.727	201.13	453.590	229.04	454.539	241.90	453.567	207.19
Firms	Hourly Earnings	2.424	1.48	2.558	1.40	2.633	1.49	2.717	2.28	2.715	1.58
	Log Hourly Earnings	0.763	0.31	0.808	0.32	0.827	0.33	0.852	0.34	0.869	0.33

Table 3a (cont.)

Descriptive Statistics, Foreign and Domestic Firms, 1991-1999

	1991		1992		1993		1994		1995	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Number of Workers	39.999	118.25	39.516	116.94	35.779	103.97	31.582	85.73	30.235	77.01
Food, Beverages	0.100		0.099		0.101		0.107		0.112	
Textiles, Clothing, Leather	0.282		0.280		0.270		0.265		0.264	
Wood, Cork	0.148		0.147		0.144		0.147		0.146	
Paper, Graphical Arts, Edition	0.078		0.080		0.080		0.085		0.083	
Chemicals	0.055		0.053		0.054		0.052		0.051	
Non-Metalic Minerals	0.079		0.083		0.081		0.079		0.081	
Metalurgic	0.014		0.014		0.015		0.013		0.012	
Metalic, Transport	0.224		0.224		0.235		0.233		0.231	
Others	0.021		0.020		0.020		0.020		0.021	
North	0.487		0.482		0.477		0.489		0.496	
Centre	0.220		0.222		0.227		0.226		0.222	
Lisboa and T. V.	0.253		0.255		0.256		0.244		0.240	
Observations	15.095		15.278		15.351		16.171		16.757	

Notes:

Monetary values are in real terms (2000 prices) and in euros.

Alentejo and Algarve regional dummies are omitted (their share ranges between .0% and 3.8%).

Table 3b

Descriptive Statistics, Foreign and Domestic Firms, 1991-1999

		1996		1997		1998		1999		Annual
	Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Average
Foreign	Monthly Earnings	809.742	406.47	831.480	410.10	874.334	446.85	888.790	481.54	805.2
Firms	Hourly Earnings	4.745	2.55	4.863	2.38	5.203	2.66	5.428	3.07	4.8
	Log Hourly Earnings	1.319	0.45	1.364	0.43	1.437	0.44	1.466	0.44	1.3
	Number of Workers	200.333	487.64	201.334	556.08	196.372	461.98	195.316	456.67	189.5
	Food, Beverages	0.119		0.104		0.112		0.093		0.10
	Textiles, Clothing, Leather	0.258		0.255		0.242		0.225		0.26
	Wood, Cork	0.028		0.045		0.045		0.061		0.05
	Paper, Graphical Arts, Edition	0.051		0.037		0.045		0.037		0.05
	Chemicals	0.157		0.139		0.162		0.154		0.16
	Non-Metalic Minerals	0.063		0.064		0.061		0.054		0.06
	Metalurgic	0.018		0.025		0.021		0.017		0.02
	Metalic, Transport	0.298		0.322		0.303		0.346		0.30
	Others	0.010		0.010		0.008		0.012		0.01
	North	0.389		0.356		0.364		0.380		0.37
	Centre	0.192		0.230		0.237		0.228		0.20
	Lisboa and T. V.	0.391		0.381		0.370		0.353		0.40
	Observations	396		404		376		408		378
Domestic	Monthly Earnings	476.766	234.64	479.462	226.46	493.402	221.24	502.278	221.16	462.8
Firms	Hourly Earnings	2.814	1.66	2.881	1.77	3.031	1.41	3.138	2.45	2.8
	Log Hourly Earnings	0.898	0.33	0.934	0.32	1.004	0.30	1.039	0.30	0.9
	Number of Workers	30.397	84.98	28.035	70.78	27.200	69.77	27.582	82.47	32.3

Table 3b

Descriptive Statistics, Foreign and Domestic Firms, 1991-1999

Variable	1996 Mean	Std. Dev.	1997 Mean	Std. Dev.	1998 Mean	Std. Dev.	1999 Mean	Std. Dev.	Annual Average
Food, Beverages	0.114		0.116		0.118		0.114		0.11
Textiles, Clothing, Leather	0.263		0.260		0.247		0.251		0.26
Wood, Cork	0.139		0.151		0.149		0.152		*0.15*
Paper, Graphical Arts, Edition	0.085		0.083		0.087		0.084		0.08
Chemicals	0.051		0.048		0.048		0.048		*0.05*
Non-Metalic Minerals	0.084		0.081		0.086		0.084		0.08
Metalurgic	0.011		0.013		0.011		0.013		0.01
Metalic, Transport	0.232		0.229		0.236		0.234		*0.23*
Others	0.020		0.019		0.019		0.019		0.02
North	0.484		0.497		0.478		0.497		*0.49*
Centre	0.236		0.231		0.240		0.231		0.23
Lisboa and T. V.	0.237		0.226		0.233		0.227		*0.24*
Observations	16.093		17.853		17.596		18.819		16.557

Notes:

Monetary values are in real terms (2000 prices) and in euros.

Alentejo and Algarve regional dummies are omitted (their share ranges between .0% and 3.8%).

Table 4

Wage Premia of Foreign Firms, 1991-1999.

Specifications - Controls	1991	1992	1993	1994	1995	1996	1997	1998	1999	Avg.	Max.	Min.
No Controls	27.1%	29.4%	32.8%	31.0%	37.2%	31.5%	34.3%	31.8%	29.4%	31.6%	37.2%	27.1%
St. Errors	[0.002]**	[0.003]**	[0.003]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**			
Adjusted R-squared	0.019	0.017	0.025	0.026	0.043	0.032	0.041	0.036	0.031			
Human Capital (HC)	23.1%	24.5%	26.6%	28.4%	32.3%	26.5%	26.9%	27.9%	25.7%	26.9%	32.3%	23.1%
St. Errors	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**			
Adjusted R-squared	0.451	0.436	0.45	0.438	0.453	0.453	0.445	0.477	0.477			
HC and Firm Characteristics	11.6%	9.6%	11.3%	10.5%	12.2%	9.0%	8.0%	12.6%	13.0%	10.9%	13.0%	8.0%
St. Errors	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**			
Adjusted R-squared	0.544	0.536	0.542	0.524	0.534	0.534	0.516	0.549	0.546			
Observations	524.230	517.007	494.461	462.406	501.256	485.522	503.643	477.692	503.156			

Notes:

Standard errors in brackets (* significant at 5%; ** significant at 1%).

Human Capital controls are: six dummies for educational degrees, a quartic in experience, a quadratic in tenure (in months), and a female dummy.

Firm Characteristics are: four regional dummies, log number of workers, a dummy for public firms and eight industry dummies.

Table 5 (cont.)

Wage Premia of Foreign Firms in Different Industries, 1991-1999.

Specifications	1991	1992	1993	1994	1995	1996	1997	1998	1999	Avg.	Max.	Min.
Food, Beverages	6.5%	8.4%	9.2%	10.1%	10.1%	9.7%	6.9%	9.7%	8.5%	8.8%	10.1%	6.5%
	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**			
Wood, Cork	-2.7%	-0.6%	-2.2%	1.4%	2.6%	4.8%	1.7%	2.3%	1.5%	1.0%	4.8%	-2.7%
	[0.002]**	[0.002]*	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**			
Paper, Graphical Arts	28.7%	21.0%	24.4%	26.5%	29.8%	29.8%	25.7%	23.7%	25.2%	26.1%	29.8%	21.0%
	[0.003]**	[0.003]**	[0.003]**	[0.003]**	[0.003]**	[0.003]**	[0.003]**	[0.002]**	[0.002]**			
Chemicals	29.0%	26.7%	29.7%	28.0%	28.9%	27.4%	23.1%	26.1%	24.5%	27.1%	29.7%	23.1%
	[0.002]**	[0.003]**	[0.003]**	[0.003]**	[0.003]**	[0.003]**	[0.003]**	[0.002]**	[0.002]**			
Non-Metalic Minerals	19.8%	26.4%	22.5%	23.9%	26.5%	29.6%	20.4%	20.1%	19.8%	23.2%	29.6%	19.8%
	[0.002]**	[0.002]**	[0.002]**	[0.003]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**			
Metalurgic	3.5%	14.7%	13.0%	22.9%	23.5%	29.3%	19.0%	22.9%	22.9%	19.1%	29.3%	3.5%
	[0.003]**	[0.004]**	[0.004]**	[0.005]**	[0.005]**	[0.005]**	[0.004]**	[0.004]**	[0.004]**			
Metalic, Transport	13.1%	15.5%	15.7%	21.3%	20.6%	21.8%	17.0%	15.7%	13.8%	17.2%	21.8%	13.1%
	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**			
Others	7.5%	12.1%	10.7%	12.3%	12.6%	10.0%	6.1%	6.2%	7.4%	9.4%	12.6%	6.1%
	[0.005]**	[0.006]**	[0.006]**	[0.006]**	[0.006]**	[0.006]**	[0.006]**	[0.006]**	[0.006]**			
Foreign Firm	5.4%	7.8%	4.8%	3.5%	5.9%	3.9%	1.9%	4.8%	3.9%	4.6%	7.8%	1.9%
	[0.003]**	[0.003]**	[0.003]**	[0.003]**	[0.003]**	[0.003]**	[0.003]**	[0.003]**	[0.003]**			
(Food, Beverages)*For.	42.5%	11.4%	22.6%	15.0%	22.6%	25.0%	32.3%	28.4%	30.3%	25.6%	42.5%	11.4%
	[0.007]**	[0.009]**	[0.007]**	[0.006]**	[0.006]**	[0.005]**	[0.006]**	[0.006]**	[0.006]**			
(Wood, Cork)*For.	28.0%	22.5%	7.0%	24.7%	15.0%	18.3%	19.6%	16.5%	14.8%	18.5%	28.0%	7.0%
	[0.014]**	[0.017]**	[0.012]**	[0.013]**	[0.016]**	[0.018]**	[0.015]**	[0.011]**	[0.009]**			

Table 5 (cont.)

Wage Premia of Foreign Firms in Different Industries, 1991-1999.

Specifications	1991	1992	1993	1994	1995	1996	1997	1998	1999	Avg.	Max.	Min.
(Paper, Graphical Arts)*For.	6.9%	14.3%	16.6%	27.4%	29.4%	18.8%	8.5%	14.7%	26.6%	18.1%	29.4%	6.9%
	[0.009]**	[0.010]**	[0.010]**	[0.011]**	[0.009]**	[0.010]**	[0.013]**	[0.009]**	[0.009]**			
Chemicals*For.	11.0%	16.5%	18.8%	19.8%	11.2%	17.5%	24.0%	16.3%	21.9%	17.4%	24.0%	11.0%
	[0.005]**	[0.006]**	[0.006]**	[0.006]**	[0.006]**	[0.006]**	[0.006]**	[0.005]**	[0.006]**			
(Non-Metalic Minerals)*For	8.8%	2.6%	0.0%	10.8%	5.4%	8.1%	13.8%	9.0%	12.2%	7.9%	13.8%	0.0%
	[0.011]**	[0.010]**	[0.009]	[0.010]**	[0.007]**	[0.007]**	[0.008]**	[0.007]**	[0.008]**			
Metalurgic*For.	31.5%	-11.6%	13.4%	-2.3%	-7.8%	-8.4%	4.0%	-0.9%	16.0%	3.8%	31.5%	-11.6%
	[0.017]**	[0.020]**	[0.021]**	[0.018]	[0.013]**	[0.012]**	[0.010]**	[0.011]	[0.013]**			
(Metalic, Transport)*For.	1.9%	-5.2%	4.9%	3.9%	4.0%	-0.4%	0.2%	5.8%	6.3%	2.4%	6.3%	-5.2%
	[0.004]**	[0.005]**	[0.004]**	[0.004]**	[0.004]**	[0.004]	[0.004]	[0.004]**	[0.004]**			
Others*For.	5.3%	-8.2%	4.7%	5.2%	4.7%	14.6%	9.1%	14.1%	14.8%	7.1%	14.8%	-8.2%
	[0.019]**	[0.031]**	[0.033]	[0.035]	[0.033]	[0.034]**	[0.032]**	[0.028]**	[0.037]**			
Adjusted R-squared	0.546	0.538	0.543	0.525	0.536	0.537	0.519	0.551	0.548			
Observations	524.230	517.007	494.461	462.406	501.256	485.522	503.643	477.692	503.156			

Notes:

Standard errors in brackets (* significant at 5%; ** significant at 1%).

Specification used includes controls for human capital and firm characteristics (see previous tables for definitions).

Table 6a

Wage Premia of Foreign Firms for Different Degrees of Control, 1991-1999.

Specifications	1991	1992	1993	1994	1995	1996	1997	1998	1999	Avg.	Max.	Min.
1%-9%	13,8%	9,6%	9,5%	16,1%	13,9%	9,7%	9,6%	2,5%	20,7%	11,7%	20,7%	2,5%
	[0.006]**	[0.006]**	[0.005]**	[0.006]**	[0.005]**	[0.006]**	[0.007]**	[0.010]*	[0.007]**			
10%-19%	17,2%	-2,2%	-5,3%	22,9%	37,2%	28,1%	25,9%	15,6%	-3,4%	15,1%	37,2%	-5,3%
	[0.008]**	[0.013]	[0.010]**	[0.008]**	[0.007]**	[0.007]**	[0.008]**	[0.007]**	[0.010]**			
20%-29%	18,1%	6,4%	3,9%	4,6%	-1,8%	14,0%	3,6%	-4,3%	-2,2%	4,7%	18,1%	-4,3%
	[0.004]**	[0.005]**	[0.006]**	[0.006]**	[0.006]**	[0.006]**	[0.006]**	[0.005]**	[0.006]**			
30%-39%	19,7%	-3,4%	-4,1%	18,9%	12,0%	0,0%	9,2%	11,9%	13,9%	8,7%	19,7%	-4,1%
	[0.007]**	[0.008]**	[0.008]**	[0.008]**	[0.008]**	[0.008]	[0.007]**	[0.007]**	[0.008]**			
40%-49%	22,5%	27,5%	19,0%	7,9%	17,2%	11,4%	13,7%	14,0%	15,0%	16,5%	27,5%	7,9%
	[0.006]**	[0.007]**	[0.007]**	[0.007]**	[0.007]**	[0.006]**	[0.005]**	[0.006]**	[0.005]**			
50%-59%	18,6%	6,5%	6,6%	11,3%	8,3%	8,5%	4,1%	-1,2%	9,7%	8,1%	18,6%	-1,2%
	[0.006]**	[0.008]**	[0.006]**	[0.007]**	[0.007]**	[0.007]**	[0.008]**	[0.007]	[0.006]**			
60%-69%	2,3%	17,2%	16,6%	40,8%	36,3%	10,4%	36,5%	36,5%	24,6%	24,6%	40,8%	2,3%
	[0.010]*	[0.007]**	[0.005]**	[0.006]**	[0.006]**	[0.008]**	[0.011]**	[0.009]**	[0.009]**			
70%-79%	2,1%	0,5%	21,2%	38,3%	26,9%	25,0%	6,2%	23,6%	-5,4%	15,4%	38,3%	-5,4%
	[0.009]*	[0.011]	[0.009]**	[0.010]**	[0.007]**	[0.008]**	[0.009]**	[0.009]**	[0.008]**			
80%-89%	7,6%	10,5%	10,5%	12,9%	12,2%	10,5%	17,4%	14,6%	14,3%	12,3%	17,4%	7,6%
	[0.005]**	[0.008]**	[0.010]**	[0.009]**	[0.008]**	[0.008]**	[0.006]**	[0.006]**	[0.007]**			
90%-99%	11,5%	8,0%	15,6%	10,6%	10,3%	15,5%	13,9%	10,8%	17,7%	12,7%	17,7%	8,0%
	[0.004]**	[0.004]**	[0.004]**	[0.004]**	[0.003]**	[0.003]**	[0.003]**	[0.003]**	[0.004]**			
100%	15,8%	11,2%	10,4%	7,9%	12,6%	7,8%	6,7%	13,7%	13,3%	11,0%	15,8%	6,7%
	[0.002]**	[0.003]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**			
Adjusted R^2	0,548	0,538	0,543	0,528	0,538	0,537	0,518	0,551	0,548			
Observations	524.230	517.007	494.461	462.406	501.256	485.522	503.643	477.692	503.156			

Notes:

Standard errors in brackets (* significant at 5%; ** significant at 1%).

Specification used includes controls for human capital and firm characteristics (see previous tables for definitions).

Table 6b

Wage Premia for Foreign Firms, Different Definitions, 1991-1999

	1991		1992		1993		1994		1995		1996		1997		1998		1999	
	Premium	St. Error	Premium	St. Error	Premium	St. Error	Premium	St. Error	Premium	St. Error	Premium	St. Error	Premium	St. Error	Premium	St. Error	Premium	St. Error
>=10%	14,8%	(0.002)**	9,9%	(0.002)**	10,3%	(0.002)**	11,3%	(0.002)**	13,1%	(0.002)**	10,5%	(0.002)**	9,3%	(0.002)**	12,3%	(0.002)**	12,3%	(0.001)**
>=20%	14,7%	(0.002)**	10,1%	(0.002)**	10,7%	(0.002)**	10,5%	(0.002)**	11,7%	(0.002)**	9,4%	(0.002)**	8,4%	(0.002)**	12,0%	(0.002)**	12,7%	(0.002)**
>=30%	13,0%	(0.002)**	10,2%	(0.002)**	11,2%	(0.002)**	10,8%	(0.002)**	12,7%	(0.002)**	8,9%	(0.002)**	8,7%	(0.002)**	13,1%	(0.002)**	13,4%	(0.002)**
>=40%	12,5%	(0.002)**	11,0%	(0.002)**	12,0%	(0.002)**	10,4%	(0.002)**	12,6%	(0.002)**	9,2%	(0.002)**	8,5%	(0.002)**	13,0%	(0.002)**	13,3%	(0.002)**
>=50%	11,6%	(0.002)**	9,6%	(0.002)**	11,3%	(0.002)**	10,5%	(0.002)**	12,2%	(0.002)**	9,0%	(0.002)**	8,0%	(0.002)**	12,6%	(0.002)**	13,0%	(0.002)**
N. Obs.	524230		517007		494461		462406		501256		485522		503643		477692			

Notes:

Standard errors in parentheses: * significant at 5% level; ** significant at 1% level

The specification considered includes human capital and firm characteristics.

Table 7

Propensity Score Matching estimates

Year/Method	Estimate	St. Error
1991	-0,34%	0,006
1992	-0,02%	0,004
1993	1,59%	0,006
1994	-5,49%	0,008
1995	1,49%	0,003
1996	2,36%	0,006
1997	2,17%	0,002
1998	2,03%	0,005
1999	5,75%	0,007
Average	1,06%	

Note:

Standard errors obtained via bootstrapping (50 repetitions)

Table 8a - Difference-in-differences Descriptive Statistics and Results, I
Control group: workers in firms that are always domestically-owned.

	Treatment1	Treatment2	Control	DID - Treat. 1		DID - Treat. 2	
Variable	**Mean**	**Mean**	**Mean**	**Coeff.**	**St. Error**	**Coeff.**	**St. Error**
Change in Log Hourly Wages	0,006	0,075	0,051				
1st Cycle	0,399	0,322	0,509	-0,014	0,003	-0,012	0,003
2nd Cycle	0,232	0,205	0,263	-0,009	0,003	-0,009	0,003
3rd Cycle	0,170	0,241	0,098	-0,014	0,004	-0,013	0,004
Secondary	0,107	0,150	0,063	-0,009	0,004	-0,013	0,004
"Bacharelato"	0,017	0,021	0,007	0,009	0,008	0,003	0,008
"Licenciatura"	0,040	0,033	0,014	0,001	0,006	-0,008	0,006
Experience	115,247	122,832	127,974	0,000	0,000	0,000	0,000
Tenure (Months)	22,570	22,907	24,350	-0,001	0,000	-0,001	0,000
Female	0,386	0,268	0,426	-0,012	0,001	-0,013	0,001
Log Firm Size	6,080	5,887	4,766	-0,008	0,000	-0,004	0,000
Difference in Years	1,410	1,634	1,332	0,030	0,001	0,023	0,001
Food, Beverages	0,088	0,126	0,109	0,013	0,002	0,011	0,002
Wood, Cork	0,039	0,064	0,081	-0,001	0,003	0,007	0,003
Paper, Graphical Arts, Edition	0,038	0,009	0,051	0,005	0,003	0,005	0,003
Chemicals	0,091	0,147	0,059	-0,026	0,003	-0,029	0,003
Non-Metalic Minerals	0,032	0,001	0,082	0,015	0,003	0,018	0,003
Metalurgic	0,017	0,000	0,015	0,015	0,005	0,008	0,005
Metalic, Transport	0,443	0,630	0,197	0,008	0,002	0,015	0,002
Others	0,001	0,000	0,008	-0,015	0,007	-0,006	0,007
Centre	0,211	0,293	0,229	-0,001	0,002	-0,002	0,002
Lisboa and T. V.	0,408	0,557	0,216	-0,020	0,002	-0,011	0,002
Alentejo	0,033	0,000	0,013	-0,008	0,005	0,000	0,006
Algarve	0,001	0,000	0,007	-0,032	0,007	-0,033	0,007
1991	0,165	0,337	0,148	0,008	0,003	0,005	0,003
1992	0,250	0,047	0,117	-0,059	0,003	-0,033	0,003
1993	0,108	0,200	0,114	-0,042	0,003	-0,039	0,003
1994	0,089	0,100	0,118	-0,024	0,003	-0,021	0,003
1995	0,063	0,000	0,131	-0,022	0,003	-0,018	0,003
1996	0,161	0,031	0,158	0,023	0,002	0,024	0,003
1997	0,069	0,069	0,116	0,001	0,003	0,005	0,003
Treatment				**-0,031**	**0,003**	**0,020**	**0,006**
Observations	18.269	2.578	250.031	268.300		252.609	
Adj. R squared				0,0165		0,0106	

Notes:
Treatment group 1: Workers in firms that are domestic in the first period and foreign in the second.
Treatment group 2: Workers in firms that are domestic in the first period, but with a positive level of foreign ownership, and foreign in the second.
The standard deviations of the change in log wages is 0,358, 0,371 and 0,31 for the first two treatment groups and for the control group respectively.

Table 8b - Difference-in-differences Descriptive Statistics and Results, II
Control group: workers in firms that are always foreign-owned.

	Control	DID - Treat. 1	
Variable	**Mean**	**Coeff.**	**St. Error**
Change in Log Hourly Wages	0,070		
1st Cycle	0,340	-0,038	0,010
2nd Cycle	0,307	-0,027	0,010
3rd Cycle	0,189	-0,031	0,010
Secondary	0,108	-0,019	0,011
"Bacharelato"	0,013	-0,003	0,016
"Licenciatura"	0,023	-0,011	0,013
Experience	111,62	0,000	0,000
Tenure (Months)	20,80	-0,003	0,000
Female	0,56	-0,021	0,003
Log Firm Size	6,38	-0,030	0,001
Difference in Years	1,53	0,074	0,002
Food, Beverages	0,064	-0,021	0,007
Wood, Cork	0,011	-0,108	0,011
Paper, Graphical Arts, Edition	0,013	0,029	0,011
Chemicals	0,128	-0,056	0,006
Non-Metalic Minerals	0,033	-0,044	0,009
Metalurgic	0,007	-0,188	0,011
Metalic, Transport	0,504	-0,034	0,005
Others	0,001	-0,103	0,039
Centre	0,238	0,038	0,005
Lisboa and T. V.	0,380	0,009	0,004
Alentejo	0,004	0,033	0,014
Algarve	0,000	0,289	0,067
1992	0,104	-0,125	0,006
1993	0,139	-0,004	0,006
1994	0,227	0,018	0,006
1995	0,211	0,000	0,006
1996	0,055	0,004	0,007
1997	0,098	-0,020	0,007
1998	0,068	0,002	0,007
Treatment		-0,042	0,003
Observations	27.229	46.476	
Adj. R squared		0,0983	

Notes:
See the definition of the treatment groups in the previous table.
The standard deviations of the change in log wages is 0,3 for the control group.

Table 9

Propensity Score Matching estimates

Control	Treatment	Effect
Always domestic firms	All domestic firms	-1,54%
Always domestic firms	Subset of domestic firms	3,85%
Always foreign firms	All domestic firms	-7,83%

Notes: See main text.

Table 10

Spillovers Effects

	Without Instruments				With Instruments			
	3-digit		4-digit		3-digit		4-digit	
	A	B	A	B	A	B	A	B
Log Domestic Size	0,0184	0,0135	0,0230	0,0116	-0,0464	-0,0479	-0,0460	-0,0473
	0,0067	0,0069	0,0047	0,0048	0,0023	0,0023	0,0020	0,0020
Log Foreign Size	-0,0005	0,0040	-0,0009	0,0023	0,0189	0,0216	0,0247	0,0265
	0,0016	0,0021	0,0011	0,0014	0,0017	0,0019	0,0015	0,0015
Observations	143.420	143.517	143.574	143.574	79.365	79.389	79.418	79.418
Adjusted R-squared	0,741	0,741	0,742	0,741				

Notes:

Standard errors reported below coefficients.

Other controls: Education Dummies, Experience, Tenure, Female Share, Firm Size, Public Ownership, and Year Dummies.

All specifications include firm fixed effects. Robust Standard Errors. Instruments are lagged foreign shares.

A - Foreign size measurement not corrected.

B - Foreign size measurement corrected.

Table A1a - Descriptive Statistics, Workers in Foreign and Domestic Firms, 1991-1999 (10% foreign ownership definition).

		1991		1992		1993		1994		1995	
	Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Foreign Firms	Monthly Earnings	729,233	562,756	748,363	641,801	811,258	696,585	775,501	659,695	817,393	717,845
	Hourly Earnings	4,324	4,434	4,544	9,743	4,699	5,272	4,637	4,698	4,812	4,979
	Log Hourly Earnings	1,264	0,573	1,297	0,576	1,324	0,601	1,318	0,586	1,365	0,576
	Schooling Years	6,211	3,166	6,322	3,186	6,417	3,318	6,780	3,230	6,974	3,261
	Experience	23,153	12,043	22,357	12,151	23,237	11,873	21,370	11,730	21,573	11,700
	Tenure (Months)	126,584	107,267	122,124	111,070	128,755	109,286	105,798	103,571	112,586	107,298
	Female	0,490		0,507		0,493		0,543		0,536	
	North	0,424		0,470		0,453		0,334		0,368	
	Centre	0,113		0,100		0,116		0,227		0,192	
	Lisboa and T. V.	0,450		0,423		0,428		0,435		0,436	
	Observations	70.335		52.362		63.084		69.466		85.323	
Domestic Firms	Monthly Earnings	483,882	360,092	547,771	474,813	552,047	491,848	549,205	1231,43	539,510	448,419
	Hourly Earnings	2,829	3,910	3,188	4,303	3,221	4,131	3,262	11,558	3,193	4,154
	Log Hourly Earnings	0,889	0,465	0,977	0,517	0,984	0,519	0,987	0,513	0,994	0,493
	Schooling Years	5,135	2,525	5,297	2,650	5,387	2,658	5,643	2,668	5,676	2,645
	Experience	22,990	12,809	23,270	12,844	23,584	12,745	23,016	12,463	23,295	12,370
	Tenure (Months)	109,846	106,141	113,417	106,747	115,418	106,809	107,201	103,164	111,715	106,740
	Female	0,444		0,439		0,432		0,440		0,443	
	North	0,542		0,527		0,518		0,532		0,553	
	Centre	0,227		0,214		0,228		0,230		0,224	
	Lisboa and T. V.	0,208		0,235		0,229		0,209		0,196	
	Observations	423.346		446.567		409.669		378.135		403.845	

Notes:
Monetary values are in real terms (2000 prices) and in euros.
Alentejo and Algarve regional dummies are omitted (their share ranges between .0% and 3.8%).

Table A1b

Descriptive Statistics, Workers in Foreign and Domestic Firms, 1991-1999 (10% foreign ownership definition).

		1996		1997		1998		1999		Annual
	Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Average
Foreign	Monthly Earnings	832,928	767,144	860,720	720,861	855,704	725,472	875,588	748,271	811,85
Firms	Hourly Earnings	4,877	7,502	4,955	4,485	5,088	4,644	5,267	6,442	4,80
	Log Hourly Earnings	1,370	0,574	1,411	0,555	1,447	0,538	1,465	0,550	1,36
	Schooling Years	7,180	3,355	7,267	3,375	7,255	3,332	7,369	3,406	6,86
	Experience	21,365	11,859	21,470	11,844	22,107	11,964	22,016	11,854	22,07
	Tenure (Months)	113,653	107,760	114,099	108,326	120,692	111,406	114,733	111,051	117,67
	Female	0,502		0,487		0,512		0,516		0,51
	North	0,364		0,334		0,389		0,393		0,39
	Centre	0,206		0,228		0,253		0,256		0,19
	Lisboa and T. V.	0,423		0,428		0,323		0,314		0,41
	Observations	84.861		88.031		80.615		79.459		74.837
Domestic	Monthly Earnings	571,277	484,371	568,369	487,239	580,210	464,515	602,491	488,942	554,97
Firms	Hourly Earnings	3,329	4,797	3,384	4,802	3,528	3,579	3,690	3,605	3,29
	Log Hourly Earnings	1,029	0,503	1,059	0,482	1,116	0,464	1,159	0,468	1,02
	Schooling Years	5,822	2,762	5,874	2,757	6,009	2,842	6,119	2,899	5,66
	Experience	23,761	12,348	23,675	12,369	24,203	12,446	24,365	12,376	23,57
	Tenure (Months)	116,938	109,310	112,610	108,318	114,152	110,524	115,158	111,048	112,94
	Female	0,445		0,460		0,447		0,453		0,44
	North	0,554		0,561		0,535		0,548		0,54

Table A1b

Descriptive Statistics, Workers in Foreign and Domestic Firms, 1991-1999 (10% foreign ownership definition).

	1996		1997		1998		1999		Annual
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Average
Centre	0,222		0,224		0,240		0,224		0,23
Lisboa and T. V.	0,196		0,187		0,194		0,201		0,21
Observations	388.000		402.799		384.213		411.037		404.572

Notes:

Monetary values are in real terms (2000 prices) and in euros.

Alentejo and Algarve regional dummies are omitted (their share ranges between .0% and 3.8%).

Table A2

Wage Premia of Foreign Firms, 1991-1999 (10% foreign ownership definition).

Specifications - Controls	1991	1992	1993	1994	1995	1996	1997	1998	1999	Avg.	Max.	Min.
No Controls	39,2%	35,3%	36,2%	35,9%	42,3%	37,6%	39,9%	36,3%	33,4%	37,4%	42,3%	33,4%
St. Errors	[0.002]**	[0.003]**	[0.003]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**			
Adjusted R-squared	0,046	0,029	0,035	0,039	0,061	0,049	0,06	0,053	0,044			
Human Capital (HC)	27,4%	26,5%	25,5%	29,2%	32,7%	28,1%	28,1%	28,0%	25,7%	27,9%	32,7%	25,5%
St. Errors	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**			
Adjusted R-squared	0,461	0,443	0,453	0,444	0,46	0,461	0,452	0,482	0,481			
HC and Firm Characteristics	15,0%	9,7%	10,4%	11,3%	13,0%	11,4%	9,2%	12,6%	12,9%	11,7%	15,0%	9,2%
St. Errors	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.002]**			
Adjusted R-squared	0,548	0,539	0,543	0,526	0,541	0,541	0,522	0,554	0,56			
Observations	524.230	517.007	494.461	462.406	501.256	485.522	503.643	477.692	503.156			

Notes:

Standard errors in brackets (* significant at 5%; ** significant at 1%).

Human Capital controls are: six dummies for educational degrees, a quartic in experience, a quadratic in tenure (in months), and a female dummy.

Firm Characteristics are: four regional dummies, log number of workers, a dummy for public firms and eight industry dummies.

References

Aitken, B., Harrison, A, and Lipsey, R. (1996) "Wages and Foreign Ownership. A Comparative Study of Mexico, Venezuela, and the United States", Journal of International Economics, 40, 345-371

Aitken, B. and Harrison, A. (1999) "Do Domestic Firms Benefit from Direct Foreign Investment? Evidence from Venezuela", American Economic Review, 89, 605-618.

Almeida, R. (2003) "The Effects of Foreign Owned Firms on the Labor Market", IZA Discussion Paper 785.

Barbosa, N. and Louri, H. (2002) "On the determinants of multinationals ownership preferences: evidence from Greece and Portugal", International Journal of Industrial Organization, 20, 493-515.

Barry, F. (1999) "FDI and Industrial Structure in Ireland, Spain, Portugal and the UK: Some Preliminary Results", University College Dublin, mimeo.

Blanchard, O. and Portugal, P. (2001) "What Hides Behind an Unemployment Rate: Comparing Portuguese and U.S. Labor Markets", American Economic Review, 91, 187-207.

Blomstrom, M. and Kokko, A. (1998) "Multinational Corporations and Spillovers", Journal of Economic Surveys, 12, 1-31.

Blundell, R., Costa Dias, M., Meghir, C. and Van Reenen, J. (2003) "Evaluating the Employment Impact of a Mandatory Job Search Assistance Program", IFS working paper W01/20.

Brown, C. and Medoff, J. (1989) "The Employer Size-Wage Effect", Journal of Political Economy, 97, 1027-1059.

Brown, D., Deardoff, A. and Stern, R. (2003) "The Effects of Multinational Production on Wages and Working Conditions in Developing Countries", National Bureau of Economic Research Working Paper 9669.

Cabral, S. (1996) "Comparative Export Behaviour of Foreign and Domestic Firms in Portugal", Banco de Portugal Working Paper 8-96.

Castellani, D. and Zanfei, A. (2002) "Multinational Companies and Productivity Spillovers: Is there a Specification Error?", University of Urbino, mimeo.

Castellani, D. and Zanfei, A. (2003) "Technology Gaps, Absorptive Capacity and the Impact of Inward Investments on Productivity of European Firms", Economics of Innovation and New Technologies, forthcoming.

Conyon, M., Girma, S., Thompson, S. and Wright, P. (2002) "The Productivity and Wage Effects of Foreign Acquisition in the United Kingdom", Journal of Industrial Economics, 50, 85-107.

Driffield, N. and Girma, S. (2003) "Regional Foreign Direct Investment and Wage Spillovers: Plant Level Evidence from the UK Electronics Industry" Oxford Bulletin of Economics and Statistics, 65, 453-474.

Dunning, J. (1977) "Trade, Location of Economic Activity and the MNEs", in B. Ohlin, P. Hesselborn and P. Wijkman (eds), *The International Allocation of Economic Activity*, Macmillan.

Feenstra, R. and Hanson, G. (1997) "Foreign direct investment and relative wages: Evidence from Mexico's maquiladoras", Journal of International Economics, 3-4, 371-393.

Feliciano, Z and Lipsey, R. (1999) "Foreign Ownership and Wages in the United States, 1987 – 1992", National Bureau of Economic Research Working Paper 6923.

Flôres, R., Fontoura, M. P. and Guerra Santos, R. (2000) "Foreign Direct Investment Spillovers: What Can We Learn From Portuguese Data?", Universidade Técnica de Lisboa, mimeo.

Gioia, C. and Thomsen, S. (2003) "International Acquisitions in the Danish Business: Selection and Performance", Copenhagen Business School, mimeo.

Girma, S., Greenaway, D. and Wakelin, K. (2001) "Who Benefits from Foreign Direct Investment in the UK?", Scottish Journal of Political Economy, 48, 119-133.

Girma, S. and Gorg, H. (2003) "Evaluating the Causal Effects of Foreign Acquisitions on Domestic Skilled and Unskilled Wages", IZA Discussion Paper 903.

Gorg, H., Strobl, E (2001) "Multinational Companies and Productivity Spillovers", Economic Journal, 111, F723-F739.

Griliches, Z. and Hausman, J. (1986) "Errors in Variables in Panel Data", Journal of Econometrics, 31, 93-118.

Guimarães, P., Figueiredo, O. and Woodward, D. (2000) "Agglomeration and the Location of Foreign Direct Investment in Portugal", Journal of Urban Economics, 47, 115-135.

Haddad, M. and Harrison, A. (1993) "Are There Positive Spillovers From Direct Foreign Investment? Evidence From Panel Data For Morocco", Journal of Development Economics, 42, 51-74.

Harris, R. and Robinson, C. (2002) "The Effect of Foreign Acquisitions on Total Factor Productivity: Plant-Level Evidence from UK Manufacturing, 1987-1992", Review of Economics and Statistics, 84, 562-568.

Harris, R. and Robinson, C. (2003) "Spillovers from Foreign Ownership in the United Kingdom: Estimates for UK Manufacturing Using the ARD", University of Durham, mimeo.

Haskel, J., Pereira, S. and Slaughter, M. (2002) "Does Inward Foreign Direct Investment Boost the Productivity of Domestic Firms?" Queen Mary, Department of Economics, Working Paper 452.

Heckman, J. Ichimura, H. and Todd, P. (1997) "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme", Review of Economic Studies, 64, 605-654.

Lichtenberg, F. and Siegel, D. (1990) "The Effect of Ownership Change on the Employment and Wages of Central-Office and Other Personnel", Journal of Law and Economics, 33, 383-408.

Lipsey, R. and Sjoholm, F. (2001) "Foreign Direct Investment and Wages in Indonesian Manufacturing", National Bureau of Economic Research Working Paper 8299.

Lipsey, R. and Sjoholm, F. (2003) "Foreign Firms and Indonesian Wages: A Panel Study", National Bureau of Economic Research Working Paper 9417.

Martins, P. (2003a) "Firm Wage Differentials in a Competitive Industry: Some Matched-Panel Evidence", International Journal of Manpower, 24, 336-346.

Martins, P. (2003b) "Rent Sharing Before and After the Wage Bill", paper presented at the Annual Conference of the European Association of Labour Economists, Seville.

Martins, P. (2003c) "Rent Sharing and Foreign Ownership", University of St Andrews, mimeo.

Martins, P. (2004) "The Foreign Ownership Wage Premium: Evidence from Agglomeration Effects", University of St Andrews, mimeo.

Mata, J. and Farinha, L. (1996) "The Impact of Foreign Direct Investment in Portuguese Manufacturing", Banco de Portugal Working Paper 16-96.

Mata, J. and Portugal, P. (2000) "Closure and Divestiture by Foreign Entrants: The Impact of Entry and Post-Entry Strategies", Strategic Management Journal, 21, 549-562.

Mata, J. and Portugal, P. (2002) "The Survival of New Domestic and Foreign Owned Firms", Strategic Management Journal, 23, 323-343.

Mata, J. and Portugal, P. (2004) "Patterns of Entry, Post-Entry Growth and Survival: A Comparison Between Domestic and Foreign Owned Firms", Small Business Economics, forthcoming.

Ministério das Finanças (2003) "Portugal no Espaço Europeu: O Investimento Directo Estrangeiro", Departamento de Prospectiva e Planeamento, Ministério das Finanças, Lisboa. [Available in www.dpp.pt]

Nogueira Leite, A., Ferreira Machado, J. and Curdia, V. (2001) "Portugal como Destino de Investimento Directo Estrangeiro. Estudo de Competitividade da Economia Nacional", Câmara de Comércio Americana em Portugal, GANEC – FEUNL, Lisbon.

Oy, W. and Idson, T. (1999) "Firm Size and Wages", in Orley Ashenfelter and David Card (eds) *Handbook of Labor Economics*, North-Holland: Amsterdam.

Pereira, P. and Martins, P. (2002) "Education and Wages in Portugal, 1991-1998", paper presented at the Bank of Portugal conference "Desenvolvimento Económico Português no Espaço Europeu", Lisbon.

Proença, I., Fontoura, M. P. and Crespo, N. (2002) "Productivity Spillovers from Multinational Corporations in the Portuguese Case: Evidence From A Short Time Period Panel Data", Universidade Técnica de Lisboa, mimeo.

Schleifer, A. and Summers, L. (1988) "Breach of Trust in Hostile Take-Overs", in A. Auerbach (ed) *Corporate Take-Overs: Causes and Consequences*, NBER: Chicago.

Sutherland, J. (2002) "The Experience of Work. Is Working for an Overseas-Owned Multinational So Different?", Employee Relations, 25, 149-167.

Vieira, J., Pereira, P. and Hartog, J. (2000) "Inter-Industry Wage Dispersion in Portugal: High but Falling", Empirica, 27, 353-364.

REDUÇÃO DO TEMPO DE TRABALHO E EMPREGO
LIÇÕES DA LEI DAS 40 HORAS

José M. Varejão
Faculdade de Economia do Porto e CETE

Lisboa, 11-12 de Março de 2004

Resumo

Em 1996, o Parlamento português aprovou a Lei 21/96 que tornou obrigatória a semana de trabalho de 40 horas. Pretendia-se aproximar os horários de trabalho portugueses dos europeus. Beneficiando de dados individuais relativos a trabalhadores e empregadores, procede-se aqui a uma avaliação dos efeitos da lei enquanto quase-experiência. Os resultados indicam que a redução do tempo de trabalho originou uma diminuição da procura de trabalho e um aumento do trabalho suplementar. Os salários horários também aumentaram, mas menos do que no resto da economia. Verifica-se, porém, uma diminuição do emprego que é particularmente significativa no caso dos estabelecimentos atingidos pela lei com mais intensidade.

1. Introdução

A tendência de diminuição da duração anual do trabalho que se verifica nos países industrializados tem sido, na Europa, pontualmente reforçada pela adopção de medidas de redução do limite máximo da duração do trabalho semanal. Tipicamente, estas medidas são implementadas centralizadamente, por iniciativa dos governos, em contextos de desemprego elevado.

A popularidade que estas medidas - habitualmente designadas como de "partilha de trabalho" - conheceram em alguns países parece, porém, resultar mais de uma crença política nos efeitos positivos de uma redução do tempo de trabalho sobre o emprego do que de argumentos teóricos sólidos.[1]

Apesar das dúvidas teóricas, vários países europeus, entre os quais a França, a Alemanha e a Itália, adoptaram nas décadas de 80 e 90, políticas destinadas a reduzir a duração máxima do trabalho semanal. Ainda que sem a motivação que resulta dos elevados níveis de desemprego, Portugal também não foi excepção. Na década de 90, por duas vezes os governos tornaram obrigatória a redução da duração normal do trabalho semanal, primeiro, em 1991, para as 44 horas e depois, em 1996, para as 40 horas. O objectivo deste artigo é, precisamente, o de avaliar as consequências da lei de 1996 em matéria de emprego, salários e horas efectivamente trabalhadas.

Na literatura não abundam os trabalhos de avaliação de medidas públicas de redução do tempo de trabalho, constituindo os trabalhos de Hunt (1999), Costa (2000) e Crépon e Kramarz (2002) as excepções mais notórias.[2] Em regra, estes estudos procedem à avaliação dos efeitos de reduções do tempo de trabalho num contexto de "quase-experiência ", isto é, admitindo que tais medidas podem ser legitimamente consideradas exógenas. Crucial é o facto de a medida afectar apenas um sub-grupo de indivíduos de entre o conjunto dos seus destinatários potenciais.

Todos estes estudos assumem como unidade de observação o trabalhador, avaliando os efeitos das medidas de redução do tempo de trabalho sobre as transições entre emprego e desemprego, as horas trabalhadas e as remunerações. O método das diferenças-nas-diferenças é utilizado no trabalho empírico (e.g., Meyer, 1995).

Neste estudo utiliza-se a mesma metodologia, mas recorre-se a dados sobre estabelecimentos que são, de facto, o local de reacção à perturbação que a medida representa.

O artigo está organizado da seguinte forma. Na secção 2 apresentam-se os principais resultados da Teoria da Procura de Trabalho sobre as horas de trabalho. A secção 3 apresenta as principais disposições da lei das 40 horas e a sua incidência. Na secção 4 descrevem-se os dados. Na secção 5 descreve-se a metodologia utilizada no trabalho empírico e na secção 6 apresentam-se os resultados. A secção 7 conclui.

---

1 Para uma descrição da evolução do tempo de trabalho na segunda metade do século XX em vários países e uma análise de várias experiências de partilha de trabalho, veja-se Hunt (1998). Para uma revisão da literatura sobre a teoria estática da procura de trabalho, veja-se Hart (1987).

2 Outros trabalhos (Trejo, 1991; Hamermesh e Trejo, 2000) ocupam-se de uma questão específica - a existência e aumento de um prémio obrigatório devido por horas de trabalho suplementar.

2. A Economia das Horas de Trabalho

Os efeitos de uma redução exógena das horas de trabalho sobre o emprego e as horas efectivamente trabalhadas são teoricamente conhecidos ainda que, por vezes, ambíguos.

O impacto de uma redução exógena da duração normal do trabalho semanal será diferente consoante na situação de partida a duração efectiva do trabalho for menor, igual ou maior do que a duração normal.

No caso em que, à partida, a duração efectiva do trabalho é inferior à duração normal, a redução da duração normal (desde que inferior, em valor absoluto, à diferença inicial entre aquelas duas durações), não produzirá quaisquer efeitos sobre o custo do factor trabalho, pelo que a combinação óptima de número de trabalhadores e duração individual efectiva do trabalho se manterá inalterada.

A situação oposta ocorre quando, à partida, a duração efectiva do trabalho excede a duração normal fixada por lei, isto é, quando há recurso a trabalho suplementar. Neste caso, uma redução da duração normal do trabalho dá origem a um aumento do custo de cada trabalhador sem que se altere o custo marginal das horas de trabalho. Assistir-se-á, por isso, a uma substituição de trabalhadores por horas de trabalho, isto é, a uma diminuição do emprego e a uma maior utilização de horas de trabalho suplementar.

Finalmente, no caso em que, na situação de partida, são iguais as durações normal e efectiva do trabalho semanal, o impacto dependerá da magnitude do choque. Para reduções da duração do trabalho normal relativamente pequenas, a empresa deslocar-se-á de uma solução de canto para outra, ocorrendo uma redução da duração efectiva do trabalho igual à da duração normal e um aumento do nível de emprego; neste caso, mantém-se a opção pela não utilização de trabalho suplementar.[3] No entanto, perante choques de maior magnitude, a empresa passará de uma solução de canto para uma solução interior com recurso a trabalho suplementar. Neste caso, os efeitos sobre o emprego serão ambíguos.

Aos efeitos de substituição entre número de trabalhadores e horas trabalhadas por indivíduo, haverá ainda que somar um efeito de escala que resulta do ajustamento da empresa ao aumento dos custos salariais que ocorre em todos os casos menos no primeiro devido quer à existência de custos fixos por trabalhador, quer à alteração na taxa de salário horário. Este efeito de escala originará, quaisquer que sejam os modelos de determinação salarial considerados, inequivocamente uma redução da quantidade de trabalho procurada.[4]

Em resumo, sabemos que de uma redução da duração normal do trabalho semanal resultará inequivocamente uma diminuição do número total de horas trabalhadas (procura de trabalho) e que os seus efeitos sobre o emprego serão ambíguos dependendo da proporção de empregadores que, antes da adopção da medida, utilizam ou não trabalho suplementar - o emprego aumentará no caso dos que não recorrem a trabalho suplementar e diminuirá no outro caso, sendo substituído pela utilização de trabalho suplementar adicional.

---

[3] Uma redução da duração normal do trabalho será considerada pequena se não for suficiente para induzir uma alteração na decisão inicial de não utilizar trabalho suplementar.

[4] Calmfors (1985) analisa a questão no contexto de um modelo de monopólio sindical, Booth e Schiantarelli (1987) fazem-no com modelos de negociação eficiente e Hoel e Vale (1986) usando modelos de *turnover* eficiente.

3. A Lei das 40 Horas

Em Portugal, compete à lei estabelecer os limites máximos dos períodos normais de trabalho. No entanto, relativamente aos máximos legais, podem os instrumentos de regulamentação colectiva de trabalho introduzir reduções dos períodos normais de trabalho, dentro dos limites consentidos pelos aumento da produtividade e sem que daí resultem inconvenientes de ordem económica ou social (DL. 409/71 de 27.9, art. 7º, nº 1). A lei impõe ainda que da redução dos limites máximos dos períodos normais de trabalho não resulte prejuízo para a situação económica dos trabalhadores, nem qualquer alteração nas condições de trabalho que lhes seja desfavorável (DL. 409/71 de 27.9, art. 8º, nº2).

Basta analisar a distribuição do período normal de trabalho dos trabalhadores por conta de outrém em Outubro de 1995 (Figura 1) quando o máximo legal estava fixado em 44 horas semanais (42 horas para os empregados de escritório) para se verificar que a negociação colectiva é uma via efectivamente utilizada para fixar (reduzir) a duração normal do trabalho semanal: nesta altura, o período normal de trabalho era, para 53.9 por cento dos trabalhadores, inferior a 42 horas por semana e para 17.7 por cento inferior a 40 horas (fonte: Quadros de Pessoal).

É, no entanto, importante reconhecer que o período normal de trabalho a que cada trabalhador se encontra obrigado (e, em particular, o facto de ele ser igual ou inferior ao máximo legal) é, em larga medida, determinado a um nível superior ao do estabelecimento/empresa, sendo comum coexistirem num mesmo estabelecimento trabalhadores com diferentes períodos normais de trabalho e abrangidos por diferentes instrumentos de regulamentação colectiva.[5]

Assim, quando em Julho de 1996, o governo fez aprovar na Assembleia da República uma lei (Lei 21/96 de 23.7) que instituía a semana de trabalho normal de 40 horas, a identificação das unidades abrangidas e não abrangidas pela medida escapava à decisão das próprias empresas. Uma empresa/estabelecimento seria abrangida pela medida se os seus trabalhadores (algum dos seus trabalhadores) estivessem eles próprios abrangidos por uma convenção colectiva que estabelecesse para a sua indústria e/ou profissão um período normal de trabalho superior a 40 horas por semana. Apenas negociando um Acordo de Empresa poderia uma empresa isolada alterar o seu estatuto perante a Lei das 40 horas.

De acordo com os Quadros de Pessoal relativos ao ano de 1995, 870374, ou seja 58 por cento do total dos trabalhadores por conta de outrém, tinham períodos normais de trabalho superiores a 40 horas semanais, sendo, por isso, abrangidos pela Lei 21/96 (Quadro 1).[6]

Estes 870374 trabalhadores trabalhavam em 97662 estabelecimentos (68.5 por cento do total de estabelecimentos com trabalhadores por conta de outrém ao serviço

---

5 De acordo com a lei, as convenções colectivas podem assumir três formas: (i) contratos colectivos, quando celebrados entre associações sindicais e associações patronais, (ii) acordos colectivos, quando outorgados por associações sindicais e uma pluralidade de entidades patronais para uma pluralidade de empresas e (iii) acordos de empresa, quando subscritas por associações sindicais e uma só entidade patronal para uma só empresa (DL 519-C1/79 de 29.12, art. 2º). Os acordos de empresa são formas raras de convenção colectiva, não abrangendo mais do que 5% do total de trabalhadores por conta de outrém. Os instrumentos dominantes, as convenções colectivas e os acordos colectivos, têm um âmbito predominantemente sectorial e/ou profissional.

6 Para uma descrição dos Quadros de Pessoal, veja-se a Secção 4.

recenseados nesse ano). Estes trabalhadores e estes estabelecimentos constituem os grupos abrangidos pelo mandato legal.

De acordo com a Lei, estes trabalhadores e os respectivos empregadores ficavam obrigados a adoptar, de forma faseada, segundo o calendário seguinte, a semana normal de trabalho de 40 horas:

- Até um de Dezembro de 1996 todos os horários superiores a 42 horas por semana deveriam ser reduzidos em duas horas semanais; os horários entre 40 e 42 horas por semana deveriam ser reduzidos para 40 horas semanais;

- Até um de Dezembro de 1997 todos os horários de trabalho ainda acima das 40 horas semanais deveriam cumprir o novo máximo legal.

A lei não alterou as disposições em vigor sobre o trabalho suplementar que está, desde 1991, sujeito a um máximo de 200 horas anuais e duas horas por dia de trabalho normal (DL. 398/92 de 16.10).[7]

A importância da obrigação de redução dos horários mais longos não deve ser subestimada - a duração média da duração normal do trabalho semanal nas unidades abrangidas era, em Outubro de 1995, de 41.9 horas, pelo que o esforço exigido corresponde a cerca de 4.5 por cento da duração média inicial (cerca de 6.5 por cento se nos referirmos apenas ao conjunto dos trabalhadores abrangidos nas unidades abrangidas).

Da comparação da distribuição do período normal do trabalho dos trabalhadores por conta de outrém antes e depois de ter entrado em vigor a Lei 21/96 - Figura 1 - verifica-se que esta produziu efectivamente o efeito pretendido - em 1995, 34.3 por cento do total praticava ainda horários superiores a 42 horas semanais (58.0 por cento acima de 40 horas); em 1997, após a conclusão da primeira fase, aquela percentagem tinha-se reduzido para 2.5 por cento (25.4 por cento acima das 40 horas); em 1998, era virtualmente nula a percentagem de trabalhadores com horários superiores a 40 horas semanais (0.8 por cento). Deve notar-se que, na sequência da aplicação da lei das 40 horas, a distribuição do Período Normal de Trabalho passou a apresentar um pico marcado para a duração 40 horas (correspondente a 80 por cento do total de trabalhadores por conta de outrém), mas que não se verificou qualquer alteração à esquerda das 40 horas.

Conclui-se, por isso, que a obrigação de redução dos horários mais longos foi, efectivamente, cumprida pelas unidades abrangidas, não se colocando questões de não cumprimento da lei, mas também que a duração do trabalho normal dos trabalhadores não abrangidos pela lei não foi afectada.

4. Os Dados

Os dados utilizados são provenientes de cinco vagas anuais (1995-1999) dos Quadros de Pessoal, uma fonte administrativa gerida pelo Departamento de Estatística do Trabalho, Emprego e Formação Profissional.

---

[7] De acordo com a lei (DL. 421/83 de 2.12, art. 7º, o trabalho suplementar prestado em dia normal de trabalho será remunerado com um acréscimo mínimo de 50% da retribuição normal na primeira hora e de 75% nas horas e fracções seguintes.O trabalho suplementar prestado em dia de descanso semanal, obrigatório ou complementar e em dia feriado será remunerado com o acréscimo mínimo de 100% da retribuição normal.

Os Quadros de Pessoal são de preenchimento obrigatório para todas as empresas e estabelecimentos com pelo menos um trabalhador assalariado. Anualmente, é recolhida informação, referida ao final do mês de Outubro, sobre cerca de 250 mil estabelecimentos e 2.5 milhões de trabalhadores. Cada empresa reporta um vasto conjunto de informação sobre si própria, cada um dos seus estabelecimentos e dos respectivos trabalhadores.

Para o trabalho empírico aqui apresentado, partiu-se do ficheiro de trabalhadores do ano de 1995. Foram apagados os registos relativamente aos quais o salário era nulo ou o período normal de trabalho não estava disponível. Apenas os trabalhadores por conta de outrém foram incluídos na amostra. Da informação contida no ficheiro retiveram-se as variáveis relativas ao nível de qualificação (agrupadas em três categorias - qualificados, não qualificados e aprendizes), salário mensal, período normal de trabalho semanal e número de horas de trabalho suplementar prestado.

Com base na informação disponível, cada trabalhador foi classificado como trabalhador a tempo inteiro ou a tempo parcial (dependendo de o seu período normal de trabalho ser ou não superior a 35 horas semanais). Foi criada uma variável dummy que identifica os trabalhadores que prestaram trabalho suplementar. Usou-se o mesmo procedimento para identificar os trabalhadores cujo salário mensal corresponde ao valor mínimo estabelecido por lei e para identificar os trabalhadores com antiguidade no estabelecimento inferior a três anos.

Finalmente, cada trabalhador foi classificado em função dos critérios estabelecidos pela Lei das 40 horas para efeitos da sua aplicação. Foram, assim, identificados os trabalhadores com horários semanais superiores e iguais ou inferiores a 40 horas, definindo-se, assim, ao nível do trabalhador, os grupos submetido ao tratamento e de controlo.

De seguida, utilizando o código de identificação dos estabelecimentos, agregou-se a informação presente no ficheiro de trabalhadores para o nível do estabelecimento. Obteve-se, assim, um segundo ficheiro com informação relativa a 142671 estabelecimentos (e 1499598 trabalhadores por conta de outrém). Dependendo de, em Outubro de 1995, empregar ou não trabalhadores com horários superiores a 40 horas semanais, cada estabelecimento foi classificado como tendo sido (ou não) abrangido pela lei, definindo-se, agora ao nível do estabelecimento, dois novos grupos de tratamento e de controlo.

Deste modo, o ficheiro de estabelecimentos contém informação sobre o sector de actividade (a um dígito), classe de dimensão, estrutura de qualificações e escolaridade dos seus trabalhadores, proporção de trabalhadores de cada sexo, a proporção no total dos trabalhadores com salário mínimo, a tempo parcial e com antiguidade inferior a três anos. Está igualmente disponível informação sobre o total de trabalhadores por conta de outrém no estabelecimento, o total de horas de trabalho suplementar utilizadas no mês de Outubro e o salário médio por trabalhador (em escudos correntes). Conhece-se o estatuto de cada estabelecimento relativamente à lei das 40 horas (abrangido ou não) e se o estabelecimento utiliza trabalho suplementar.

Para além de ser teoricamente a solução mais adequada, o recurso a dados de empregadores tem a vantagem de permitir comparar os efeitos produzidos pela medida em estabelecimentos que são por ela atingidos com diferentes intensidades. Para tal, calculou-se o rácio entre o número de horas que cada estabelecimento é, por força da lei, obrigado a reduzir (igual à diferença entre o somatório das horas normais de trabalho semanal antes da aplicação da medida e o produto entre o número de trabalhadores do estabelecimento e o novo máximo legal) e o número total de horas trabalhadas antes da aplicação da lei.

O ficheiro de estabelecimentos de 1995 assim obtido foi, posteriormente, utilizado como ficheiro mestre para seleccionar dos ficheiros correspondentes relativos a cada um dos anos seguintes (os estabelecimentos que, estando presentes, em 1995 sobreviviam em cada um dos anos de 1996, 1997, 1998 e 1999). Obteve-se, assim, um painel equilibrado de 67939 estabelecimentos (a que correspondem, em 1995, 947553 trabalhadores) que estavam continuamente presentes na base de dados entre 1995 e 1999. Dos estabelecimentos presentes na amostra, 46584 foram abrangidos pela lei das 40 horas.

5. Identificação dos Efeitos da Lei

Conforme foi descrito, a redução da duração do trabalho para um máximo de 40 horas foi súbita e inesperada. A sua motivação foi, sobretudo, política e determinada por um desejo de aproximação à legislação europeia, devendo ser considerada exógena relativamente ao nível de emprego. No momento em que a lei foi aprovada, a taxa de desemprego era de cerca de 7 por cento e estava a diminuir.

A medida, tendo como alvo os trabalhadores, atingia as empresas. No momento em que a lei entrou em vigor, era possível identificar trabalhadores que já trabalhavam 40 horas semanais ou menos (trabalhadores não abrangidos) e outros que ainda se encontravam acima daquele limite (trabalhadores abrangidos). Dadas as características do sistema de negociação colectiva, existiam também empresas/estabelecimentos que apenas empregavam trabalhadores não abrangidos pela lei (empresas não abrangidas) e empresas em que pelo menos um trabalhador foi abrangido (empresas abrangidas). Dentro do grupo das empresas abrangidas, é muito variável a intensidade da exposição à obrigação de reduzir o tempo de trabalho - desde um só trabalhador abrangido (num universo também ele variável) até à totalidade dos trabalhadores abrangidos.

Estas duas características do processo de redução do tempo de trabalho constituem as duas fontes de identificação dos efeitos da lei - a medida pode, legitimamente, ser considerada exógena e existe um grupo submetido e outro não submetido à medida.

A lei das 40 horas reúne, pois, as características necessárias para poder ser considerada uma "quase-experiência". É nesse contexto, que os seus efeitos serão avaliados.

Seguindo a literatura teórica avaliar-se-ão os efeitos da obrigação da redução do tempo de trabalho sobre o emprego, o salário mensal e a duração do trabalho suplementar. Vimos já que o efeito negativo sobre o emprego será tanto menos provável quanto maior for a contenção salarial e menor a substituição de horas normais de trabalho por horas suplementares.

A condição de identificação dos efeitos da lei é, então, que na ausência da medida a evolução das variáveis "resultado" - horas totais, emprego, horas suplementares e salários - seria a mesma para os estabelecimentos abrangidos pela obrigação de reduzir os horários de trabalho (grupo de tratamento) e os não abrangidos por aquela obrigação (grupo de controlo).

A estratégia de estimação consiste em comparar os resultados observados para as variáveis de interesse entre o grupo submetido ao tratamento e o grupo de controlo, utilizando o estimador de "diferenças-nas-diferenças" (Card e Sullivan, 1988):

$$\Delta = (Y_{t+1}^{A} - Y_{t}^{A}) - (Y_{t+1}^{NA} - Y_{t}^{NA}) \qquad (1)$$

em que Y é a variável de interesse, A designa o grupo abrangido pelo tratamento e NA o grupo de controlo, t refere-se ao período antes da medida e t+1 a um período após a implementação da medida.

As estimativas de Δ podem ser obtidas a partir da seguinte equação:

$$Y_i = \alpha + \beta_1 T_i + \beta_2 A_i + \beta_3 T_i A_i + \varepsilon_i \quad (2)$$

em que T é uma variável dummy que assume o valor um no período pós-tratamento e A é uma variável dummy igual a um se o indivíduo foi abrangido pela medida e zero no caso contrário e ε é uma perturbação aleatória. O coeficiente β mede a dupla diferença definida em (1).

A equação (2) é estimada em primeiras diferenças, escrevendo-se a equação de regressão na forma:

$$\Delta Y_i = \gamma_1 + \gamma_2 A_i + \nu_i \quad (3)$$

onde Δ representa a variação observada na variável *Y* entre *t* e *t+1*. γ é agora o parâmetro cuja estimativa corresponde à diferença-nas-diferenças definido em (1).

O estimador "diferenças-nas-diferenças" será cêntrico e consistente se for legítimo admitir que a selecção dos participantes no programa é aleatória, isto é, que a variável indicativa de exposição à medida (A) é estatisticamente independente dos resultados que seriam produzidos com e sem exposição à medida (Y_1 e Y_0).

Na presença de selectividade, será ainda possível obter uma estimativa consistente do efeito médio do programa. Para tal, bastará garantir que, a hipótese de independência entre A e (Y_1 e Y_0) se verifique mas condicionada a um vector de caraterísticas individuais observáveis (hipótese de desconhecimento do tratamento ou de selecção com base em variáveis observáveis).

Se o vector X incluir o controlo de um conjunto suficiente de características será possível garantir a validade da hipótese de independência entre A e (Y_1 e Y_0), o que será suficiente para garantir a consistência do estimador dos efeitos da participação no programa na presença de selectividade.

O controlo das características individuais pode ser efectuado a partir da equação (3) pela adição de um vector X que inclui as variáveis relevantes, ou seja:

$$\Delta Y_i = \gamma_1 + \gamma_2 A_i + \gamma_3 X_j + \nu_i \quad (4)$$

6. Resultados

6.1. Diferenças entre as Médias Amostrais

Como se viu, espera-se que as consequências de uma redução exógena da duração do trabalho normal se manifestem em termos do volume total de trabalho utilizado (isto é, o produto entre o número de trabalhadores empregados e a duração média individual do trabalho efectivo, aqui designado por Horas Totais) e o volume de emprego. Vimos também que a possibilidade de substituição de trabalhadores por horas de trabalho suplementar e a evolução dos salários influenciam a magnitude dos efeitos sobre o emprego e o volume de trabalho.

Nesta secção apresentam-se os resultados da estimação da equação de regressão (3), tendo-se seleccionado como variáveis dependentes - variáveis "resultado" - o total de horas trabalhadas (Horas Totais), o total de horas de trabalho suplementar (Horas Suplementares), o emprego (Número de Trabalhadores), o salário horário e o salário mensal. Todas as variáveis dependentes surgem na forma de primeiras diferenças (entre os respectivos logaritmos, no caso das variáveis "salário").[8]

Todas as diferenças são calculadas relativamente ao ano de 1995. Os resultados reportados no Quadro 2 são as estimativas obtidas para o coeficiente da variável binária indicadora do estatuto de cada estabelecimento relativamente à lei das 40 horas (igual a um se abrangida) nas equações em que a variável dependente é a variação da variável "resultado" indicada em cada linha.

Quadro 1

Cobertura da Lei das 40 Horas (1995)

	Trabalhadores	%	Estabelecimentos	%
Abrangidos	870374	58.0	97662	68.5
Não abrangidos	629224	42.0	45009	31.5
Total	1499598		142671	

[8] Nesta secção, como nas restantes, as equações em que a variável dependente corresponde às variações do logaritmo dos salários (horários ou mensais) foram ponderadas pelo emprego. Em qualquer caso, os salários estão medidos em escudos correntes.

Quadro 2

Impacto da Lei das 40 Horas - Variável de Tratamento Binária (Sem Regressores)

	1997		1998		1999	
Horas totais	-11.327	(8.64)	-20.004	(9.63)	-4.119	(10.11)
Horas suplementares	6.609 *	(2.55)	4.991***	(2.66)	7.827 *	(2.85)
Nº trabalhadores	0.042	(0.19)	0.054	(0.22)	0.460	(0.00)
Salário horário (log)	0.033 *	(0.00)	0.051 *	(0.00)	0.042 *	(0.00)
Salário mensal (log)	-0.004 *	(0.00)	-0.0003	(0.00)	-0.010 *	(0.00)

Desvios padrões entre parênteses; valores iguais a (0.00) significam que a estimativa do desvio-padrão é inferior a 0.001; *, ** e *** indicam estimativas significativas a 1, 5 e 10 por cento.

Relativamente à variável Horas Totais, os resultados indicam que a procura de trabalho aumentou menos nas unidades abrangidas pela lei do que nas que o não foram, sendo essa diferença no primeiro ano após a adopção da medida de -11.3 horas por estabelecimento.

As estimativas obtidas para as restantes equações indicam que a obrigação de reduzir a duração do trabalho semanal para 40 horas teve como consequências adicionais uma maior utilização de trabalho suplementar e um aumento do volume de emprego. Como se esperaria, o crescimento dos salário horários pagos pelas unidades abrangidas foi superior ao verificado no grupo de controlo, o que significa que a redução do número de horas de trabalho normal não foi acompanhada por uma redução proporcional dos salários mensais. Verifica-se, porém, que relativamente aos salários mensais, as unidades abrangidas exerceram alguma contenção que se traduziu num crescimento menor dos salário mensais.

Os resultados obtidos - todos eles consistentes com a teoria - indicam a ocorrência de um efeito de escala negativo potenciado pelo aumento dos custos salariais (horários). No entanto, as estimativas obtidas, com excepção das que se referem às diferenças no crescimento dos salários, são, geralmente, não significativas ao nível de 10%. Este facto não deve, porém, ser interpretado como não sendo estatisticamente diferentes os resultados observados para os dois grupos - abrangidos e não abrangidos pela medida. Recorde-se que o estatuto de cada estabelecimento perante a medida depende de o estabelecimento empregar pelo menos um trabalhador cujo período normal de trabalho excedia, em 1995, o novo máximo legal. Ora, entre as unidades classificadas como abrangidas pela medida com base nesse critério, a intensidade da exposição à medida é muito variável, esperando-se, naturalmente, que a sua reacção varie com essa mesma intensidade. A possibilidade de controlar para a intensidade da exposição à medida é, precisamente, uma das vantagens da utilização de dados relativos aos empregadores para analisar o efeito da redução das durações máximas do tempo de trabalho pois que, só neste caso, a medida é administrada com intensidades variáveis, oferecendo uma fonte adicional de identificação dos seus efeitos.

Para este efeito, foi construída a variável Intensidade do Tratamento - uma variável contínua que mede a proporção entre o número total de horas que cada

estabelecimento é obrigado a suprimir e o total de horas utilizadas em 1995 (cf. Secção 3).[9] A equação (3) foi reestimada usando esta variável como alternativa à variável binária indicadora de tratamento. Os resultados são apresentados no Quadro 3.

Quadro 3

Impacto da Lei das 40 Horas - Variável de Tratamento Contínua (Sem Regressores)

	1997		1998		1999	
Horas totais	-329.66 *	(105.03)	-464.435 *	(116.99)	-228.300 *	(122.79)
Horas suplementares	22.719	(30.93)	22.879	(32.29)	77.550 **	(34.65)
Nº trabalhadores	-4.662**	(2.33)	-5.709 **	(2.72)	0.009	(2.81)
Salário horário (log)	0.566 *	(0.01)	0.967 *	(0.01)	0.786 *	(0.01)
Salário mensal (log)	-0.082 *	(0.00)	-0.020 *	(0.01)	-0.229 *	(0.01)

Desvios padrões entre parênteses; valores iguais a (0.00) significam que a estimativa do desvio-padrão é inferior a 0.001; *, ** e *** indicam estimativas significativas a 1, 5 e 10 por cento.

Globalmente, estes resultados são consistentes com os do Quadro 2 indicando o mesmo padrão de resposta à aplicação da medida. Conforme previsto, o coeficiente da variável relativa à intensidade do tratamento é agora significativo, nomeadamente nas equações relativas às horas totais e número de trabalhadores. Deve, porém, salientar-se que a estimativa do coeficiente da variável de tratamento na equação relativa ao número de trabalhadores é agora negativa, indicando que quanto maior é a intensidade da exposição à medida mais negativo é o resultado sobre o emprego, sendo também negativo o efeito sobre as horas totais. Isto é, o padrão de resposta que emerge é que nas unidades atingidas mais fortemente o efeito de escala é particularmente negativo, assistindo-se também a uma substituição de trabalhadores por horas de trabalho suplementar.

Porque a existência de diferentes grupos expostos, com intensidades variáveis, ao tratamento oferece uma possibilidade adicional de fortalecimento dos resultados obtidos com quase-experiências baseadas em diferenças-nas-diferenças, estimou-se ainda um modelo que, mantendo as variáveis dependentes anteriores, incluía como variáveis independentes três variáveis binárias indicadoras do percentil de intensidade de exposição à lei a que cada estabelecimento pertence (a categoria omitida corresponde às unidades não abrangidas pela lei). Os resultados são apresentados no Quadro 4.

9 O valor médio amostral da variável Intensidade de Tratamento é 4,6%, variando entre o valor mínimo de 0% (unidades não submetidas a tratamento) e o valor máximo de 18,4%.

Quadro 4

Impacto da Lei das 40 Horas (Por Percentil de Intensidade de Tratamento)

	1997		1998		1999	
Horas totais						
Percentil 25	14.235	(11.15)	14.516	(12.42)	29.204*	(13.04)
Percentil 50	-14.503	(12.80)	-26.994	(14.26)	-16.428	(14.97)
Percentil 95	-27.355	(10.06)	-40.451	(11.20)	-21.395 ***	(11.78)
Nº trabalhadores						
Percentil 25	0.553**	(0.25)	0.746*	(0.29)	1.064*	(0.30)
Percentil 50	0.148	(0.29)	0.054	(0.33)	0.455	(0.34)
Percentil 95	-0.354	(0.22)	-0.417	(0.26)	0.052	(0.27)
Salário horário (log)						
Percentil 25	0.030 *	(0.00)	0.036*	(0.00)	0.041*	(0.00)
Percentil 50	0.050 *	(0.00)	0.066*	(0.00)	0.068*	(0.00)
Percentil 95	0.068 *	(0.00)	0.097*	(0.00)	0.097*	(0.00)

Desvios padrões entre parênteses; valores iguais a (0.00) significam que a estimativa do desvio-padrão é inferior a 0.001; *, ** e *** indicam estimativas significativas a 1, 5 e 10 por cento.

Os resultados obtidos não apenas confirmam a validade do exercício anterior - as respostas relativas dos diferentes grupos são consistentes com o que a teoria prevê - como indicam um padrão de resposta à lei interessante: quando aumenta o grau de exposição à medida, o efeito de escala torna-se cada vez mais negativo, passando-se de uma situação em que tal efeito se esgota numa redução das horas normais de trabalho para situações em que o próprio nível de emprego é atingido negativamente.

6.2. Resultados com Controlo das Características dos Estabelecimentos

O estimador "diferenças-nas-diferenças" mantém, como se viu, a sua validade na presença de selectividade se for possível incluir na equação de regressão um conjunto de variáveis que, controlando para a heterogeneidade observada, reponham a hipótese de independência entre a exposição ao tratamento e os respectivos resultados (com e sem exposição). O que se pretende é, essencialmente, eliminar a influência de todos os factores para além da exposição à lei que possam influenciar o comportamento das variáveis resultado.

Nesta secção apresenta-se o resultado da estimação da equação (4) que inclui um vector X de características das unidades incluídas na amostra. Pretendeu-se controlar o efeito de características que influenciam a capacidade de ajustamento dos estabelecimentos a choques sobre a procura de trabalho e, em particular, sobre a duração normal do trabalho. O vector X inclui, por isso, a proporção dos trabalhadores do estabelecimento com um horário semanal normal inferior a 35 horas (Tempo Parcial), a proporção dos trabalhadores do estabelecimento cuja remuneração coincide com o salário mínimo nacional (Salário Mínimo) e a proporção de trabalhadores do estabelecimento com uma antiguidade inferior a três anos (Antiguidade Reduzida) que funciona como uma *proxy* para a proporção de trabalhadores com contratos a termo, mas que também pode captar, dada a ausência do controlo correspondente, o efeito da idade do estabelecimento. Foi ainda incluída uma

variável dummy - Trabalho Suplementar - que indica se, em 1995, os estabelecimentos utilizavam (valor um) ou não trabalho suplementar.

Os resultados, apresentados nos Quadros 5 e seguintes, são, no que se refere ao efeito da exposição ao tratamento sobre as variáveis "resultado" consistentes com os da secção anterior: ser abrangido pela lei está associado a um efeito de escala negativo que se traduz, por comparação com o que ocorreria na ausência da medida, numa diminuição do número total de horas trabalhadas e de trabalhadores e num aumento do salário horário.

A presença de trabalhadores pagos ao nível do salário mínimo acentua a magnitude do efeito de escala e a evolução desfavorável do emprego. Este resultado é o esperado uma vez que para este conjunto de trabalhadores a redução da duração do trabalho origina, irremediavelmente, um aumento da remuneração horária, sem que seja possível contratar qualquer compensação.

O trabalho a tempo parcial influencia positivamente o desempenho das variáveis Horas Totais e Número de Trabalhadores e surge associada a menores aumentos do salário horário. Isto é, o trabalho a tempo parcial surge, efectivamente, como uma forma de trabalho flexível (sobretudo em termos de horários, mas também de custos salariais) que permite um melhor desempenho em termos de procura de trabalho e emprego.

Ao contrário do que se poderia esperar, dados os menores custos de despedimento de trabalhadores contratados mais recentemente e possivelmente com contratos a termo, a presença de trabalhadores com baixa antiguidade não surge associada a evoluções mais negativas das horas totais trabalhadas ou do emprego. O que pode acontecer é que esta variável esteja a captar o efeito da idade dos próprios estabelecimentos e que as estimativas correspondentes apenas reflictam o facto de estabelecimentos jovens, desde que sobrevivam, tenderem a crescer rapidamente.

Finalmente, é particularmente interessante o sinal da estimativa obtida para o coeficiente da variável Trabalho Suplementar na equação correspondente ao número de trabalhadores. Como se viu, nas unidades que já utilizam trabalho suplementar, a redução da duração normal do trabalho origina um aumento do custo dos trabalhadores relativamente ao custo das horas de trabalho. É este o único caso em que o efeito sobre o número de trabalhadores é inequivocamente negativo. É também isso que indica o resultado obtido para esta variável na equação relativa ao número de trabalhadores.

Quadro 5 (cont.)

Impacto da Lei das 40 Horas - Horas Totais

(Com Regressores)

	1997		1998		1999	
Indicador de tratamento	-17.098**	(8.92)	-31.389*	(9.93)	-22.794**	(10.40)
Tempo parcial	36.356	(31.36)	37.582	(34.90)	80.198**	(36.55)
Salário mínimo	-16.613	(12.09)	-23.596	(13.46)	-19.340	(14.09)
Antiguidade reduzida	83.394*	(10.61)	115.327	(11.81)	142.060*	(12.36)
Trabalho suplementar	-98.880	(19.20)	-201.390*	(21.37)	-384.016*	(22.38)
Constante	-19.825**	(9.46)	-26.034*	(10.53)	-49.113*	(11.03)
Intensidade de tratamento	-441.59*	(109.54)	-671.906*	(121.92)	-596.060*	(127.70)
Tempo parcial	22.685	(31.34)	20.753	(34.88)	63.143***	(36.53)
Salário mínimo	-10.610	(12.20)	-15.337	(13.58)	-11.728	(14.22)

Quadro 5 (cont.)

Impacto da Lei das 40 Horas - Horas Totais

(Com Regressores)

	1997		1998		1999	
Antiguidade reduzida	82.254*	(10.60)	11.350*	(11.80)	143.120*	(12.36)
Trabalho suplementar	-104.686*	(19.26)	-209.325*	(21.44)	-391.371*	(22.45)
Constante	-12.431*	(8.77)	-18,284***	(9.75)	-40.080*	(10.22)

Desvios padrões entre parênteses; valores iguais a (0.00) significam que a estimativa do desvio-padrão é inferior a 0.001; *, ** e *** indicam estimativas significativas a 1, 5 e 10 por cento.

Quadro 6

Impacto da Lei das 40 Horas - Número de Trabalhadores

(Com Regressores)

	1997		1998		1999	
Indicador de tratamento	-0.007	(0.20)	-0.060	(0.23)	0.209	(0.24)
Tempo parcial	1.047	(0.70)	1.589**	(0.81)	2.081**	(0.84)
Salário mínimo	-0.646**	(0.27)	-0.967*	(0.31)	-0.808**	(0.32)
Antiguidade reduzida	1.872*	(0.24)	2.722*	(0.27)	3.289*	(0.28)
Trabalho suplementar	-1.085*	(0.43)	-2.702*	(0.50)	-5.547*	(0.51)
Constante	-0.325	(0.21)	-0.555**	(0.25)	-1.192*	(0.25)
Intensidade de tratamento	-5.794**	(2.44)	-8.022*	(2.83)	-4.706***	(2.93)
Tempo parcial	0.707	(0.70)	1.155	(0.81)	1.646**	(0.84)
Salário mínimo	-0.531**	(0.27)	-0.818*	(0.32)	-0.680**	(0.33)
Antiguidade reduzida	1.895*	(0.24)	2.751*	(0.27)	3.319*	(0.28)
Trabalho suplementar	-1.197*	(0.43)	-2.850*	(0.50)	-5.675*	(0.52)
Constante	-0.087	(0.20)	-0.257	(0.23)	-0.859*	(0.23)

Desvios padrões entre parênteses; valores iguais a (0.00) significam que a estimativa do desvio-padrão é inferior a 0.001; *, ** e *** indicam estimativas significativas a 1, 5 e 10 por cento.

Quadro 7 (cont.)

Impacto da Lei das 40 Horas - Salário Horário

(Com Regressores)

	1997		1998		1999	
Indicador de tratamento	0.040*	(0.00)	0.057*	(0.00)	0.057*	(0.00)
Tempo parcial	-0.121*	(0.01)	-0.129*	(0.01)	-0.147*	(0.01)
Salário mínimo	0.069*	(0.00)	0.091*	(0.00)	0.103*	(0.00)

Quadro 7 (cont.)

Impacto da Lei das 40 Horas - Salário Horário

(Com Regressores)

	1997		1998		1999	
Antiguidade reduzida	0.016*	(0.00)	0.023*	(0.00)	0.030*	(0.00)
Trabalho suplementar	0.012*	(0.00)	0.009*	(0.01)	0.018*	(0.01)
Constante	0.079*	(0.00)	0.130*	(0.00)	0.170*	(0.00)
Intensidade de tratamento	0.591*	(0.03)	0.879*	(0.03)	0.845*	(0.03)
Tempo parcial	-0.116*	(0.01)	-0.119*	(0.01)	-0.138*	(0.01)
Salário mínimo	0.064*	(0.00)	0.083*	(0.00)	0.095*	(0.00)
Antiguidade reduzida	0.015*	(0.00)	0.023*	(0.00)	0.029*	(0.00)
Trabalho suplementar	0.017*	(0.00)	0.017*	(0.01)	0.026*	(0.01)
Constante	0.080*	(0.00)	0.130*	(0.00)	0.172*	(0.00)

Desvios padrões entre parênteses; valores iguais a (0.00) significam que a estimativa do desvio-padrão é inferior a 0.001; *, ** e *** indicam estimativas significativas a 1, 5 e 10 por cento.

7. Conclusões

Pelas suas características, a lei das 40 horas reúne as condições para ser considerada uma quase-experiência. A medida foi inesperada e motivada, sobretudo, por razões extra-económicas. No momento em que a lei foi aprovada, existiam trabalhadores com horários inferiores e superiores ao novo máximo pelo que apenas alguns foram abrangidos pela lei. De igual modo, alguns estabelecimentos (a unidade de análise que aqui se adopta) apenas empregavam trabalhadores com horários iguais ou inferiores a 40 horas semanais não sendo, por isso, abrangidos pela lei. Acresce que a dimensão dos grupos (de trabalhadores e de estabelecimento) submetidos e não submetidos ao tratamento é grande.

Beneficiando da existência de dados recolhidos junto dos empregadores (provenientes dos Quadros de Pessoal), procedeu-se a uma avaliação dos efeitos da lei, recorrendo-se ao estimador "diferenças-nas diferenças".

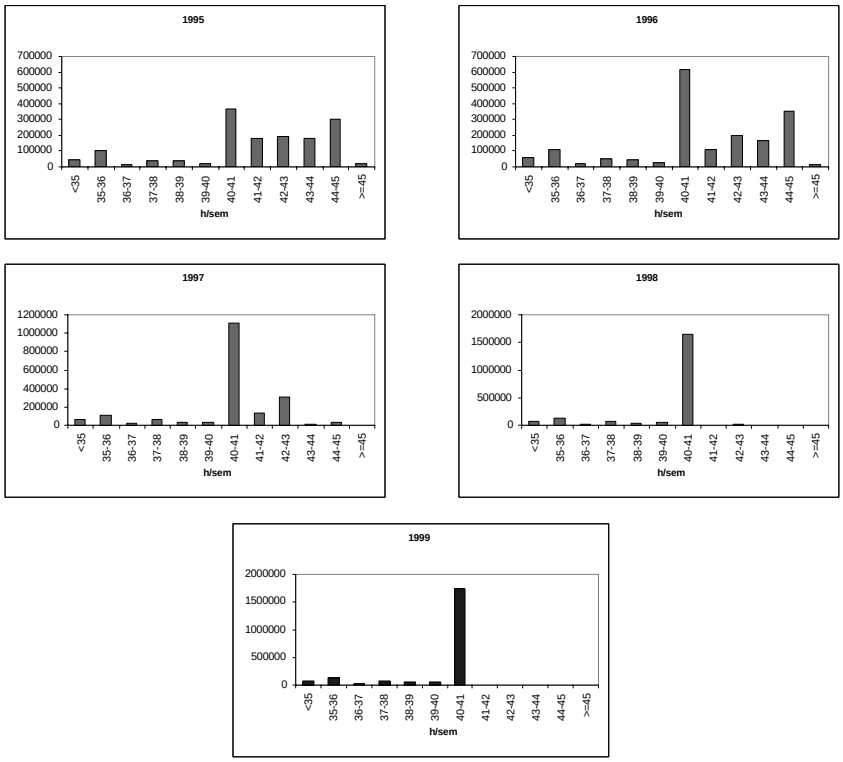
Os resultados obtidos indicam que a redução da duração máxima do trabalho semanal por via legal produziu um efeito de escala negativo que se traduziu numa redução do volume total de trabalho utilizado, acompanhada por um aumento da utilização de trabalho suplementar e por uma diminuição do emprego total. Embora se assista a um aumento dos salários horários, a exposição aos efeitos da lei é muito claramente acompanhada por uma contenção salarial que se traduz num menor crescimento dos salários pagos pelas unidades expostas à medida do que dos pagos pelas unidades que integram o grupo de controlo. Salienta-se que o efeito adverso sobre o emprego é muito ligeiro e ocorre apenas no sub-grupo dos estabelecimentos atingidos com mais intensidade pela medida.

A exploração das múltiplas fontes de identificação dos efeitos da lei, nomeadamente, a existência de grupos que recebem o tratamento com diferentes

intensidades e de múltiplos grupos de controlo (resultados não reportados) reforça a nossa confiança nos resultados aqui apresentados.

Figura 1

Distribuição do Período Normal de Trabalho



Referências

Booth, Alison e Fabio Schiantarelli (1987), "'The Employment Effects of a Shorter Working Week'", *Economica*, 54: 237-248.

Calmfors, Lars (1985), "'Work Sharing, Employment and Wages'", *European Economic Review*, 27: 293-309.

Card, David e Daniel Sullivan (1988), "Measuring the Effect of Subsidized Training In and Out of Employment", *Econometrica*, 56: 497-530.

Costa, Dora L. (2000), "Hours of Work and the Fair Labor Standards Act: A Study of Retail and Wholesale Trade", *Industrial and Labor Relations Review*, 53: 648-664.

Crépon, Bruno e Francis Kramarz (2002), "Employed 40 Hours or Not-employed 39: Lessons from the 1982 Mandatory Reduction of the Workweek", *Journal of Political Economy*, 110: 1355-1389.

Hamermesh, Daniel S. e Stephen J. Trejo (2000), "The Demand for Hours of Labour: Direct Evidence from California", *Review of Economics and Statistics*, 82: 38-47.

Hart, Robert (1987), *Working Time and Employment*, Boston: Allen and Unwin.

Hoel, Michael and Bent Vale (1986), "Effects on Unemployment of Reduced Working Time in an Economy Where Firms Set Wages", *European Economic Review*, 30: 1097-1104.

Hunt, Jennifer (1998), "Hours Reductions as Work-sharing", *Brookings Papers on Economic Activity*, 1: 339-381.

Hunt, Jennifer (1999), "Has Work-sharing Worked in Germany", *Quarterly Journal of Economics*, 114: 117-148.

Meyer, Bruce D. (1995), "Natural and Quasi-experiments in Economics", *Journal of Business and Economic Statistics*, 13:151-161.

Trejo, Stephen (1991), "The Effects of Overtime Pay Regulation on Worker Compensation", *American Economic Review*, 81: 719-740

Sessão 4 - Mercados e Desenvolvimento
Session 4 - Markets and Development

BUSINESS CYCLE AND LEVEL ACCOUNTING: THE CASE OF PORTUGAL

Tiago V. Cavalcanti[*]

Abstract

This paper studies what accounts for the recent business cycles phenomena in Portugal and why it is depressed relatively to the United States. The business accounting procedure applied here suggests that most of the changes in output per worker in Portugal over the period (mainly, from 1979 to 1991) can be attributed to changes in economic efficiency. For instance, the strong economic recovery in output per worker just after Portugal joined the European Union until the first years of the 1990s can be essentially attributed to improvements in economic efficiency. From 1979 to 2000, Portugal caught up with the industrial leader. Its output per worker is currently depressed by about 46% relative to the United States level (it was depressed by 57% in 1979). In the 1980s all of this depression in output per worker relative to the United States was due to the productivity factor. By 2000 Portugal depression relative to the United States was a mix of the French and Japanese depression. The labor factor accounted for roughly 24% of this depression, while the economic efficiency accounted by about 89%.

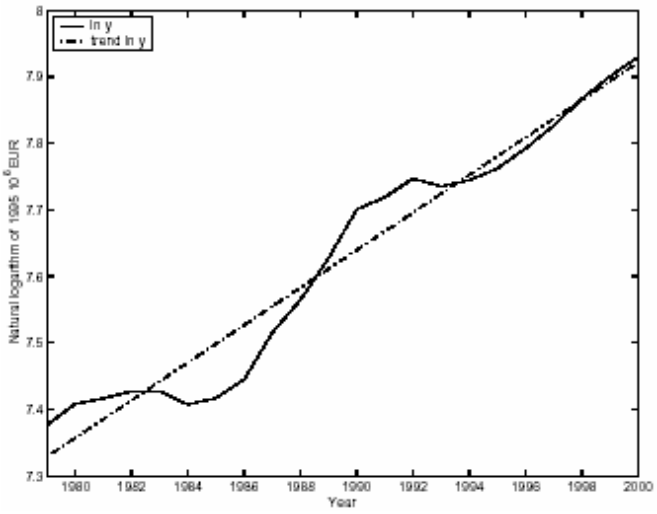
[*] Departamento de Economia, Universidade Federal de Pernambuco, and INOVA, Universidade NOVA de Lisboa. I thank António Antunes for helpful discussion, comments, and for providing some of the data and codes. I am also indebted to Isabel Horta Correia for helpful comments. I am responsible for any remaining error. This article was presented at the II Conferência do Banco de Portugal: *Desenvolvimento Económico Português no Espaço Europeu.*

1. Introduction

Prescott 2002 suggests that depression and prosperity are relative concepts. For instance, in 1979 Portugal was depressed relative to the United States, since Portugal's income per capita was roughly 57 percent lower than the United States level (see table 1). However, in the same year, output per worker in Portugal was about 5 percent higher than its long run trend-corrected level[1] (see figures 1 and 4) and therefore it was prosperous relative to its long-run trended correct level. Over the period Portugal caught up with the United States, but its output per worker is still 46% lower than the United States level.

Figure 1

Log of Output per Worker in Portugal, 1979 to 2000



This paper studies what accounts for the recent business cycles phenomena in Portugal. In particular, the article addresses the following questions: i) what explains Portugal's economic slowdown from 1979 to 1985 and from 1992 to 1996? ii) What are the factors that account for the Portuguese strong economic recovery from 1986 to 1991? iii) Lastly, but not least, what are the factors that explain why output per worker in Portugal is still depressed by about 46 percent relative to the United States level?

In order to see the importance of such analysis, Prescott (2002) found that France is depressed by about 30 percent relative to the United States. In addition, *the labor factor accounts for nearly all of the depression.* The capital and productivity factors are essentially equal in the two countries. Finally, according to his estimates the welfare gain of France reforming its labor market institutions are substantial (about 19 percent of actual consumption). On the other hand, total factor productivity explains why Japan is depressed

[1] The long run trend level calculated from 1979 to 2000 using Hodrick and Prescott filter procedure.

relative to the United States. In fact, productivity is approximately 20 percent higher in the United States than in Japan (c.f., Hayashi and Prescott (2002).

The methodology to address these questions was already applied to several countries to study some business cycles phenomena[2] and it is well explained in a recent article by Chari, Kehoe and McGrattan (2002). The authors show that a large class of economic models, including those with various detailed frictions (e.g., financial frictions, entrepreneur decisions, and credit market imperfections) are equivalent to a prototype growth model with time-varying wedges that look like time-varying productivity, labor taxes, and capital income taxes. They call these "distortions" as the efficiency, labor and productivity wedges, respectively. Institutions, public policies, fiscal and monetary policies affect the three wedges and therefore the allocations of capital and labor, and productivity in the economy.

Using this prototype growth model, Chari et al. (2002) propose an accounting procedure for guiding researchers in developing quantitative models of economic fluctuations. This procedure uses the real data together with the equilibrium conditions of a prototype growth model to measure these wedges. Once the wedges are constructed it is possible to assess what fraction of the output fluctuation can be attributed to each wedge separately and in combination. This is similar to the analysis conducted by Prescott to study the French Post War performance, and is exactly what is applied here to study the Portuguese economy.

We found that most of the movements in output per worker in Portugal can be attributed to the efficiency and labor wedges. For instance, the strong economic recovery in output per worker just after Portugal joined the European Union until the first years of the 1990s can be essentially attributed to improvements in economic efficiency. In terms of the output levels, Portugal is currently depressed by about 46% relative to the United States level, and its depression is a mix of the French and Japanese depression. The labor factor accounted for roughly 24% of this depression, while the economic efficiency accounted by about 89%.[3]

The remained of this paper is organized as follows: section 2 develops the level accounting. It decomposes the differences in output per worker between Portugal and the United States in three factors: productivity, labor, and capital inputs. Section 3 develops the business accounting. It describes the model economy and implements the quantitative exercises. Concluding remarks and policy implication are collected in the last section.

2. Level Accounting: Portugal Relative to the United States

Idea behind the Methodology

Since Robert Solow's (1956) seminal work, there is a long tradition in economics to account for the economic performance of the countries. Solow's growth accounting procedure decomposes the changes in economic performance of nation in a given period in changes in the labor force, productivity, and capital accumulation. Following the same

---

2 See the January 2002 issue of *The Review of Economic Dynamics, edited by Timothy J. Kehoe and Edward C. Prescott (2002), and the May 2002 issue of the American Economic Review.*

3 The capital factor goes in the opposite direction. It is higher in Portugal than in the United States.

methodology Prescott (2002) and others provide a simple framework to study the relative performance of the nations.

From 1979 to 2000 Portugal caught up with the United States. Its output per worker was about 46% and 54% of the American output per worker in 1979 and 2000, respectively. In this section we investigate what accounts for the Portuguese depression relative to the United States. Following Prescott (2002), if we define the production function as[4] $Y_{it} = \left(A_{it}\gamma^{t}\right)^{1-\alpha} K_{it}^{\alpha}\left(L_{it}h_{it}\right)^{1-\alpha}$, we can decompose the log of output into four factors:

$$\log\left(y_{it}\right) = t\log\left(\gamma\right) + \log\left(A_{it}\right) + \frac{\alpha}{1-\alpha}\log\left(\frac{k_{it}}{y_{it}}\right) + \log\left(h_{it}\right). \qquad (1)$$

The first factor is the trend growth, the second one is the productivity factor, the third is the capital factor, and the last one is the labor factor. Prescott (2002) argues that the production function is the same among nations. Therefore, the capital share, α, and the common trend growth, γ, are the same across countries.[5] Differences among nations appear in differences in the productivity, capital, and labor factors, which are in turn due to differences in institutions and economic policies.

Level Accounting

Table 1 shows this decomposition of output per worker in Portugal relative to the United States[6] in 1986 (just when Portugal joined the European Union) and in 2000.[7] In 1986 Portugal was depressed relative to the United States by about 57%. Table 1 shows that most of the depression was due to a low productivity of the factors of production. The other factors were essentially the same between the two countries in the period.

The last row of table 1 shows that Portugal is currently depressed relative to the United States by about 46%. The labor and productivity factors account for nearly all of the depression 2000.[8] The labor factor accounts for roughly 24%, while the productivity factor accounts for about 89% of the depression.[9]

---

4 γ is the common trend, A_{it} is country specific and corresponds to a productivity factor, L_{it} is the working population. The other variables are standard in the literature. Notice that this production function is almost identical to the one used previously.

5 α is roughly 0.4, which is consistent to (Gollin (2002) and γ is the growth in the world knowledge that can be accessed at low cost and that enhances production possibilities. Prescott (2002) assumes a 2 percent growth in world knowledge, which corresponds to the United States annual growth rate in the last century.

6 We normalize the values of the factors in the United States to one.

7 For comparison reasons, we used the series from the OECD Labor Statistics for the labor input. For Portugal the first observation for the labor input is 1986. The data for capital-output ratio in Portugal is described in appendix A. Capital-output ratio in the United States has been remarkably stable around 2.3 in the last two decades (see Prescott (2902)).

8 This is consistent to the next section, since it shows that most of the movements in output per worker from 1979 to 2000 can be attributed to changes in the efficiency and labor wedges (see figure 4). The efficiency wedge alone, for instance, accounts for almost all the recovery in output per worker from 1986 to 1991 (see figure 4).

9 Capital to output ratio is higher in Portugal than in the United States.

Table 1

Level Accounting Relative to the United States

	GDP Factor	Productivity Factor	Capital Factor	Labor
France*, 1998	-31%	6%	1%	-37%
Japan*, 1998	-31%	-33%	3%	-1%
Portugal**, 1986	-57%	-59%	2%	0.40%
Portugal**, 2000	-46%	-41%	4%	-11%

*See Prescott (2002). **Author's calculation. GDP per worker are from Heston et al. (2002). The labor input are from the Labor Statistics of the OECD.

Notice that Portugal depression relative to the United States is a mix of the French and Japanese depression. The French depression is due entirely to the depressed labor factor, while the Japanese is due to a depressed productivity. The accounting procedure suggests that labor market reform can lead to a substantial increase in output per worker in Portugal. Indeed, hours worked in Portugal are 12% lower than in the United States, and the model suggests that Portugal might make a better use of the available working time introducing less distorted labor market policies and more flexible labor market institutions. The potential for rapid economic growth in Portugal, however, is through improvements in productivity. In section 4 we discuss some policies that aim to improve economic efficiency.

3. Business Cycle Accounting

Idea behind the Methodology

This methodology follows a similar principle as the growth accounting procedure developed by Solow. As Chari (2002) suggest the method has two components: an equivalence result and an accounting procedure. The equivalence result follows from the fact that a large class of detailed economic models are equivalent to a stochastic growth model with time-varying wedges (they will be defined shortly). These wedges represent distortions that affect economic productivity and the allocation of inputs in the production. Given the model, we use real data to construct the wedges (distortions).

We then follow Lucas (1980) who argues that one of the functions of theoretical economies is to provide economic models that can serve as laboratories. Once we have the wedges, we can use our model as a laboratory to, for instance, assess how much of an economic recovery (Portugal from 1986 to 1991) can be accounted by improvements in productivity, or changes in policies that effected the capital and labor inputs. We do this by counterfactual exercises: we feed the values of the wedges into the growth model one at a time and in combination keeping the others at their long rung trend level to assess what fraction of the output movements can be attributed to each wedge separately and in combination.

3.1. The Model

Our model is based on the stochastic growth model proposed by Chari et al. (2002). In this model there are time-varying wedges, that look like time-varying productivity, A_t, one minus labor taxes, $1-\tau_t^n$, and one minus capital income taxes, $1-\tau_t^k$. These wedges are "distortions"' and represent policies and institutions, which affect productivity, capital accumulation, and hours worked. The problem of a representative agent in this economy is to choose consumption, c_t, hours worked, h_t, and capital, k_{t+1}, to maximize

$$U = E_t\left\{\sum_t \beta^t u\left(c_{t,} h_t\right)\right\}, \beta \in (0,1), \tag{2}$$

subject to

$$c_t + k_{t+1} - (1-\delta)k_t = \left(1-\tau_t^h\right) w_t h_t + \left(1-\tau_t^k\right) r_t k_t + T_t, \tag{3}$$

where the notation is standard, except that T_t are lump-sum taxes.

Firms choose K_t and H_t to maximize profits,

$$A_t F\left(K_t, H_t\right) - r_t K_t - w_t H_t. \tag{4}$$

In equilibrium, the solution is summarized by the resource constraint

$$c_t + K_{t+1} + g_t = Y_t + (1-\delta)K_t \text{ and } Y_t = A_t F\left(K_t, H_t\right), \tag{5}$$

the standard market clearing conditions, and the following marginal conditions:[10]

$$\frac{u_h(t)}{u_c(t)} = -\left(1-\tau^h\right) A_t F_H(t), \tag{6}$$

$$u_c(t) = \beta E_t\left\{u_c(t+1)\left[A_{t+1} F_K(t+1)\left(1-\tau_{t+1}^k\right) + 1 - \delta\right]\right., \tag{7}$$

$$A_t = \frac{Y_t}{F\left(K_t, H_t\right)\}}. \tag{8}$$

These wedges $\left(1-\tau^h,\ 1-\tau^k,\ A_t\right)$ represent the shocks that hit the economy and affect total factor productivity, and the allocations of capital and labor. The intra-temporal labor-consumption choice (equation (6)) and the inter-temporal consumption-savings

[10] Notice that the model is a closed economy. Since Portugal is a small open economy we plan to extend the analysis to the case of a open economy following the lines of Correia, Neves and Rebelo (1995).

substitution (equation (7)) amplify the shocks and propagate them through time.[11] Among the shocks that affect these wedges, some important candidates are: technology shocks (e.g., the introduction of new machines and computer), natural disasters and whether shocks (e.g., earthquake, El Niño), monetary shocks (e.g., interest rate changes), political shocks (e.g., elections), taste shocks, institutional shocks (e.g., changes on bankruptcy laws), and policy reforms (e.g., labor market reform). For instance, institutions represented by the risk of confiscation and expropriation of private investment would clearly be linked to the investment wedge (equation (7)). Countries with more secure property rights have smaller τ_t^k compared to those where property rights are not well defined and policies are more distortional. However, Chari, Kehoe and McGrattan (2002) show that more detailed models with frictions that are linked to the investment wedge (e.g., credit market frictions) might also have some impact on the efficiency wedge, A_t.[12]

3.2. Quantitative Exercises

3.2.1. Constructing the Wedges

Notice that the efficiency and labor wedges are obtained from two static equations, (6) and (8). Therefore, in order to construct these two wedges we just need to: i) define functional forms for the utility function, and production function; ii) assign values for the parameters of the model; and iii) use the data[13] and the first order conditions (6) and (8).

We follow Chari et al. (2002) and assume that the production function is Cobb-Douglas, $F\left(K_t, H_t\right) = K_t^{\alpha} H_t^{1-\alpha}$, and the utility function has a logarithm form,[14] $U\left(c,h\right) = \log c + \psi \log\left(\bar{h} - h\right)$. We used standard parameter values for capital share, depreciation rate, subjective discount factor, and preference parameters, ψ and $\bar{h}$.[15] Table 2 reports these parameter values.

---

11 Shocks that have a negative impact on productivity lower the marginal return of the factors of production. Since consumers prefer to smooth consumption over time they might prefer to lower savings for some time when a shock hits. On the aggregate level, this leads to lower investment and a lower capital stock in the future.

12 See Antunes and Cavalcanti (2003) for a model in which credit market policies affect total efficiency.

13 We need data for y_t, k_t, h_t, and c_t. Details about the data used are in appendix A.

14 These functions are standard in the literature. We use them for comparison reasons.

15 Details about the parameterization procedure are in appendix B.

Table 2

Parameter Values.

α	δ	β	ψ	$\bar{h}$
0.44	0.06	0.97	2.24	1

The raw series of the efficiency and labor wedges are displayed in figure 2. We can see that the efficiency wedge increased substantially in the period,[16] while the distortions that manifested themselves in the labor wedges (e.g., labor market institutions and labor income taxes) worsened (except in the period from 1986 to 1991). These results are consistent to the findings in section 2. There, we show that Portugal caught up to the United States and that while in 1986 most of the difference in output per worker between the two countries was due to differences in productivity, in 2000 this difference was explained not only by differences in economic efficiency, but also by the depressed labor input of Portugal relative to the United States.

Given the raw series of the labor and productivity wedges, we estimate the stochastic process for τ_t^h and A_t, by estimating a first-order vector autoregressive model. Finally, we use a finite grid for k_t, A_t, and τ_t^h, and value function interactions to find the optimal policy function for the investment decision.[17]

[16] It increased about 46% from 1979 to 2000. The period, for instance, is associated to a rapid exposure to foreign competition, which forced industries to adopt new technologies. The country's public infrastructure also benefited from European Union's transfers (structural funds), which increased productivity in many economic activities (see Baer and Leite (2003)).

[17] The Matlab code is available upon request. See appendix C and Chari et al. (2002) for more details. In a deterministic version of the model the data generated by the model would be the same as the real data.

Figure 2

Efficiency and Labor Wedges

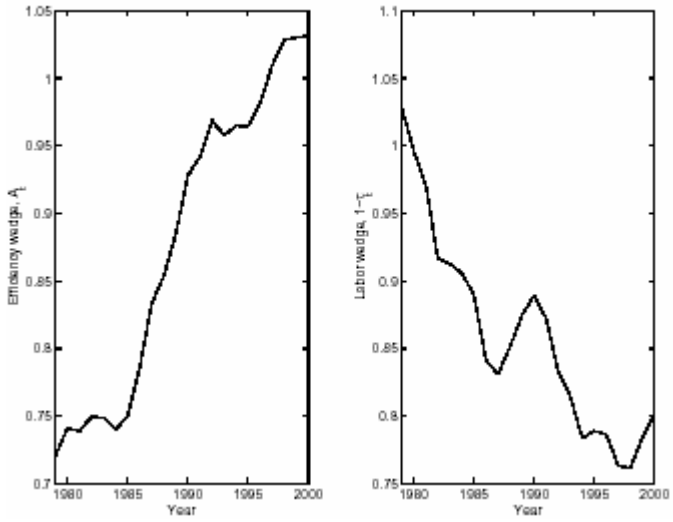


Table 3 reports the steady-state values of the endogenous variables, as well as the steady-state wedges and output per worker growth rate consistent to the data. Notice that the estimated investment frictions (e.g., capital income taxes, bankruptcy laws), τ^k, that affect the economy by raising the cost of investment are substantial. It is, for instance, about 2.3 times the labor wedge.

Table 3

Steady-State Values of Endogenous Variables, and Wedges and Growth Rate Consistent With the Model.

y	c	i	g	k	l	A	τ^h	τ^k	γ
1	0.62	0.22	0.16	2.46	0.26	1.43	0.14	0.32	1.03

Figure 3

Detrended Output per Worker, Detrended Efficiency and Labor Wedges, 1979-2000

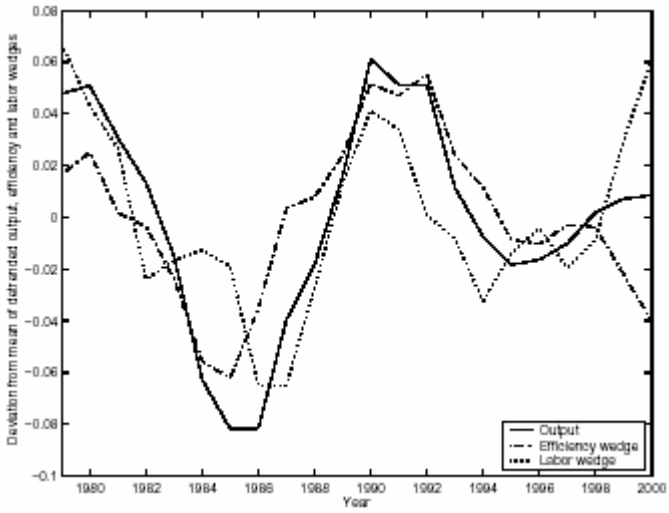


Figure 3 displays detrended output per worker, and the detrended efficiency and labor wedges from 1979 to 2000.[18] The efficiency and labor wedges followed a similar path of the output per worker. They are both pro-cyclical and their cycles have a similar amplitude as the output cycle. We notice that from 1979 to 1985 output in Portugal declined monotonically from the long-run trend level. In 1985 it was about 8% below the trended level. On the other hand, at the time that Portugal joined the European Union in 1986, output started to increase sharply. Its peak was in 1991 when it was about 6% above trend. Then it started to decrease until 1995 when it had a slight recovery. Notice, however, that while in 2000 the detrended output was slightly above the long-run trend, efficiency was 4% below. The labor wedge, on the other hand, was in 2000 6% above the trended level.

3.2.2. Counterfactual Analysis

Now we start to assess the separate contributions of the efficiency and labor wedge. In order to accomplish that, we include these wedges one at a time in the model and set the other wedges at their long run trend level.[19] In terms of output per worker (see figure 4), the model with the efficiency wedge alone mimics well the Portuguese economic slowdown from 1979 to 1985 and the strong economic recovery from 1986 to 1991. In fact,

[18] The trend of the variables is based on data from 1979 to 2000. The data from 1953 to 2000 yield a similar trend.

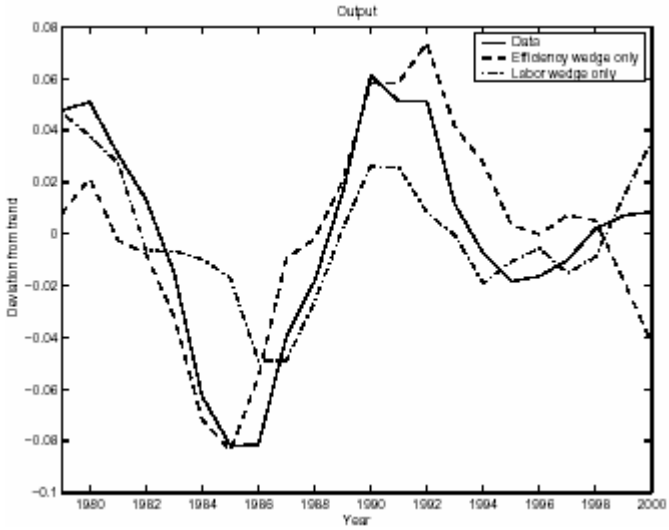
[19] As in Chari et al. (2002) the contribution of the investment wedge is measured residually, as the difference between the contribution of the efficiency and labor wedges and the data.

it accounts for essentially all of the movements in output from 1982 to 1991. From 1992 onward the model, however, yields a continue economic slowdown, while in the data there is an economic recovery around 1995. The reason is that the model with the efficiency wedge alone misses the strong improvements in the labor wedge from 1995 onward. Notice also that this model cannot account for the changes in investment and labor intensity in the period (see figures 7 and 8 in appendix D).

The model with only the labor wedges misses the sharp decline in output per worker from 1979 to 1985. Output per worker falls only about a half as much by 1985 as output actually fell relatively to the 1979 trended level: 6% vs. 12%. The economic recovery from 1986 to 1991 is also not so strong in the model as it was in the data. However, the economic recovery, in the second half of the last decade, is stronger in the model than in the data. The model with the labor wedge alone does a poorer job than the model with only the efficiency wedge in accounting for the output fluctuations in Portugal in the last 22 years. Mainly for the period from 1979 to 1991. It, however, accounts for essentially all variations in labor hours worked in the period (see figure 8 in appendix D).

Figure 4

Data for Detrended Output in Portugal and in the Model.

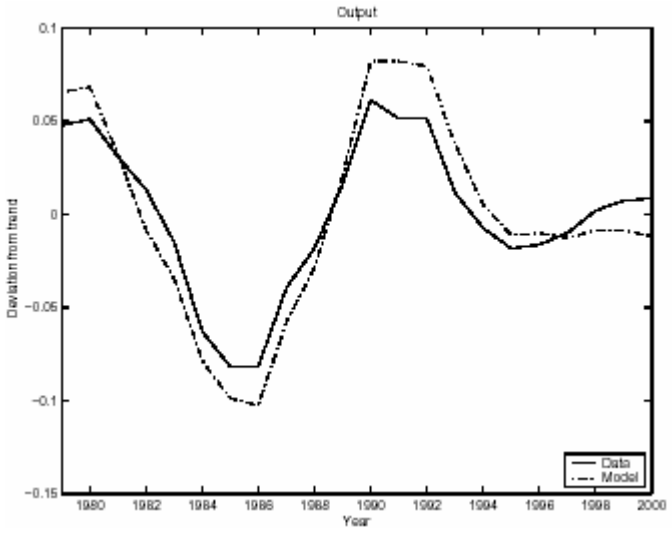


We now keep the investment wedge at its steady state level, and simulate the model with the estimated series for the efficiency and labor wedges. Figure 5 shows that the model with both wedges replicates well the real data and accounts for all the movements in output. The model deviates from the real data in the amplitude of the economic recession and economic boom in 1985 and 1991, respectively. The stronger amplitude of a recession in the model than in the data means that the model with both wedges misses some improvements in investment frictions, which lowered the investment cost. Comparing to the results of Chari et al. (2002) for the United States, we notice that the

investment wedge had a stronger effect in the recent business cycle phenomena in Portugal than in the United States.

Figure 5

Data for Detrended Output in Portugal and in the Model.



Notice in figure 9 in appendix D that the investment wedge is key to explain changes in the investment in the period. Mainly, from 1988 to 1994. The model with the efficiency and labor wedges misses the economic slowdown in the growth of investment from 1989 to 1994. This explains the discrepancy in output per worker for the same period in the data and in the model with both wedges.

4. Concluding Remarks and Policy Implications

This paper studied the business cycles phenomena in Portugal from 1979 to 2000. Over the period, Portugal joined the European Union and went through important economic changes:[20] (i) it followed the world trend towards privatization and liberation;[21] (ii) the economy went through a continuing opening process, exposing domestic producers to outside competition; and (iii) it used European Union's transfers ("structural funds") to improve substantially its public infrastructure.

The accounting procedure developed by Chari et al. (2002) and applied here suggests that most of the changes in output per worker in Portugal over the period (mainly, from 1979 to 1991) can be attributed to changes in economic efficiency. For instance, the strong economic recovery in output per worker just after Portugal joined the European

[20] See Baer and Leite (1992) and Baer and Leite (2003) for a recent overview of Portugal's recent economic performance and policies.

[21] It sold over 100 State enterprisers throughout the 1990s Baer and Leite (2003).

Union until the first years of the 1990s can be essentially attributed to improvements in economic efficiency.[22] Therefore, most of the economic changes and policies implemented in the period that have an effect on output manifested themselves in the efficiency wedge. For example, the availability of "structural funds" certainly decreased the investment cost (investment wedge), but the investments made (e.g., highways) improved the overall economic efficiency.

Over the period, Portugal caught up with the industrial leader. Its output per worker is currently depressed by about 46% relative to the United States level (it was depressed by 57% in 1979). In the 1980s all of this depression in output per worker relative to the United States was due to the productivity factor. By 2000 Portugal depression relative to the United States was a mix of the French and Japanese depression. The labor factor accounted for roughly 24% of this depression, while the economic efficiency accounted by about 89%.[23] These results suggest that labor market reform can have a substantial increase in output per worker and welfare in Portugal. Indeed, hours worked in Portugal are 12% lower than in the United States, and the model suggests that Portugal might make a better use of the available working time introducing less distorted labor market policies and more flexible labor market institutions. The *1999 OECD Employment Outlook*, for instance, shows that employment protection (e.g., mandated severance pay and advanced notice periods) in Portugal is one of the highest among OECD countries. Recent studies (cf., Alvarez and Veracierto (1999), (Bertola and Rogerson (1997), and Blanchard and Portugal (2001)) have shown that, in general, layoff costs not only increase job tenure, but they also have a negative impact on labor demand. When firms decide to hire a worker, they take into account the expected cost of hiring, and dismissal costs are clearly part of it. Therefore, to avoid layoff costs firms hire workers less frequently.

Improvements in productivity, on the other hand, are not so straightforward as changes in hours worked through labor market reform. However, since Portugal still lags behind the industrial leaders in productivity, there is a *potential* for rapid economic growth in the country (see Parente and Prescott (2000)).[24] Portugal can certainly increase productivity by increasing competition, as suggested by Bailly and Solow (2001). This can clearly be done by reducing the monopoly rights of some industries (e.g., trade barriers - see Parente and Prescott (2000)). Another way to increase productivity is through improvements in the legal and enforcement system. A better legal system, for instance, increases the availability of funds to finance projects and the adoption of new technologies. Financial entities are willing to loan more when there is a higher probability that debtors will pay their financial obligations. Antunes and Cavalcanti (2003) show that credit market institutions and policies have a strong effect on economic productivity. Credit constrained entrepreneurs reduce the size of their project and productivity.

---

22 The model with only the efficiency wedge, however, cannot account for the changes in hours worked and investment in the period.

23 The capital factor goes in the opposite direction. It is higher in Portugal than in the United States.

24 The growth of rich countries is driven mainly by the growth of productivity knowledge (technological progress), while for poor countries growth might be also driven by reducing the barriers (e.g., institutional changes) to technology adoption and the better use of the factors of production.

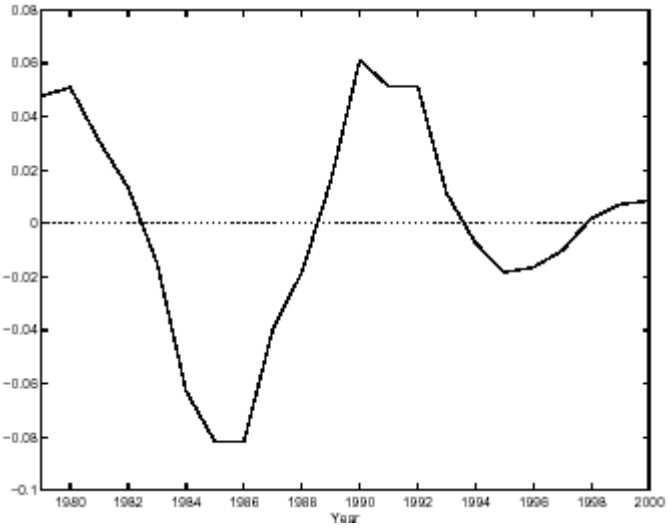
A Data and Parameterization of the Model

We use three main data sets in this paper. The first is the *Séries Longas para a Economia Portuguesa*,[25] published by Banco de Portugal, which covers the period from 1953 to 1995. The second is the annual macro-economic database of the European Commission's Directorate General for Economic and Financial Affairs, also known as the Ameco database. It goes from 1960 to 2003. Finally, we use the OECD data on average yearly hours worked per person employed from Groningen Growth and Development Centre and The Conference Board (2003). They go from 1950 to 2002, but have annual periodicity only from 1979 on. For the period from 1996 to 2000 we also used data published by Banco de Portugal on the growth rates of the GDP and several components of domestic expenditure.

Because of lack of data on yearly hours worked per person employed prior to 1979, and on several components of GDP per capita after 2000, we studied the period from 1979 to 2000, which encompasses two complete business cycles. (See figure 6.) To extend the *Séries Longas* after 1995, we used the growth rates published by Banco de Portugal for period 1996–2000.

Figure 6

Deviations of Output From Trend Over Period 1997–2000.



Source: Pinheiro et al. (1997), Banco de Portugal and Ameco.

As in Prescott (2002), we map labor effort h to the average working hours of a working-age person. We used the working-age population (that is, people aged 15–64) and total employment from the Ameco database.

[25] See Pinheiro et al. (1997) for a complete description of this data set.

In order to estimate the capital stock over this period, we used the perpetual inventories method. From the extended *Séries Longas*, we computed the logs of per capita investment, $i_t = \ln I_t$, where I_t is the level of investment at constant prices per working-age person. We then used OLS to compute i_t's linear trend, $\hat{i}_t = a + bt$.

The 1950 capital stock per working-age person was taken to be the capital stock steady-state level consistent with the 1950 steady-state level of investment, that is, $K_{1950} = (\delta + \gamma_{\hat{i}} - 1)^{-1} \exp \hat{i}_{1950}$, where δ is the depreciation rate and $\gamma_{\hat{i}} - 1$ is the steady-state growth rate of investment. The actual value we used was the slope of the previously calculated linear trend. We then used the accumulation relationship between the capital stock and investment to obtain capital stock for the whole period of time:

$$K_{t+1} = (1-\delta) K_t + I_t.$$

Deviations from the real initial capital stock are hampered by the relatively long period of time that goes from 1950 until 1979. We chose a depreciation rate of 6% a year.

Labor effort intensity per working-age person was calculated as follows. From the OECD data on annual hours worked per person employed, L_t, and Ameco's data on working-age population and total employment, W_t and E_t, we estimated the labor effort intensity per working-age person by $h_t = \frac{L_t}{\bar{L}} \frac{E_t}{W_t}$, where $\bar{L}$ is the individual endowment of hours for market activities, which was taken to be 5000 hours, as in Chari et al. (2002).

We estimated the share of capital income in total income over period from the *Repartição do Rendimento Nacional* sheet of Pinheiro et al. (1997), which contains the GNP income components. We estimated α as the average ratio of gross operational surplus to GDP at factor cost over the total time span of the *Séries Longas* database. As in Chari et al. (2002), parameters β and ψ were taken to be 0.97 and 2.24, respectively.

Table 2 reports the values of all parameters of the model.

B Numerical Solution of the Model

We used a three-dimensional grid for the state variables of the model: k, A and τ^h. The random process governing the stochastic vector $z_t = \begin{bmatrix} A_t & \tau_t^h \end{bmatrix}'$, was taken to be a VAR(1) process with zero mean:

$$z_t = D z_{t-1} + \varepsilon_t.$$

We used the detrended wedges $\{A_t, \tau_t^h\}_{t=1979,\ldots,2000}$ to estimate D and the variances-covariances matrix of ε

We calculated the steady state of the model consistent with the data. Table 3 reports the steady-state values of the endogenous variables, as well as the steady-state wedges and growth rates consistent with the data. The investment wedge, $1-\tau^k$, is computed using the steady-state version of the model's Euler equation.

Because business cycle variations during the period under analysis are potentially very high, it is advisable that a rigorous computation of the policy functions be performed. This, however, implies that none of the usual linear approximation methods can be used. Moreover, we must construct a discrete approximation of the VAR(1) process governing the wedges. This was performed using Tauchen's (1986) method. The policy functions were calculated by iterating over $h(k,z)$, the labor decision variable, and recalculating the

remaining decision variables in each iteration. To solve the model, we assumed that government consumption g was exogenous, and allowed it to fluctuate around its steady-state value.

C Additional Figures

Figure 7

Data for Detrended Investment in Portugal and in the Model.

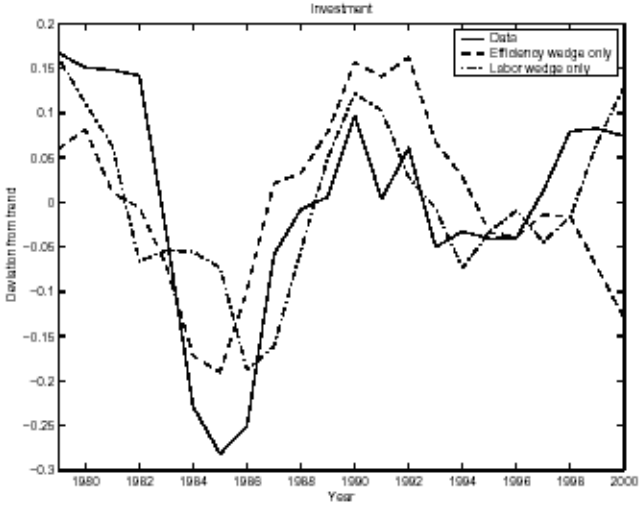


Figure 8

Data for Detrended Labor in Portugal and in the Model.

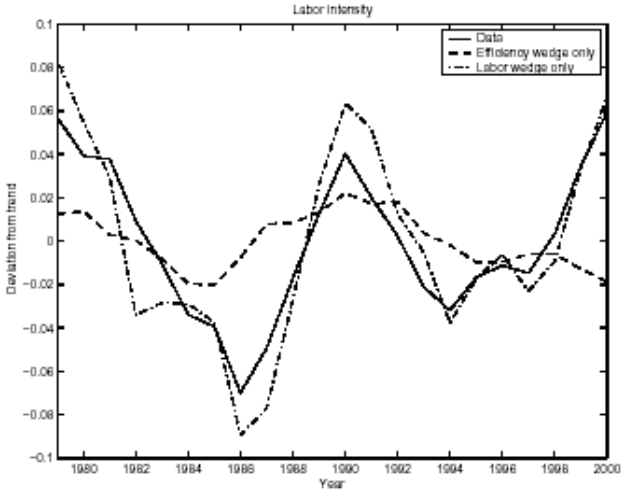


Figure 9

Data for Detrended Investment in Portugal and in the Model (Both Wedges).

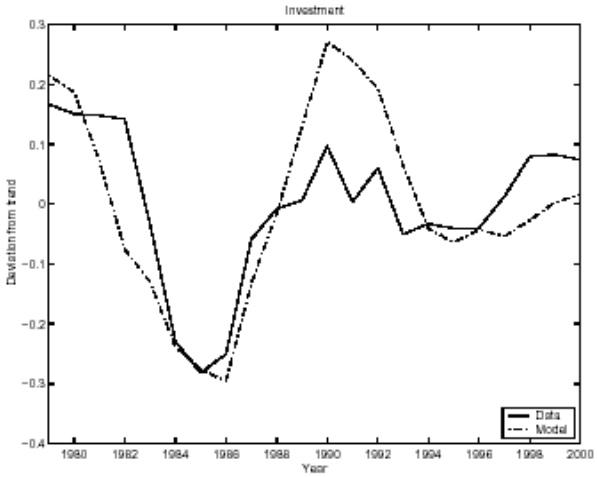
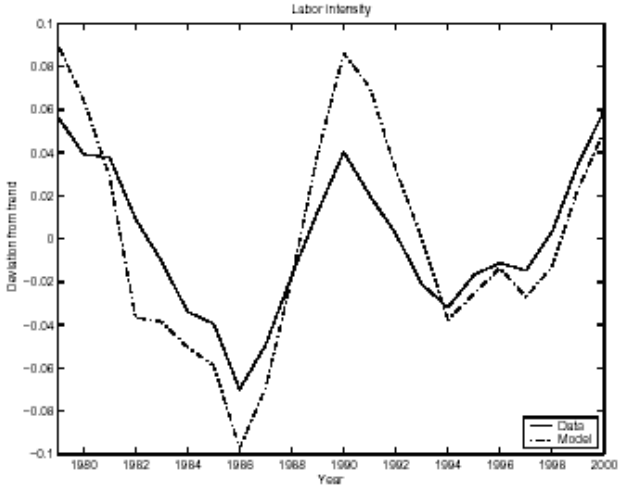


Figure 10

Data for Detrended Labor in Portugal and in the Model (Both Wedges).



References

Alvarez, F. and Veracierto, M. (1999), "Labor market policies in an equilibrium search model", *1999 NBER Macroeconomic Annual* 14, 263-304.

Antunes, A. and Cavalcanti, T. V. (2003), "Corruption, credit market imperfections, and economic development", *Quarterly Review of Economics and Finance* 43(4), 627-642.

Baer, W. and Leite, A. N. (1992), "The peripheral economy, its performance in isolation and with integration: The case of Portugal", *Luso-Brazilian Review* 29(2), 1-43.

Baer, W. and Leite, A. N. (2003), "The economy of Portugal within the european union: 1990-2002", *Quarterly Review of Economics and Finance* 43(5), 738-754.

Bailly, M. N. and Solow, R. M. (2001), "International productivity comparisons built from the firm level", *Journal of Economic Perspective* 15(3), 151-172.

Bertola, G. and Rogerson, R. (1997), "Institutions and labor reallocation", *European Economic Review* 41(6), 1147-1171.

Blanchard, O. and Portugal, P. (2001), "What hides behind an unemployment rate: Comparing Portuguese and U.S. labor markets", *American Economic Review* 91, 187-207.

Chari, V. V., Kehoe, P. J. and McGrattan, E. R. (2002), Business cycle accounting, Mimeo, Federal Reserve Bank of Minneapolis.

Correia, I., Neves, J. C. and Rebelo, S. (1995), "Business cycles in a small open economy", *European Economic Review* 39, 1089-1113.

Gollin, D. (2002), "Getting income shares right", *Journal of Political Economy* 110(2).

Groningen Growth and Development Centre and The Conference Board (2003), Total *Economy Database*. http://www.ggdc.net.

Hayashi, F. and Prescott, E. (2002), "The 1990s in Japan: A lost decade", *Review of Economic Dynamics* 5(1), 206-235.

Heston, A., Summers, R. and Aten, B. (2002), "Penn World Table Version 6.1". Center for International Comparisons at the University of Pensylvania (CICUP).

Lucas, Jr, R. E. (1980), "Methods and problems in business cycle theory", *Journal of Money, Credit, and Banking* 12(4), 696-715.

Parente, S. L. and Prescott, E. (2000), Barriers to Riches, Cambridge, MA. MIT Press.

Pinheiro, M. et al. (1997), "Series longas para a economia portuguesa". Banco de Portugal.

Prescott, E. (2002), "Prosperity and depression", *American Economic Review* 92(2), 1-15.

Tauchen, G. (1986), "Finite state markov-chain approximations to univariate and vector autoregressions", *Economics Letters* 20, 177-181.

SECTORAL REGULATORS AND THE COMPETITION AUTHORITY: WHICH RELATIONSHIP IS BEST?*

Pedro P. Barros
FEUNL, CEPR

Steffen H. Hoernig
FEUNL, CEPR

January 2004

Abstract

Inspired by the creation the new Competition Authority in Portugal, we consider the interplay between regulatory agencies with overlapping competencies, for example a competition authority and a sectoral regulator. We analyse how authorities' incentives are affected if they can decide independently, or must follow each others' opinions, respectively, and consider how this relationship performs in the presence of institutional biases and lobbying efforts. It is found that the best results tend to be achieved when the authorities act independently of each other: Decisions are made more effectively, and are less vulnerable to lobbying.

Keywords: Competition Authority, Sectoral Regulators, Institutional Relationship, Strategic Substitutes and Complements, Lobbying

JEL: L51

* We would like to thank an anonymous referee for his suggestions, which led to considerable improvements in this paper.

1. Introduction

New legal dispositions enacted in May 2003 have established the jurisdiction of the Portuguese Competition Authority (*Autoridade da Concorrência*) over all sectors of economic activity. Unlike previous competition laws, this latest one exempts only legal monopolies, and even these only for the breadth of the concession awarded in the contract establishing the legal monopoly. The broad range of markets covered by the *ex post* control exerted by the Competition Authority clearly creates the scope for conflict with sectoral regulatory authorities. While this possibility is recognized in the definition of powers of the Competition Authority, the boundaries of its power are not clearly defined.

The resulting overlap of jurisdictions implies that either the Competition Authority and the sectoral regulatory authorities reach an agreement about timing and scope of their interventions, or that in some cases intervention by both authorities may occur. The possibility of simultaneous intervention of the sectoral regulators and the Competition Authority has been publicly discussed mainly in legal terms and in relation to the expertise and to the nature of instruments available to each authority. The typical argument runs as follows: The Competition Authority acts as a monitoring entity. It verifies *ex post* whether or not firms had adopted anti-competitive conduct. On the other hand, sectoral regulatory authorities intervene *ex ante*, enacting regulatory measures aimed at avoiding the exercise of monopoly power in regulated markets, for example the setting of high prices or inefficient investment decisions. Alternatively, one could imagine that both authorities are on an equal footing, and each investigates where it sees a need to intervene. Whether the relationship between the Competition Authority and the Sectoral Regulators will be one of control, or of cooperation between equals, is not established in the Law.

We add to these legal and economic considerations a third type of argument the determination of incentives for each authority to intervene, and the informational advantages of two independent views over the same case. The explicit acknowledgment of the incentives facing the relevant economic authorities is, of course, crucial to institutional design, since unforeseen incentive effects may lead to outcomes quite distinct from the intended ones. We address this issue, and implications for institutional design follow. In the following, for simplicity we generically refer to the commitment, objectivity and depth of investigations produced by an economic authority as its "effort", a measure of its input into the decision process.

Our first result states that overlap of jurisdictions leads to lower effort by both authorities, as compared to a single one. This is so irrespective of whether "cooperation" or "competition" between economic authorities prevails. Cooperation, or joint decision making, is intended to mean that to reach a decision both economic authorities must have started and concluded successfully an investigation into the same matter. Thus, it requires both authorities to agree on having a case of anti-competitive behavior. It takes to the extreme the duty of consulting the other economic authority, turning the opinions of one authority binding to the other (from the Competition Authority to the sectoral regulator or vice-versa).

Under competition, or independent decision making, a case is closed as soon as one of the authorities concludes successfully the investigation, irrespective of whether the other authority has come to a conclusion or even opened a case. We find that if each authority's equilibrium probability of finding anti-competitive behavior is larger than $1/2$,

then independent interventions, even if overlapping, induce lower efforts by each authority than joint decisions (and vice-versa if this probability is less than $1/2$). Still, and more importantly, the probability that anti-competitive behavior is not challenged by economic authorities is always lower under independent decisions, and total welfare is higher.

Therefore, a first implication is that institutional design should allow for "competition" between economic authorities, based on their independence from each other, instead of forcing a joint decision process. Naturally, apart from and on top of the incentive problem focussed on above, the institutional framework of joint decisions would most probably involve higher administrative costs, anyway. These result from the need to accommodate the different cultures and views of each economic authority.

There is, nonetheless, a third option to be considered - a protocol defining that one authority should decide first, with the other authority controlling the latter: Only when the latter does not intervene, will the other economic authority be able to step in with its own investigation.[1] We find that this arrangement strongly distorts downward the incentives of the authority to be controlled, while the other authority, once it steps in, investigates as if it were alone. As a result, welfare is lower than under independent decisions.

The arguments reported above rely on the assumption that economic authorities value consumer surplus and firms' profits equally, a rather common assumption in the economics literature. However, according to public statements, the Competition Authority essentially values consumer surplus. On the other hand, the sectoral regulatory authorities tend to consider both consumer surplus and firms' profits, though not necessarily in an equal manner. In the case where at least one economic authority has a bias in its objective function (by comparison with the social welfare standard), the optimum allocation of efforts cannot be achieved.

Interestingly, a bias of one economic authority cannot be fully compensated by giving the other authority a corrective bias. If the authorities decide jointly, an opposite bias of the other authority raises welfare to the second-best but does not restore the first-best: If the legislator considers sectoral regulators to be more industry-oriented than socially optimal, then the Competition Authority should have a pro-consumers' bias, even though this is not as good as having both economic authorities unbiased. Surprisingly, if authorities decide independently then the second-best is achieved by giving the other authority the *same* bias as the first authority. The intuition behind these results is that the bias of one authority distorts the incentives of the other authority. In the case of competing authorities this distortion is best corrected by making authorities' biases more similar rather than more different to avoid excessive investigations by the other authority.

A final, and related, issue is regulatory capture. Typically, it is believed that sectoral regulators are more easily captured by firms than a general Competition Authority. Therefore it may be argued that one should rely more on the Competition Authority to ensure a level playing field in the marketplace. Still, the existence of two economic authorities with jurisdiction over the same market may render capture by firms harder. Therefore, it is of interest to address the issue of whether or not overlapping economic authorities are an instrument against regulatory capture. Though this intuition seems reasonable, one must not forget the incentives of economic authorities to intervene, which may lead to a different conclusion.

---

[1] It could be determined that in a specific set of situations one authority acts first, while in all other cases the other authority takes precedence.

To this conventional wisdom we add that it does matter how the economic authorities interact with each other. If a joint decision framework is imposed, by concentrating lobbying efforts on the more industry-friendly economic authority firms can quite effectively deter the launch of investigations and a final decision contrary to firms' interests. This problem does not exist in the case of independent decisions. Thus, whenever the possibility of capture exists, our analysis suggests that institutional design must allow for intervention of both authorities, without requiring a joint decision. This implies that opinions that one authority seeks from the other should not be made binding.

Overall, to sum up our contribution: Having two economic authorities with overlapping jurisdictions, one taking care of sector-specific regulation, and a general Competition Authority, can often be welfare improving. Consequently, the recent changes in competition policy in Portugal, extending the application of competition law to all sectors of economic activity, do not imply or require, in terms of the underlying economics, the dismissal of sectoral regulators on matters of anti-competitive behavior of firms.

Although some cooperation between authorities may be desirable (in the sharing of information, sector-specific or technical expertise, for example), it would be problematic to impose a joint decision process, where a final decision on a case would be reached only under agreement of both authorities. Retaining independence in decisions is key to achieving the beneficial effects of overlapping authorities.

In the following section we start with an overview of recent legislation in Portugal and public statements by some of the actors involved to shed light on what are the intentions of legislators and regulators. Section sets out our model, and Section discusses the relation between the authorities. In Sections and we consider biases and regulatory capture. Section presents some extensions of the basic model, while Section concludes.

2. Institutional Background

The sectoral regulatory authorities (SRAs) in Portugal are the following, in alphabetical order: ANACOM (telecommunications), Banco de Portugal (banking sector), CMVM (stock markets), ERSE (energy), IMOPPI (construction), INAC (Civil air transport), INTF (railways), IRAR (fresh and waste water), ISP (insurance). Of these ANACOM, Banco de Portugal e ERSE are independent. The new Competition Authority (CA), Autoridade da Concorrência, officially started its activity on March 24th, 2003, after it had been created by a law passed in 2002 (*Lei n.º 24/2002 de 31 de Outubro*) and a decree-law of 2003 (*Decreto-Lei n.º 10/2003, de 18 de Janeiro*) which set out its statutes. In June 2003 the new Competition Law came into effect (*Lei 18/2003, de 11 de Junho*), with the intention of harmonizing Portuguese legislation with EU law. All of these documents, and others, mention to various degree the relation with the SRAs. The following is meant to give a quick overview of the most important points (all translations are the authors', and not necessarily precise), notwithstanding our modest legal knowledge.

- *Lei n.º 24/2002 de 31 de Outubro*, allowing the government to create the CA;Article 2 g) “The Authority will establish the necessary collaboration (“conveniente articulação”) between its activities and the activities of the sectoral regulatory authorities, and will receive from firms and public

authorities the information and cooperation necessary for the fulfillment of its purpose."[2]

- *Decreto-Lei n.º 10/2003, de 18 de Janeiro*, containing the Statutes of the CA as an annex. The introduction mentions the "desirable and necessary collaboration with the respective sectoral regulatory authorities".[3] Its article 6 deals exclusively with the relation to the SRAs. In its points 2 and 3 it states that future law [the new competition law, see below] will define the "modes of intervention" and "participation" of the CA in areas subject to sectoral regulation, "as much as necessary in the interest of the objectives of the competition legislation", and will define the "obligations of the sectoral regulators with respect to anti-competitive practices that they come to know about in their activities", as well as the "collaboration with the authority on issues subject to sectoral regulation", while point 4 enumerates the most important SRAs.

- The CA's strategy document "*Estratégia de Desenvolvimento (2003-2005)*": Two of the five main points of action identified are concerned with the relation with the SRAs (p. 16-18). On p. 17 it is stated that "The fundamental principle of work in the relations between the different regulatory authorities is that it is the [Competition] Authority that takes on the responsibility of the evaluation of competition, structure and behavior of markets, while the sectoral authority takes on the responsibility the technical aspects of the sector."[4] It is acknowledged that this principle needs to be adapted and made more precise for each sector, and it is proposed to sign "protocols" between the CA and the respective SRA. Furthermore, it is stated that the CA will make recommendations and try to influence the regulatory process so that restrictions on competitions are avoided.

- *Lei 18/2003, de 11 de Junho 2003*, the new Competition Law: Sets out the powers of the CA, which essentially will deal with the enforcement of the rules of competition. The articles dealing with the relation to the SRAs are:

  – Article 15: CA e SRAs "collaborate in the application of the competition legislation" in the terms described by the following Articles about processes on forbidden practices (27, 28, 29) and on concentrations (39);

  – Article 27, no. 4: Provisional measures (medidas cautelares) can only be imposed after the SRA states its evaluation ("*parecer prévio*"), for which it has 5 working days.

---

2 *Artigo 2 g) "A Autoridade procederá à conveniente articulação das suas actividades com a das autoridades reguladores sectoriais e receberá das empresas e das autoridades públicas as informações e a cooperação necessárias ao cabal desempenho das suas atribuições."*

3 *"…desejável e necessária articulação com as respectivas autoridades reguladoras sectoriais".*

4 *"O princípio fundamental de trabalho nas relações entre as diferentes agências reguladoras é o de que a Autoridade se responsabiliza pela avaliação da concorrência, estrutura e comportamento dos mercados, enquanto a agência sectorial se responsabiliza pelos aspectos técnicos do sector."*

- Article 28, no. 2: Any decision other than closing the case can only be imposed after the SRA states its evaluation ("*parecer prévio*"), for which the CA fixes a deadline.

- Article 29, Collaboration ("*Articulação*") with the sectoral regulatory authorities.

  no. 1: The CA informs the SRA immediately about any suspicion, and the SRA has to comment.

  no. 2: The SRA informs the CA immediately about possible violations of the Law of Competition, and supplies material for investigation.

  no. 3: Given no. 1 and 2 the CA can put on hold own investigations.

  no. 4: Before a final decision the SRA informs the CA, so that CA can respond within a time limit set by itself.

- Article 39, no. 1: The CA before taking a decision asks the SRA for its opinion.no. 2: The exercise of powers of the SRA are not affected by no. 1.

- It is interesting to note that the CA's powers were reduced as compared to an earlier version of the law: By Article 29, no. 3 it could have sent an investigator to accompany permanently at the SRA processes about forbidden practices. This clause has been removed in the final version.

- As the Competition Law was discussed in the Portuguese Parliament, the Minister of Economics, Carlos Tavares, said in his speech to the Parliament: "...Finally, we underline the clear definition of the terms of collaboration between the Competition Authority and the Sectoral Regulatory Authorities, which constitutes a clear option for advantages of a close collaboration between the two types of entities, while not putting into question the respective own competencies."[5]

- In his intervention during the presentation of the new Directive Council of ANACOM in July 2002, the Minister of Economics, said that "In the area of competition – a crucial question in the sector – ANACOM must collaborate closely ["*...articular e colaborar estreitamente...*"] with the Competition Authority, to be created in the near future and whose role is nor to be

[5] *"Finalmente, saliente-se a definição clara dos termos da articulação entre a Autoridade da Concorrência e as autoridades reguladoras sectoriais, o que constitui uma aposta clara nas virtualidades de uma cooperação estreita entre os dois tipos de entidades, sem que com isso se ponham em causa as respectivas competências próprias." in "Intervenção do Ministro da Economia sobre o novo regime jurídico da Concorrência" (13 de Fevereiro de 2003, Assembleia da República), em http://www.portugal.gov.pt.*

confused with, nor to be superimposed ["*...nem se sobrepõe...*"] on the role of the sectoral regulators".[6]

- In a recent interview (*Público, Suplemento de Economia*, June 2nd, 2003), the president of the new CA, Abel Mateus, is quoted saying that the law did not specify all aspects of the relation between CA and SRAs, but that "In competition issues the CA must prevail", while exhorting the SRAs to show an "attitude of large cooperation" ("*uma atitude de grande cooperação*"). He also said that his authority would check the government proposals for the liberalization of the electricity market. In Semanário Económico, on April 10th 2003, he is quoted as assuring that the CA would not deal with tariffs established by regulators (here electricity tariffs set by ERSE), but with the underlying behavior of the firms involved. This is because the SRA has the necessary technical expertise which the CA has not. As concerns the possible opening of the network and/or ownership of the Portuguese cable TV operator TVCabo to competitors in telecommunications, at a different occasion he is quoted as saying "...We are working together with ANACOM [telecoms regulator], but we will have to take the decision".[7]

- Finally, at a dinner-debate of APDC in June 2003 Âlvaro Dâmaso, the president of ANACOM, said that it would be necessary to develop concerted modes of action ("*modos de actuação concertados*"), to avoid "superpositions or omissions in the activities" of the two institutions responsible for the sector.[8]

It seems that the different laws enumerated above, and several statements by the Minister of Economics, point to a relationship equal-to-equal between the Authority of Competition and the SRAs, even if the CA may have more powers in communication between the authorities. It also seems clear that the law does not strictly delimit the boundaries of responsibilities between the CA and the SRA. They do have to inform each other when they know of facts that of interest to the other authority, and can be investigating in parallel; they have to ask for the other authority's opinion before coming to a decision, but it is not stated what will happen if they disagree.

On the other hand, the direction of the statements coming from the CA is quite different and does not only point to the idea of a clear separation of activities, but even to an idea of superiority of the CA over the SRAs. The future will show which point of view will prevail in practice. It is also worth mentioning that the CA and several of the SRAs have already expressed interest in establishing protocols governing the respective dominions of action and spelling out the details of their relationship.

---

6 *"Intervenção do Ministro da Economia na posse do Conselho de Administração da Autoridade Nacional de Comunicações" (8 de Julho de 2002), em http://www.portugal.gov.pt.*

7 *"...estamos a trabalhar em conjunto com a ANACOM, mas nós é que vamos ter de tomar a decisão", Semanário Económico*, May 9th, 2003.

8 *Jantar-debate da APDC, 17 de Junho, texto em http://www.apdc.pt.*

3. The Basic Setup

The aim of the government is to maximize social welfare originating in certain markets, $W = S + \Pi$, for which it uses two regulatory authorities, 1 and 2. These authorities, have an $ex-aequo,$ non-hierarchical relationship, unless explicitly stated otherwise. S and Π are consumer surplus and industry profits, respectively. Apart from section we will assume that both authorities are not functionally different.[9] Furthermore, since we are mainly interested in how their decisions related we assume that they do not exchange information to not distract from our main point. If information were exchanged our results comparing different decision processes would still hold, while welfare would simply be higher.

For each case it considers, authority i has a utility of

$$U = S + \lambda\Pi, \tag{1}$$

where λ is authority i's "bias" with respect to this case.[10] For $\lambda = 1$ the authority is neutral and maximizes total welfare, while for $\lambda > (<) 1$ it is industry - (consumer-) friendly. That is, in both the latter cases the authority does not try to implement the social optimum. This bias may spring from the authority's rules, from the choice of personnel,[11] or may be the result of lobbying. The latter case is especially important, and will be dealt with in section.

Each market under scrutiny is characterized by an unobservable state of the world taking one of two values, and observable behavior of firms that the authorities investigate.

For example:

- Marginal production cost of all firms can be low or high; the authorities investigate whether high prices stem from collusion;
- Costs of giving access to a network can be low or high; the authorities investigate whether high access prices are cost-based;
- A proposed merger raises welfare or not; the authorities investigate whether the merger should be allowed.

Let us call the states of the world "violation" and "no violation" (of competition law or regulatory rules). The probability of the "violation"-state occurring $\pi \in (0,1)$ is exogenous and common knowledge. An investigation is successful if and only if the authorities come to a decision about the true value of the underlying state of nature. In this case remedies will be imposed if the state of nature is "violation", and nothing will be done

9 We are aware that authorities may intervene at different times and with different cost, but since we are addressing the incentives to intervene a useful benchmark is identical intervention. Qualitatively the results would not differ much if we assumed that authorities were somewhat different.

10 This is the traditional weighted welfare function of partial equilibrium analysis, with the weight λ on profits indicating the bias. Other formulations would lead to qualitatively similar results.

11 For example, the first director of the new Portuguese Competition Authority stated in an interview that he was principally interested in preserving consumer surplus, not in firms' stock market value (*Público, 2003*).

if the state is "no violation". If on the other hand they do come not to a decision then the *status quo* is maintained on the presumption of no violation.[12]

Each market will then be characterized by three pairs of consumer surplus and profits:

1. (S_V, Π_V), with index V for "violation" of the rules of competition; that is, high prices when costs are low, or a welfare-decreasing merger;

2. (S_R, Π_R), with index R for "remedies": low prices with low cost, or no merger if welfare decreases; we naturally assume $S_R > S_V$, $\Pi_R < \Pi_V$ and $S_R + \Pi_R > S_V + \Pi_V$, or

$$\Delta_S = S_R - S_V > \Delta_\Pi = \Pi_V - \Pi_R > 0 \tag{2}$$

3. (S_N, Π_N), with index N for "no violation": high prices if high costs, or merger if welfare increases.

Authorities 1 and 2 spend (or exert effort measured in terms of units of consumer surplus) $e_1, e_2 \geq 0$ to resolve a given case, which leads to a probability $P(e_1, e_2) \in [0,1]$ of "success", independently of the true state of the world. "Success" is supposed to mean that enough verifiable data has been collected so that either an enforceable decision can be made, or it is concluded that there has been no violation of the rules. "No success" in this interpretation means that the authorities' investigations were inconclusive, in which case the *status quo* is maintained. The probability P is assumed to be twice differentiable and concave, increasing in each effort, and with $P(0,0) = 0$ As we will see below, its economic properties are determined by the institutional relationship between the two authorities.

Each authority's effort devoted to a given case will be determined by maximizing its utility with respect to this case.[13] The first decision to be made is whether the case is to be investigated at all, which is the case if authority i's utility, but not necessarily total welfare, under the remedy is higher than in the "violation" state. That is, if

$$U_{iR} = S_R + \lambda_i \Pi_R > U_{iV} = S_V + \lambda_i \Pi_V,$$

or

$$\Delta_i = U_{iR} - U_{iV} = \Delta_S - \lambda_i \Delta_\Pi > 0.$$

An industry-friendly authority may be so biased that it perceives that there are no adverse welfare effects ($\Delta_i < 0$) even though the case should be investigated ($\Delta = \Delta_S - \Delta_\Pi > 0$), and it will therefore sometimes not even start investigations which could lead to an increase in welfare. In this case disagreement may appear if the other authority starts and successfully finishes investigations. Whether a decision will emanate from these investigations then depends on how the relations between the two authorities are structured.

---

12 Implicitly, we here assume that the authorities do not impose "remedies" by mistake since they can infer the true state of the world.

13 Below we will also consider budget constraints. These constraints intertwine all cases at hand, but do not qualitatively change the results unless preferences over outcomes differ widely.

Given the other authority's effort e_j, if $\Delta_i > 0$ then the optimal effort of authority i (its "best response") is determined through the maximization of its expected utility,

$$\begin{aligned}\max_{e_i \geq 0} E[U_i] = P(e_i, e_j)\left(\pi U_{iR} + (1-\pi)U_{iN}\right) \\ + \left(1 - P(e_i, e_j)\right)\left(\pi U_{iV} + (1-\pi)U_{iN}\right) - e_i.\end{aligned} \tag{3}$$

This objective function can be rewritten as

$$E[U_i] = P(e_i, e_j)\pi\Delta_i - e_i + \left[\pi U_{iV} + (1-\pi)U_{iN}\right].$$

Leaving out the constant terms in square brackets, the problem reduces to

$$\max_{e_i \geq 0} P(e_i, e_j)\pi\Delta_i - e_i, \tag{4}$$

the expected gross increase in welfare due to the authority's investigation minus the cost of investigation. The necessary first-order condition is also sufficient for a global maximum because P is concave, and defines the best response $\hat{e}_i(e_j)$ as

$$\frac{\partial P}{\partial e_i}\left(\hat{e}_i(e_j), e_j\right) = 1/\pi\Delta_i \text{ if } \frac{\partial P}{\partial e_i}(0, e_j) \geq 1/\pi\Delta_i, \tag{5}$$

$$\hat{e}_i(e_j) = 0 \text{ if } \frac{\partial P}{\partial e_i}(0, e_j) < 1/\pi\Delta_i. \tag{6}$$

That is, effort is zero if the expected gain $\pi\Delta_i$ from a successful investigation is too small; otherwise effort is positive. We immediately arrive at the following comparative statics results:

Proposition 1 *Authority i's optimal effort is*

1. *increasing in the probability that observed behavior constitutes a violation of the rules of competition π, and in the utility differential Δ_i. It is decreasing in its bias λ_i.*

2. *increasing (decreasing) in the other authority's effort e_j if $\partial^2 P / \partial e_i \partial e_j > (<)\, 0$.*

Proof. In the objective function (4) e_i has increasing differences with π and Δ_i, and increasing or decreasing differences with e_j depending on the signal of $\partial^2 P / \partial e_i \partial e_j$. ■

Depending on the cross-derivative of P, authorities' efforts are strategic complements or substitutes; in section we show how this originates from the institutional relationship. Furthermore, any increase in Δ_S or decrease in Δ_Π leads to more effort, while any increase in λ_i, for example due to lobbying by firms, decreases effort.

The outcomes in the authorities' game are the Nash equilibria (e_1^*, e_2^*) given by the simultaneous solutions to each authority's maximization problem. How these equilibria depend on the parameters, in particular the biases (λ_1, λ_2), will be discussed below in Sections 4 and 5.

4. The Relation between Authorities

In this section we focus on the definition of the probability of coming to a decision P. Assume that each authority has an individual probability $p_i(e_i)$ of completing its part of the investigation successfully. The dependence of total success P on p_i depends on how the relations between the authorities are structured. We consider here two extreme cases of most interest if the relationship between both authorities is one of equals:

1. *Joint Decision (JD)*: A case comes to a decision if and only if both authorities have started and successfully completed their investigation. Thus:

$$P^{JD}(e_1,e_2) = P_2 = p_1(e_1)\,p_2(e_2). \tag{7}$$

2. *Independent Decisions (ID)*: The case is decided if either one of the authorities has started and successfully concluded its investigation:

$$\begin{aligned} P^{ID}(e_1,e_2) &= 1-\big(1-p_1(e_1)\big)\big(1-p_2(e_2)\big) \\ &= p_1(e_1)+p_2(e_2)-p_1(e_1)\,p_2(e_2). \end{aligned} \tag{8}$$

Intermediate cases can be modelled through different forms of P, for example the case where an investigation can be concluded if one authority passes on information that can be inconclusive in itself, but may be combined with material existing in the other authority to allow for a decision.

For further reference we will also define where a hierarchical decision process, where without loss of generality authority 1 only investigates if authority 2 has not previously come to a decision (and authority 1 is notified of this, which we assume). In this sense the second authority controls the work of the first one, by stepping in if it seems that the investigation has not been diligent enough.

3. *Hierarchical Decisions (HD)*: Authority 2 investigates first, and authority 1 investigates afterwards if and only if there was no decision at the first stage:[14]

$$\begin{aligned} P^{HD}(e_1,e_2) &= p_2(e_2)+\big(1-p(e_2)\big)\,p_1(e_1) \\ &= p_1(e_1)+p_2(e_2)-p_1(e_1)\,p_2(e_2). \end{aligned} \tag{9}$$

It is remarkable that the expression for the resulting probability of success is identical to the one under independent decisions, for any (e_1,e_2); nevertheless the resulting equilibrium effort levels will be different because the sequence of moves and informational assumptions.

Important observations are the following:

[14] The logic of the resulting game is analogous to the well-known Stackelberg model.

Proposition 2

1. *Under Joint Decisions efforts are strategic complements (if one authority increases effort the other follows), while under Independent Decisions efforts are strategic substitutes (if one authority increases effort the other works less as a result).*

2. *Under both Joint and Independent Decisions each authority exerts less effort than as if it was investigating alone. If the probability that the other agency concludes successfully its part of the investigation is smaller (larger) than* $1/2$ *, more (less) individual effort is exerted under Independent Decisions than under Joint Decisions.*

3. *If in (symmetric and stable) Nash equilibrium the chosen individual probability of concluding the investigation is larger (smaller) than* $1/2$*, then efforts are higher (lower) under Joint Decisions than under Independent Decisions.*

4. Under Hierarchical Decisions authority 1 selects the effort level that would be optimal for a single authority, but authority 2 investigates even less than under Independent Decisions.

Proof. Under JD we obtain

$$\frac{\partial P^{JD}}{\partial e_1} = p_1' p_2, \ \frac{\partial^2 P^{JD}}{\partial e_1 \partial e_2} = p_1' p_2', \tag{10}$$

and under ID

$$\frac{\partial P^{ID}}{\partial e_1} = p_1'(1-p_2), \ \frac{\partial^2 P^{ID}}{\partial e_1 \partial e_2} = -p_1' p_2', \tag{11}$$

We see that $\partial^2 P^{JD}/\partial e_1 \partial e_2 > 0$ and $\partial^2 P^{ID}/\partial e_1 \partial e_2 < 0$, which proves the first statement.

As for the second one, the first-order conditions are $p_1' p_2 = 1/\pi\Delta_1$, and $p_1'(1-p_2) = 1/\pi\Delta_1$, which lead to a higher p_1' at the solution than in the single-authority first-order condition $p_1' = 1/\pi\Delta_1$, because $p_2 \in [0,1]$. Then the optimal e_1 must be smaller because p_1 is concave. Now for the second part of this statement: Comparing the first-order conditions we see that the best responses under the two decision processes cross at $p_j = 1/2$, where they have the same value because $p_j = 1 - p_j$. Since ID best responses are decreasing and the JD ones increasing, this implies that for $p_j < 1/2$ ID best responses are larger, while for $p_j > 1/2$ it is the JD ones who are larger.

The third statement follows from the observation that the symmetric Nash equilibria are found where best responses cut the diagonal (and since the are stable the ID (JD) best response cuts the diagonal from above (below)). In the equilibrium probabilities $p_1 = p_2$ are less than $1/2$ under ID, then the intersection of the two types of best responses at $p_j = 1/2$ lies below the diagonal. Since the JD best response remains below the ID best

response to the left of this point, this implies that the JD best response cuts the diagonal below the ID Nash equilibrium. Furthermore, the opposite holds if the ID Nash equilibrium involves $p_1 = p_2 > 1/2$.

Now consider the fourth statement: The most common method of solving games where players move at different points in time is backward induction, that is, solving the game backwards. If stage 2 is reached then all that is relevant for authority 1 is that not decision has been made yet, so that this authority is in the same situation as if it were alone. Therefore is chooses the optimal effort level e^* of a single authority, leading to a probability of success, say, p_1^*. At the first stage, authority 2 foresees the other authority's future choice, and selects its effort to maximize its utility

$$\max_{e_2 \geq 0} P^{HD}\left(e^*, e_2\right)\pi\Delta_2 - e_2.$$

Since the functional form of P^{HD} and P^{ID} are identical, the optimal choice e_2^{HD} is decreasing in e^*. Since e^* is larger than the equilibrium effort under ID, e_2^{HD} must be smaller.■

It must be noted, though, that the amount of effort, the input in investigation, is not what counts in the end, but the output: the total probability of a decision, and the resulting welfare. In this respect, more effort under a given institutional design does indeed lead to a higher probability of success, but this probability may be higher under a different design involving less effort. On the other hand, high efforts may even be excessive from the point of view of total welfare, so that too much investigation is not welfare maximizing, either.

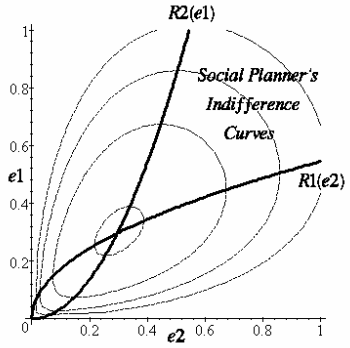
Furthermore, in the simple hierarchical relationship defined above, the authority that is being controlled has much lower incentives to do its job, since it knows that the other authority will substitute it. Thus instead of making it work better, control actually makes it work less.

The fundamental results contained in the above proposition tell us that the institutional relations between both authorities matter much. Under JD each authority puts in effort if the other does so, but may also almost neglect a case if the other one does – the essence of strategic complementarity. On the other hand, under ID we have strategic substitutes: An authority lays back if the other authority seriously deals with a case, while it will step in if the other authority does not work hard enough.

As an illustration, consider the following example: Let $p_i(e_i) = \min\left\{e_i^{1/3}, 1\right\}$ and $\pi\Delta = 2$. The following Figures 1 and 2 contain the social planner's indifference curves (thin lines) and the authorities' best responses (thick lines) when they are unbiased; best responses intersect precisely at the social optimum. A consumer-friendly bias will shift a best response upwards, while an industry-friendly bias shifts it downwards. In Figure 1 we see the representation of Joint Decisions, which lead to increasing best responses. Since in this example the slope of p_i at zero is infinite, best responses are increasing at zero and there are exactly two Nash equilibria: the socially optimal one, and the zero-effort equilibrium. Figure 2 depicts the case of Independent Decisions, which result in a completely different picture. Here authorities' efforts crowd each other out, and there is no zero-effort equilibrium.

Figure 1

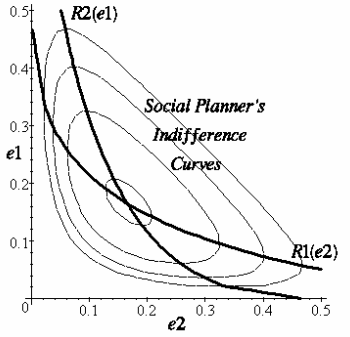
Unbiased Best Responses and Indifference Curves with Joint Decisions



$(p_i(e) = e^{1/3}, \pi\Delta = 2)$

Figure 2

Unbiased Best Responses and Indifference Curves with Independent Decisions



$(p_i(e) = e^{1/3}, \pi\Delta = 2)$

Still, both authorities exert *less effort* than if they were working alone. The reason is that the marginal value of effort is lower, for two different reasons: Under ID because the other authority may solve the case anyway, so the additional effort could be wasted; under JD because the other authority may not be able to successfully conclude its own investigation, and as a result any effort would be wasted. Yet, this does not mean that the *sum* of efforts is necessarily lower or higher than the effort a single authority would put in.

Since p is concave there are decreasing returns to scale in investigation, which favors a higher effort level by each of the authorities - the total result is ambiguous.

More important than efforts are the equilibrium probability P of coming to a decision, and the ensuing comparison of total welfare. Here we find the following:

Proposition 3 *Assume that both authorities are unbiased.*

1. *Under Independent Decisions, the Nash Equilibrium probability* $P\left(e_1^*, e_2^*\right)$ *of a successful investigation is higher than with a single authority, while it is even lower than the latter under Joint Decisions.*

2. *Total welfare is higher under Independent Decisions than under both Joint Decisions and Hierarchical Decisions.*

Proof. Note first that each authority's maximization problem is equivalent to maximizing $\pi\Delta P\left(e_1, e_2\right) - e_1 - e_2$ with respect to its own decision variable e_i since authorities are unbiased; therefore the Nash equilibrium coincides with the outcome of the joint maximization over both variables. Second, note that for *any* given pair $\left(e_1, e_2\right)$ we have the following chain of inequalities, letting $p_i = p_i\left(e_i\right)$,

$$\begin{aligned} 1-\left(1-p_1\right)\left(1-p_2\right) &= p_1 + p_2 - p_1 p_2 \\ &\geq p_i \\ &\geq p_1 p_2, \end{aligned}$$

which are the success probabilities under ID, one authority, and JD, respectively, at identical effort levels. Furthermore, these inequalities are strict for positive levels of effort, i.e. $p_1, p_2 > 0$. Consider the following maximization problem:

$$P(\lambda) = \max_{e_1, e_2} \lambda\left[1-\left(1-p_1\right)\left(1-p_2\right)\right] + \left(1-\lambda\right) p_i.$$

Clearly $P(0)$ and $P(1)$ correspond to the equilibrium probabilities of successful investigation with a single authority, and under Independent Decisions, respectively. By the envelope theorem,

$$\frac{dP(\lambda)}{d\lambda} = \left[1-\left(1-p_1\right)\left(1-p_2\right)\right] - p_i.$$

By the above inequalities this is non-negative for all $\lambda \in [0,1]$, so that P is non-decreasing in λ. Furthermore, since for almost all λ the solution involves positive effort levels the inequalities become strict, and we have that $P(1) > P(0)$. The corresponding argument applies to the comparison between a single authority and Joint Decisions, and to the welfare comparison between Independent and Joint Decisions.

As for HD, we have already shown above that the equilibrium effort levels are different from the optimal ones under ID. Since the functional form of the welfare objective under HD and ID are the same, the HD effort levels must yield less welfare than the ID ones.■

We thus see that Hierarchical Decisions lead to lower welfare than Independent Decisions; it does not matter much that we cannot directly compare the probabilities of success in this case. The main effect involved was mentioned above: The control relationship distorts the incentives of the authority subject to control in the "wrong" direction.

From the above arguments it does not follow, though, that welfare is necessarily higher under Independent Decisions with *two* authorities than with just one single authority. The question whether there should be two authorities or just one is straightforward to formalize if one only takes into account the variable and fixed costs of running authorities, but less so if we take into account other arguments, to be considered below.

If we assume that running each authority involves a fixed cost F, then the total expected increase in welfare achieved with one authority is

$$\Delta W^1 = \Sigma_k \left[p(e_k)\pi\Delta_k - e_k \right] - F, \tag{12}$$

where the sum is over the cases to be analyzed, and $\Delta_k = \Delta_{Sk} + \Delta_{\Pi k}$.

If responsibilities of the authorities are completely separated, i.e. if there is no overlap in their responsibilities, then the duplication of fixed cost must be compensated by gains in specialization. These could arise from accumulated expertise in certain types of cases, for example mergers on the one hand, and access pricing on the other hand. On the other hand, with two authorities and overlapping responsibilities we have

$$\Delta W^2 = \Sigma_k \left[P(e_{1k}^*, e_{2k}^*)\pi\Delta_k - e_{1k}^* - e_{2k}^* \right] - 2F, \tag{13}$$

where stars denote equilibrium actions. Clearly in terms of cost it is better to have two authorities if $\Delta W^2 > \Delta W^1$ this may happen if fixed costs are sufficiently low, and more importantly, if there are strongly decreasing returns to scale in investigation, that is if p_i is relatively more concave than P. In the latter case e_k may be larger than $e_{1k}^* + e_{2k}^*$ while the probability of success is smaller. Following the above proposition, clearly this can only happen with Independent Decisions. That is, if Joint Decisions are adopted then advantages other than cost reduction must come into play which are not modelled here.

The most important alternative approach, as in Laffont and Tirole (1993) and Martimort and Laffont (1999) assumes that regulators' objectives only include direct monetary benefits, not social targets. Therefore their model is couched in terms of "bribes" from firms to regulators, and incentive payments to regulators, in order to make the latter refrain from taking these bribes. Consequently, having two regulators, who both have their own signals of the relevant information, instead of just one, lowers the value of the bribes that firms are willing to pay. This leads to a reduction in incentive pay, which may be enough to cover the cost of having an additional regulator.

Apart from cost-based ones there are other arguments to be taken into account which can affect the trade-off analyzed above:

- An important issue is lobbying, to be dealt with in section, because the institutional arrangements between authorities may alleviate or aggravate the problem of capture by interest groups.

- One authority may have the implicit or explicit role to control the decisions of the other authority. Such will be the case in Portugal, where the new Competition Authority has been assigned to review the decisions by the SRAs that affect competition. In this case one may argue that the relation between authorities is hierarchical rather than equal-to-equal.

- If, as is also the case in Portugal, one authority (the CA) supervises competition many different markets, and each SRA different aspects in its assigned regulated market, then their roles are different, and new problems can arise. For example, the delineation of responsibilities may not be clear, or the CA's activities in various markets may overstretch its resources.

5. Biased Authorities

In this section we model authorities as biased maximizers of welfare, in the spirit of Olsen (1965) and Stigler (1971). These biases may stem from various sources: An orientation towards maximizing consumer surplus may explicitly be mentioned an authority's mission statement; the personality, education and employment history of the authority' directors will affect the closeness to industry or specific firms. While the first bias is institutional, the second one can at infrequent moments in time be influenced by the government, through the choice of new directors.

We will consider, first, the welfare effects of biased authorities; second, how the bias of an authority should be designed to correct for a bias of the other authority. The point of departure is the following observation:

Lemma 4 *If the authorities are not biased* $(\lambda_1 = \lambda_2 = 1)$ *then the socially optimal levels of effort given the relation between the authorities constitute a (stable) equilibrium.*

Proof. The government's objective function is

$$TW = \Sigma_k \left[P\left(e_{1k}, e_{2k}\right)\pi\Delta_k - e_{1k} - e_{2k} \right] - 2F,$$

and the socially optimal effort levels are found through its first-order conditions $\pi\Delta_k \partial P\left(e_{ik}, e_{jk}\right)/\partial e_{ik} = 1$ and $\pi\Delta_k \partial P\left(e_{ik}, e_{jk}\right)/\partial e_{jk} = 1$. Since these are exactly the first-order conditions of the unbiased authorities, at the socially optimal effort levels e_{jk}^S and e_{ik}^S of the other authority each authority also chooses its socially optimal effort level. This equilibrium is stable because the best responses cross from above.■

This result means that, given the institutional relations captured by P, unbiased authorities do decentralize the social optimum, and constitutes a benchmark for the analysis of the behavior of biased authorities. However, a *caveat* is necessary: This lemma does not rule out other equilibria. Indeed, with JD there always exists an additional equilibrium where both authorities exert zero effort. In this equilibrium none of the authorities will invest in a case because each expects the other authority, whose collaboration it needs, to not investigate. This equilibrium is unstable if $\frac{\partial P}{\partial e_i}(0,0) \geq 1/\pi\Delta_i$ because the slightest increase in the other authority's effort will meet a positive response. On the other hand it is stable if

$\frac{\partial P}{\partial e_i}(0,0) < 1/\pi\Delta_i$ because then a significant effort of the other authority is needed to elicit positive effort – the best response will be zero on an interval starting at zero.

In the following we will consider the cases where both authorities are biased in the same direction, or biased in different directions.

Proposition 5 *As compared to the socially optimal equilibrium with unbiased authorities,*

1. *Under Joint Decisions (or strategic complements in general) both authorities' equilibrium efforts increase (decrease) if they are both consumer-friendly (industry-friendly), while the outcome is ambiguous if biases are opposite.*

2. *Under Independent Decisions (or strategic substitutes in general) at least one authority exerts more (less) effort if they are both consumer-friendly (industry-friendly), while the other authority's effort may increase or decrease. If biases are opposite then the industry-friendly authority works less while the other authority works more.*

Proof. Follow from upward- or downward shifts of increasing or decreasing best responses, respectively. These results hold more generally for strategic substitutes and complements.■

Because the equilibrium efforts with biases are always different from the social optimum, by implication none of the outcomes with biased authorities is socially optimal, not even in the case where authorities have opposite biases.[15] This implies that the bias of one authority *cannot* be fully compensated even by an optimal bias by the other authority. Nevertheless, it is a valid question to consider which bias a new authority should have depending the bias of the existing authority and their institutional relationship. As we will see, there is an interesting interaction between the decision process and biases that produces the answer – it will not simply be "correct a bias with the opposite bias".

More precisely, we find the following:

Proposition 6

1. *Under Joint Decisions (or strategic complements in general) the effects of a bias by one authority can be reduced by a bias in the* opposite *direction.*

2. *Under Independent Decisions (or strategic substitutes in general) the effects of a bias by one authority can be reduced by a bias in the* same *direction.*

3. *Still, even at the optimal bias of the new authority welfare will never achieve the first best.*

Proof. We start from the following fundamental observation (also see Figures 1 and 2, and also Figure 3 below): Since the best response R_2 of an unbiased authority 2, seen as a function of the other authority's effort e_1 on the vertical axis, stems from the first-order condition identical to the one of welfare maximization, the social planner's indifference curves cut this best response when they are horizontal. This fact is precisely the expression

[15] Dewatripont and Tirole (1999) consider "advocates" who specialize in defending one of the two sides representing an issue. Their approach is different because they assume that utility depends only on winning the case for one side, which is more likely to occur if only information in favour of one side is found. This muffles incentives of a single entity to investigate in both directions.

of the maximization of welfare given the other authority's effort, with welfare increasing at this intersection in the direction towards the welfare maximum. Furthermore, at these points any given biased best response R_1 of authority 1 (seen as a function of the authority 2's effort e_2 on the horizontal axis) is either increasing or decreasing, depending on whether efforts are strategic complements or substitutes. We are interested in the side of this intersection where R_1 enters the region of higher welfare: This is where the new intersection (and therefore Nash equilibrium) is to be constructed by giving the right bias to the new authority 2 and shifting its best response R_2.

Under strategic complements best responses are increasing, and if the bias is $\lambda_1 > 1$ (industry-friendly) then the best response R_1 of authority 1 is shifted downwards. Since at the intersection with authority 2's unbiased best response welfare then is increasing upwards and R_1 cuts the indifference curve from below, higher welfare can be achieved if R_2 is shifted rightward, or in other words, if authority 2 is given a consumer-friendly bias $\lambda_2 < 1$. An analogous argument shows that for $\lambda_1 < 1$ the optimal choice involves $\lambda_2 > 1$.

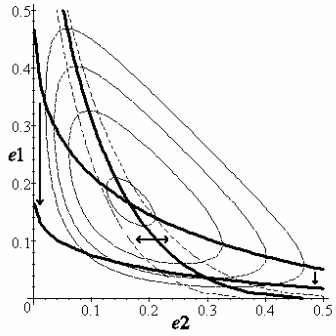
With strategic substitutes best responses are decreasing, there R_1 will cut the indifference curves from above. If it is shifted downwards ($\lambda_1 > 1$, industry-friendly) then welfare is higher to the left, and the optimal R_2 is shifted to the left. This implies that in this case authority 2 should also be industry-friendly, $\lambda_2 > 1$ The case $\lambda_1 < 1$ leads to the analogous result $\lambda_2 < 1$.

In all cases the highest indifference curve reached in this way is below the social optimum – because a biased R_1 does not pass through the social optimum in any circumstance.■

The reasoning behind these results is that biases change how much authorities spend on investigating, so they create over- or under-spending from a social point of view. These effects are non-trivial because one authority's bias not only determines how well this authority works, but through a strategic effect also distorts the incentives of the other authority. Anyhow, the more surprising result is clearly the second one: How can a bias be corrected by "worsening" the problem? Let us see what happens if in the case of $\lambda_1 > 1$ authority 2 would be endowed with a $\lambda_2 < 1$: As one can easily confirm in Figure 3, as a result authority 2 would heavily overspend to compensate, while authority 1 spends even less than before – total welfare decreases even further because of decreasing returns to scale. Giving the *same* bias to authority 2 raises authority 1's equilibrium effort e_1 and lowers e_2 , which raises welfare because authority 2 had been spending too much and authority 1 too little.

Figure 3

With Independent Decisions The Best Remedy To A Downward-Shifted R_1 Is A Left-Shift of R_2 (Same Bias), While A Right-Shift Lowers Welfare.



So far we have assumed that the introduction of a new authority leaves the bias of the existing authority unaffected. This may not always be the case, as the old authority may feel that since now it shares the responsibility of market outcomes with someone else it will be less bound by pressure from interest groups. We leave this issue for further exploration.

6. Lobbying and Capture

The decision to establish two regulatory authorities can be influenced by attempts of the government to make lobbying and capture more difficult. The argument could go like this: Not only would lobbyists have to knock on two doors to advance their interests, but also the authorities could control each other. Yet the validity of this argument must be seen in the light of the precise relationship between the two authorities. We have seen above that the welfare effects of biases can be unexpected depending on this relationship.

For concreteness, we assume that the lobby trying to influence the authorities is an industry pressure group. This is also the more relevant case because of the known large difficulty in organizing a joint effort by dispersed consumers. The questions to be answered are: Is lobbying more or less successful with two authorities? Will lobbyists concentrate on one authority, or lobby with both?

One caveat is in order: We consider lobbying as a long-term investment by lobbying groups which they make before actually knowing for which specific issues they will need the authority's bias. This seems a reasonable assumption given that regulators part from given standpoints (which they defend), and that argumentation and counter-argumentation take time. Furthermore, if lobbying only occurred in the event of a violation then it in itself would be proof of violation – which is logically inconsistent.

To answer the first of the above questions we will consider the extreme case where the lobby attempts to prevent the investigation altogether; then we consider how

would be spent on lobbying if the investigation cannot be stopped. Assume the industry pressure group must spend the amount of $m_i(\lambda_i)$ to move authority i's bias to $\lambda_i \geq \lambda_{i0}$, where m_i is increasing in λ_i, and λ_{i0} is the initial bias of the authority due to other reasons than lobbying. The aim of our analysis is to determine the minimum lobbying expenditure necessary to prevent the investigation.

If there is a single authority, it will not investigate if

$$\pi\left(\Delta_S - \lambda_i \Delta_\Pi\right) p_i'(0) \leq 1. \tag{14}$$

For this it is sufficient in any case to convince the authority that there is no case altogether, for which it is sufficient to set $\lambda_i = \Delta_S / \Delta_\Pi$, resulting in $\Delta_i = 0$. If $p_i'(0) < \infty$ then this outcome can be achieved more cheaply, with

$$\lambda_i \geq \frac{\Delta_S}{\Delta_\Pi} - \frac{1}{\pi \Delta_\Pi p_i'(0)}, \tag{15}$$

which happens when the authority acknowledges that there is a case, but believes it to be too insignificant to warrant investigation.

Whether this outcome can be achieved with two authorities depends on their decision process. This can be most easily seen in the cases of Joint or Independent Decisions. With Joint Decisions it is clearly enough that one authority does not investigate, that is,

$$\lambda_1 \geq \frac{\Delta_S}{\Delta_\Pi} - \frac{1}{\pi \Delta_\Pi p_1'(0)} \text{ or } \lambda_2 \geq \frac{\Delta_S}{\Delta_\Pi} - \frac{1}{\pi \Delta_\Pi p_2'(0)}. \tag{16}$$

Yet, with Independent Decisions the case will not be investigated if

$$\lambda_1 \geq \frac{\Delta_S}{\Delta_\Pi} - \frac{1}{\pi \Delta_\Pi p_1'(0)} \text{ and } \lambda_2 \geq \frac{\Delta_S}{\Delta_\Pi} - \frac{1}{\pi \Delta_\Pi p_2'(0)}, \tag{17}$$

Though at a rather basic level, these arguments show clearly that the structure of the decision process between authorities strongly influences the vulnerability to lobbying, and that the mere creation of a second authority as such does not solve the problem if not done carefully.

Let us now consider lobbying efforts in the more realistic case where authorities will investigate in any case. Here we are interested in knowing to which authority the lobby will turn its attention. The following results are less conclusive, and need more future investigation. The lobby would like to minimize the expected loss of a decision against them plus the amount spend on lobbying:

$$\min P(e_1, e_2) \pi \Delta_\Pi + m_1(\lambda_1) + m_2(\lambda_2). \tag{18}$$

We obtain the following result:

Proposition 7 *If the lobby cannot prevent the investigation it spends more on the authority that is more industry-friendly. Under Independent Decisions (strategic substitutes) it may*

also lobby more with the authority that has less strongly decreasing returns to scale in investigation.

Proof. The Lagrangian of the lobbyist's maximization problem is the following:

$$\begin{aligned} L(\lambda_1,\lambda_2,e_1,e_2;\mu_1,\mu_2) = P(e_1,e_2)\pi\Delta_\Pi + m_1(\lambda_1) + m_2(\lambda_2) \\ +\mu_1\left(\frac{1}{\partial P/\partial e_1} - \pi(\Delta_S - \lambda_1\Delta_\Pi)\right) \\ +\mu_2\left(\frac{1}{\partial P/\partial e_2} - \pi(\Delta_S - \lambda_2\Delta_\Pi)\right) \end{aligned} \tag{19}$$

The first line contains the objective function to be minimized. The second and third lines contain the first-order conditions prescribing positive efforts as an equilibrium of the authorities's game given the λ's; the multipliers μ_i are negative because the first-order conditions are binding upwards. The Kuhn-Tucker conditions describing the minimum are, for $i=1,2$

$$\frac{\partial L}{\partial \lambda_i} = m_i' + \mu_i\pi\Delta_\Pi \geq 0,\ \lambda_i \geq \lambda_{i0},\ \frac{\partial L}{\partial \lambda_i}(\lambda_i - \lambda_{i0}) = 0, \tag{20}$$

$$\frac{\partial L}{\partial e_i} = \frac{\partial P}{\partial e_i}\pi\Delta_\Pi - \mu_i\frac{\partial^2 P/\partial^2 e_i}{(\partial P/\partial e_i)^2} - \mu_j\frac{\partial^2 P/\partial e_i\partial e_j}{(\partial P/\partial e_j)^2} = 0, \tag{21}$$

$$\frac{\partial L}{\partial \mu_i} = \frac{1}{\partial P/\partial e_i} - \pi(\Delta_S - \lambda_i\Delta_\Pi) = 0. \tag{22}$$

Solving $\partial L/\partial e_i = 0$ and $\partial L/\partial e_j = 0$ for μ_i yields

$$\mu_i = \Delta_\Pi\pi\left(\frac{\partial P}{\partial e_i}\right)^2 \frac{\left(\partial^2 P/\partial^2 e_j\right)\left(\partial P/\partial e_i\right) - \left(\partial P/\partial e_j\right)\left(\partial^2 P/\partial e_i\partial e_j\right)}{\left(\partial^2 P/\partial^2 e_i\right)\left(\partial^2 P/\partial^2 e_j\right) - \left(\partial^2 P/\partial e_i\partial e_j\right)^2} < 0 \tag{23}$$

In this expression the denominator is positive because P is concave; the numerator will be strictly negative if $\partial^2 P/\partial e_i\partial e_j > 0$, and may approach zero only if $\partial^2 P/\partial e_i\partial e_j < 0$.

From (20), little money will be spend on authority i if and only if μ_i is small compared with $m_i'(\lambda_{i0})/\pi\Delta_\Pi$. Under strategic complements or substitutes this will occur if $\partial P/\partial e_i \approx 0$, or by (22) if the authority is strongly consumer-friendly. Under strategic substitutes there is the second possibility that the numerator of μ_i could be close to zero. This happens if the curvature of P with respect to e_j is low (because of the concavity of P then the curvature with respect to e_i must be sufficiently large), or that $\partial P/\partial e_j$ is rather flat. The effect of this is that an increase in λ_j induces a large reduction in e_j.■

This result implies that regardless of the relation between authorities lobbyists will not distribute their efforts evenly if they cannot prevent the investigation. In this case they will not bother to change the position of the "hostile" authority, but rather invest in the support from the "friendly" authority.

On the other hand, with strategic substitutes, or Independent Decisions, the money is spent on the authority whose effort is easier to influence. Mathematically this effect is also present with strategic complements, for example Joint Decisions, but is always overshadowed by the interaction between authorities. Since in the latter case a reduction in one authority's effort also reduces the other's, the lobby finds it optimal to exploit this effect on both authorities.

Finally, the above proposition is quite straightforwardly adaptable to consumer groups' lobbying. The objective function () is changed to

$$\max P\left(e_1, e_2\right)\pi\Delta_S - M_1\left(\lambda_1\right) - M_2\left(\lambda_2\right), \tag{24}$$

where M_i denotes the consumer group's spending and is decreasing in $\lambda_i \le \lambda_{i0}$. Again, spending will be low on authority i if μ_i is close to zero. Since conditions (21) essentially remain the same, and therefore the expressions for μ_i, consumer groups will lobby more heavily with the industry-friendly authority.

7. Extensions

In this section we present some extensions of the basic model. They present paths for further research, and do not yet lead to conclusive answers.

7.1. The Effects of Unclear Responsibilities

Here we will shortly discuss the issue of unclear responsibilities, modelled as a modification to authorities' utility functions. This is not the only way to capture this aspect, but the one most consistent with our approach.

The problem of unclear responsibilities is a complex one, which may stem from the political and juridical process that defines the field of action of each of the authorities. A simple way to capture this in our simple model is through the weight that a certain issue has in the authorities' objective functions

$$U = \sum\nolimits_k \left(w_k P\left(e_{ik}, e_{jk}\right)\pi\Delta_{ik} - e_{ik} \right). \tag{25}$$

Here we assume that the importance a case assumes is measured by the factor w_k, while expenditures on cases e_{ik} naturally count as equally for all cases. The first-order condition for an interior optimum is

$$w_k \pi\Delta_{ik} \frac{\partial}{\partial e_{ik}} P\left(e_{ik}, e_{jk}\right) = 1, \tag{26}$$

which makes clear that a case with less weight will receive less effort. More importantly, remembering from (6) the fact that no effort will be expended on a case if

$$w_{ik}\pi\Delta_{ik}\frac{\partial}{\partial e_{ik}}P(0,0)<1, \tag{27}$$

we see that if $\partial P/\partial e_i$ is finite at zero then for weights w_{ik} sufficiently small no effort at all will be dedicated to this case. This means that if none of the authorities has this type of case on their list of priorities then it may disappear in a "vacuum of responsibility".

These effects are reinforced if authorities are subject to a budget constraint (money or time) $\sum_k e_{ik} \le E_i$. If the Lagrange multiplier of this constraint is $\mu_i \ge 0$, then the first-order conditions for an interior solution become

$$\pi\Delta_{ik}\frac{\partial}{\partial e_{ik}}P\left(e_{ik},e_{jk}\right)=\frac{1}{w_{ik}}+\frac{\mu_i}{w_{ik}}. \tag{28}$$

The resulting effects are: All efforts are smaller if the constraint is binding. Furthermore, if the weight w_{ik} is small then it magnifies the effect of the budget constraint: *Ceteris paribus*, effort is more strongly reduced as compared to the situation without a budget constraint.[16]

Is there a way out? Returning for simplicity to the case without budget constraint in (27), we see that it is more likely that no effort is the outcome if the marginal increase in probability $\partial P(0,0)/\partial e_{ik}$ is small. If we compare the extreme cases of Joint or Independent Decisions, we see that

$$\begin{aligned} &\text{Joint Decisions:} && \tfrac{\partial}{\partial e_{ik}}P(0,0)=p_i'(0)\,p_j(0)\\ &\text{Independent Decisions:} && \tfrac{\partial}{\partial e_{ik}}P(0,0)=p_i'(0)\left(1-p_j(0)\right)\end{aligned} \tag{29}$$

Given that $p_j(0)$ is zero or close to zero, $\partial P(0,0)/\partial e_{ik}$ will always be larger with Independent Decisions than with Joint Decisions. The same will be true with any relation between the two authorities where a failure to investigate by one of them does not destroy the incentives for the other authority to do so.

Conclusion: The responsibilities of the authorities should be made clear, otherwise some events may not receive due attention. Furthermore, this situation is less likely to appear with Independent Decisions than with Joint Decisions.

7.2. Transversal and Sectoral Regulation

Assume now that authorities do make errors: They may receive a wrong signal that a violation has occurred. Given that a SRA has more knowledge about the sector is regulates than a transversal CA it seems reasonable to assume that the probability of making such an error will be higher for the CA. On the other hand, the SRA may be more subject to lobbying and therefore have a bias $\lambda_s \neq 1$. If the decision is to made which of the two authorities (alone) should look into a particular case then there will be a trade-off between the CA's higher probability of errors and the SRA's bias. We will consider how the bias of the CA should be set, and which authority should take on which cases.

---

[16] Actually, this argument applies to the increase in $\partial P/\partial e_i$, while the exact reduction in e_i still depends on the curvature of $\partial P/\partial e_i$ which has no direct economic meaning.

Assume that during a successfully concluded investigation both the SRA and the CA receive the correct signal that a violation has occurred as before with probability π, and the wrong signal that a violation has occurred while in fact none has occurred with probabilities $\pi_s < \pi_c$. The remaining event, a correct signal of no violation, is received with probabilities $1-\pi-\pi_s$ and $1-\pi-\pi_c$, respectively.

The effect on consumer surplus and on industry profits of an authority's error are $\delta_S, \delta_\Pi > 0$, but this time we assume $\delta_S - \delta_\Pi < 0$: the error lowers total welfare. Given its bias λ_s, the SRA maximizes

$$\begin{aligned} U_s &= p(e_s)\left(\pi\left(\Delta_S - \lambda_s\Delta_\Pi\right) + \pi_c\left(\delta_S - \lambda_s\delta_\Pi\right)\right) - e_s \\ &= p(e_s)\left(\pi\Delta + \pi_s\delta - (\lambda_s - 1)\left(\pi\Delta_\Pi + \pi_s\delta_\Pi\right)\right) - e_s. \end{aligned} \tag{30}$$

We assume that it raises welfare to investigate even the presence of errors, therefore $(\pi\Delta_\Pi + \pi_s\delta_\Pi) > 0$. Then social welfare is maximized for λ_1, while any $\lambda_1 > (<) 1$ yields under-/(over-)investment. The welfare resulting from the SRA's action is $W_s^* = p\left(e_s^*\right)(\pi\Delta + \pi_s\delta) - e_s^*$.

As concerns the CA, its optimal bias is trivially $\lambda_c = 1$: Given its higher probability of error, social welfare is maximized as a solution of the problem $W_c^* = \max_e p(e)(\pi\Delta + \pi_c\delta) - e$, which the CA replicates when it has precisely the same objective function. This result becomes clear when we take into account that the CA will act instead of the SRA, and not in conjunction with it – there is nothing to compensate given that the SRA will not be active.

If the SRA had no bias it would clearly be preferable that it conduct the investigation and not the CA, therefore this situation defines the welfare benchmark. With a biased SRA welfare under both authorities' investigations will be lower, but which of the two authorities will be active depends on whether $W_c^* > W_s^*$.

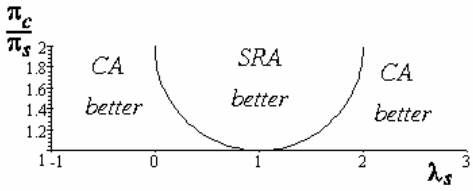
Consider the following example: Let $p(e) = \sqrt{e}$ for $e \le 1$. Then CA is better if

$$(\pi\Delta + \pi_c\delta)^2 \ge (\pi\Delta + \pi_s\delta)^2 - (\lambda_s - 1)^2\left(\pi\Delta_\Pi + \pi_s\delta_\Pi\right)^2. \tag{31}$$

The first observation is that the SRA is better if λ_s is close enough to 1, and that the SRA's biases in both directions are equally bad (This may be special to this example). Second, the left-hand side is decreasing in the CA's probability of error π_c, so π_c must be small enough. The trade-off becomes clear in the following Figure 4: The CA is better than an almost unbiased SRA only if π_c is close to π_s, otherwise the case is better investigated by the SRA. On the other hand, if the bias of the SRA is large enough then it may always be better to leave the case to the CA.

Figure 4

The trde-off between bias λ_s and the relative probability of error



$$\pi_C / \pi_S \left(\text{with } \pi\Delta=2,\ \pi_S \delta=-1, \left(\pi_{\Delta\Pi}+\pi_S \delta_\Pi\right)^2=1 \right).$$

8. Concluding Remarks

In the case of Portugal, it is seems that the law does not define very clearly the boundaries between the areas of responsibility of the Competition Authority (CA) and the Sectoral Regulatory Authorities (SRAs). In this work we have focused on how the relationship between the two authorities determines the incentives both authorities have to investigate cases.

Between the two alternatives of having the authorities decide jointly or independently, we found that the second hypothesis performs far better. First, with independent decisions the probability that cases are solved is highest, even though each authority may give less attention to the case than if it were alone. Second, Independent Decisions are less vulnerable to lobbying, while, third, it is also less likely that no authority feels responsible for a given case.

A further interesting result is that, if the bias of one authority is to be corrected by a bias of the other authority, the best choice is not necessarily an opposite bias. Indeed, with Independent Decisions the optimum is achieved at the same bias, avoiding excessive distortions in incentives.

References

Concorrência, Autoridade Da (2003), Estratégia de Desenvolvimento (2003 - 2005). http://www.autoridadedaconcorrencia.pt.

Dewatripont, Mathias and Tirole, Jean (1999), Advocates. *Journal of Political Economy*, 107(1):1-39.

Laffont, Jean-Jacques and Martimort, David (1999), Separation of Regulators Against Collusive Behavior. *RAND Journal of Economics*, 30(2):232-262.

Laffont, Jean-Jacques and Tirole, Jean (1993), A Theory of Incentives in Procurement and Regulation. MIT Press.

Olsen, Mancur (1965), The Logic of Collective Action: Public Goods and the Theory of Groups. Cambridge, Harvard University Press.

Stigler, George J. (1971), "The Theory of Economic Regulation". *Bell Journal of Economics and Management Science*, 2(1):3-21.

LOCATION AND LOCALIZATION OF PORTUGUESE MANUFACTURING INDUSTRIES

Paulo Guimarães
Universidade do Minho and CEMPRE

Octávio Figueiredo
Universidade do Porto and CEMPRE

8 Março 2004

Abstract

This paper offers an empirical investigation into the level of localization of Portuguese manufacturing industries. Using the conceptual framework of , we develop an alternative method to measure localization. We find that most Portuguese industries exhibit modest levels of localization and that the industries with the highest tendency to cluster are more traditional sectors. By regions, levels of localization are particularly important in the Oporto-Lisbon coastal corridor and around Serra da Estrela. We also find that accessibility to the Oporto-Lisbon coastal corridor (and transportation costs) is the variable with the highest potential to reshape the Portuguese industrial landscape.

JEL classification: R12; R32

1. Introduction

In the last two decades, the recognition of the importance of agglomeration economies as a source of increasing returns for individual firms has led researchers in international trade, growth and industrial organization to join urban and regional economists in investigating why and to what extent they explain observed spatial concentration of economic activity. There is a wealth of examples from many countries showing that many individual industries tend to be localized (i.e. to concentrate over and above overall economic activity). There are the classic examples of Alfred Marshall for the clustering of cutlery industry in Sheffield and the jewellery industry in Birmingham as well as the more contemporaneous and prominent examples of the high-tech industry in Santa Clara County, California (the Silicon Valley), Boston's Route 128 in Massachusetts, and Austin, Texas. But are these examples the rule or the exception? And how general and how strong is the tendency of industry to agglomerate?

The debate reignited by the advent of the new economic geography, with its emphasis on the importance of external economies, has again brought these questions to the forefront of many scientists' research agendas. But clear answers to these questions have been marred by the lack of an adequate approach to the measurement of the localization level of an industry. More recently, Ellison & Glaeser (1997) tackled this problem. Based on a Random Utility (Profit) Maximization model of industrial location, they proposed an index that avoids several of the pitfalls of past measures. In a short period of time, their work spawned a significant number of studies and rapidly emerged as the standard approach for measuring the degree of localization of industry. Nevertheless, we contend that the index proposed in Ellison & Glaeser (1997) (henceforth the EG index) has some serious drawbacks. Thus, we elaborate on the work of these two authors and propose an alternative index which is more faithful to the theoretical construct that was exposed in their original work [Ellison & Glaeser (1997)].

In this paper we abstract from the dynamic aspects that lead to changes in the landscape of economic activity overtime. Instead, we try to characterize the level of localization at a given point in time. Using the most recently available data we provide a thorough characterization of the level of localization of the Portuguese industry and query into the mechanisms that underlie the location decisions of firms.

The rest of the paper is comprised of 4 sections. The following section reviews the measures of spatial concentration and attendant problems. In section 3 we explore the methodological questions and develop the alternative method for measuring localization. Section 4 explains the data and presents our main results on localization of Portuguese manufacturing industries. Finally, our last section (section 5) offers some concluding comments and points to policy implications.

2. The Measurement of Spatial Concentration

Past economists had no shortage of tools for measuring geographical concentration. Most prominent are Hoover's location quotient and the Gini coefficient as applied by Krugman (1991). These measures quantify the discrepancy between the distribution of regional employment in a particular industry against the distribution of overall manufacturing employment. But are these measures able to capture the concept of agglomeration? A first obvious problem is that they are sensitive to the levels of

concentration within the industry. Take as an example two industries which have identical measures for the Gini index. The first industry is composed of many independent firms, all equally sized and located in a single region, while the second industry is composed of just one firm operating a large establishment. The first case agrees more with the notion of agglomeration, which may explain the clustering of all firms in that industry. But for the second industry, it is obvious that the concept of agglomeration does not apply. In the first case, external economies associated with firms' clustering are most likely the source of geographic concentration, while in the second case the spatial concentration is entirely explained by industrial concentration, and then by returns to scale.

Another obvious problem is that these measures do not account for the inherent randomness of the underlying location decisions. Firms may exhibit some level of spatial concentration by chance. This idea can be easily explained by resorting to the well-known balls and urns example often used in statistics. If one has, say, 10 urns (regions) and 10 balls (firms) and drops the balls at random into these urns then, even though all urns are equally probable, it is very unlikely that we will observe exactly one ball in each urn. Some clustering will necessarily occur and that is perfectly compatible with the idea that the balls were thrown at random (the firms' decisions were random). The above indexes are not able to control for this type of clustering.

From the above discussion it should be obvious that these indices do not adequately measure an industry's degree of agglomeration. The recent index proposed in Ellison & Glaeser (1997) to measure localization overcomes the limitations that were just discussed. Like the Gini coefficient it attempts to measure the tendency of one industry to agglomerate in relation to the general tendency of all industries to agglomerate. But, unlike its predecessors, it accounts for the inherent discreteness (lumpiness) that will be observed if location decisions are driven by chance alone and it expurgates the effect of industrial concentration. Thus, it allows for a standardized measure of the industry agglomerative forces which can be readily used for temporal or inter-sectorial comparisons. Most notably, the EG index is rooted in the location choice model of Carlton (193), which in turn is based on McFadden's Random Utility Maximization framework (henceforth RUM framework) and has been the workhorse for the empirical literature on industrial location [e.g. Bartik (1985), Luger & Shetty (1985), Hansen (1992), Schmenner, Huber & Cook (1987), Coughlin, Terza & Arromdee (1991), Woodward (1992), Friedman, Gerlowski & Silberman (1992), Head, Ries & Swenson (1995), Guimarães, Figueiredo & Woodward (2000), and Figueiredo & Woodward (2002)]. Other authors have proposed indices which are very closely related to the EG index. Based on a different set of theoretical arguments, Maurel & Sedillot (1999) constructed an index which is similar to the EG index. By comparing the two formulas, they show that the difference between the indices has an expected value of zero. Also noteworthy is the work of Devereux, Griffith & Simpason (2003). They showed that the index of EG can be conveniently approximated by the difference between an index that measures geographic concentration and another that measures industrial concentration.

More recently, Duranton & Overman (2002) have proposed a different approach to the measurement of spatial concentration. Their approach draws directly from methods well-known to spatial statisticians to measure concentration of spatial phenomena. They treat space as continuous and compute their measurements based on the cartesian distances between each pair of plants. Treating space as continuous has an inherent appeal but their

approach lacks a theoretical underpinning. Moreover, it is an essentially descriptive procedure that requires precise information on the exact location of each business unit.

The new wave of literature initiated with Ellison & Glaeser (1997) has already generated a substantial amount of applied work. Besides the U.S. [Ellison & Glaeser (1997), Dumais, Ellison & Glaeser (2002) and Holmes & Stevens (2002)], studies characterizing the localization levels of industries have been produced for France [Maurel & sedillot (1999) and Houdebine (1999)], Belgium [Bertinelli & Decrop (2002)], UK. [Devereux et al. (2003)] and Spain [Callejón 1997)]. Common to all studies is the finding that the majority of industries are localized.

In the following, we offer an approach to the measurement of industrial localization that is grounded more solidly on the RUM framework of industrial location, yet it borrows from the conceptual approach of Ellison & Glaeser (1997). As will become obvious in the next section, the link between the RUM location literature and the EG index is feeble. We show how the two can be better integrated. Also, contrary to the general trend in the literature, we argue that employment figures are a confounding factor in the measurement of localization [as proposed by Ellison & Glaeser (1997)] and advocate the use of counts of plants instead.

3. Methodologic Issues

Industrial location models based on the RUM framework provide an explanation for the agglomeration of industry - idiosyncratic factors aside, firms choose locations that yield the highest profits. In practice, most empirical models of industrial location have been applied to new firm births and accounted for the importance of localization economies in an indirect way - by introducing variables that measure the presence of that same industry. If we abstract from the dynamic questions we can use the theoretical framework of the RUM to justify the agglomeration of an industry. This was the approach of Ellison & Glaeser (1997). Yet, as argued here, the integration between the RUM and the derivation of the EG index can be strengthened and thus it should provide a better way to measure localization of an industry.

3.1. The EG Index

To motivate our approach, we take a more in-depth look at the derivation of the EG index. Let us assume at the outset that the economy is divided into J geographical units (regions). Also, we take as our reference a given industry which has exactly n_j plants located in each region j. Thus, $n = \sum_{j=1}^{J} n_j$ represents the total number of existing plants in our reference industry. Next, we briefly sketch how the EG index is obtained taking as a reference their model of natural advantages. If firm i chooses to locate in region j then its profits will consist of

$$\ln \pi_{ij} = \ln \overline{\pi}_j + \varepsilon_{ij} \tag{1}$$

where $\overline{\pi}_j$ is a non-negative random variable reflecting the profitability of locating in area j for a typical firm in the industry. In this formulation of the model, Nature introduces the randomness in $\overline{\pi}_j$ by selecting for each region the characteristics that make it unique (their

natural advantages). ε_{ij} is a random disturbance. If we assume that ε_{ij} is an identically and independently distributed random term with an Extreme Value Type I distribution[1] then, conditional on a realization of $\bar{\pi}_j$, we can apply result to obtain,

$$p_{j/\pi} = \frac{\exp(\ln \bar{\pi}_j)}{\sum_{j=1}^{J} \exp(\ln \bar{\pi}_j)} = \frac{\bar{\pi}_j}{\sum_{j=1}^{J} \bar{\pi}_j} \ , \tag{2}$$

which denotes the probability of a firm locating in region j. Thus, p_j is obtained from the Random (Profit) Utility Maximization framework of Carlton (1983) which, as mentioned earlier, gives support to most recent studies of industrial location. To derive their index, Ellison & Glaeser (1997) introduced two parametric restrictions regarding the expected value and variance of p_j. Thus, they assume that the distribution of $\bar{\pi}_j$ is such that:

$$E(p_j) = x_j \ , \tag{3}$$

and that,

$$V(p_j) = \gamma x_j (1 - x_j) \ , \tag{4}$$

where x_j may be thought of as the probability of a firm locating in region j in the absence of any region specific advantages for that industry. Thus, the larger the discrepancy between x_j and p_j, the larger the influence that these region specific effects (say, natural advantages) play in the location decisions of firms in that industry. That difference is captured by the parameter γ (which we will refer to as the EG parameter) which belongs to the unit interval. It is easy to see that if $\gamma = 0$ then the industry will tend to replicate the pattern observed for the x_j (what Ellison and Glaeser call the dartboard model) and we can conclude that there is no spatial concentration in excess of what we would expect to occur. If, however, $\gamma > 0$, then the actual location probabilities of the industry will differ from x_j and in the limit when $\gamma = 1$, each p_j has the largest variance and becomes a Bernoulli random variable. Thus, in the limit, all the investments for that industry would be located in a single region.

Ellison & Glaeser (1997) also show that the γ parameter may be derived from an alternative model that emphasizes industrial spillovers as the force leading to excessive concentration. Whichever theoretical motivation one uses is irrelevant because they are observationally equivalent and lead to the same functional form for the index, the practical implication being that we can not readily distinguish the two sources of geographic concentration (natural advantages and industrial spillovers).

To estimate γ for a particular industry they let x_j denote area j's share of total manufacturing employment. Here, the idea is that the model should on average reproduce the overall distribution of manufacturing activity. In a next step they considered the following "raw concentration index" of employment:

---

[1] In the past this distribution has been referred to by other names such as Weibull, Gumbel and double-exponential [Louviere, Hensher & Swait (2000)].

$$G_E = \sum_{j=1}^{J} (s_j - x_j)^2 \tag{5}$$

where, s_j denotes area j's share of employment in that industry and the $x_j s$ are as described above. Now, taking the expected value of G_E they obtain a function of γ and the authors use that relation to propose an estimator for γ. Their proposed estimator for γ (the EG index) is then

$$\hat{\gamma}_{EG} = \frac{G_E - \left(1 - \sum_{j=1}^{J} x_j^2\right) H_E}{\left(1 - \sum_{j=1}^{J} x_j^2\right)\left(1 - H_E\right)}\ , \tag{6}$$

where H_E is the employment Herfindhal index for the industry and the expected value of G_E is replaced by its actual value. Note that the computation of the Ellison and Glaeser measure of concentration only requires employment information. In that sense this measure is remarkable because it provides a framework to extract information about the spatial concentration of industry based exclusively on regional employment information.

3.2. A EG Index Based on Counts of Plants

One could as well derive an alternative estimator based on counts of plants. To see this, define

$$G_F = \sum_{j=1}^{J} \left(\frac{n_j}{n} - x_j \right)^2 \tag{7}$$

and proceeding in a fashion similar to EG (see Appendix A) we derive the following alternative estimator for γ :

$$\hat{\gamma}_A = \frac{nG_F - \left(1 - \sum_{j=1}^{J} x_j^2\right)}{(n-1)\left(1 - \sum_{j=1}^{J} x_j^2\right)} \tag{8}$$

The above expression is very similar to that of the Ellison-Glaeser index. It replaces the Herfindhal index by $1/n$ and the "raw concentration index" is expressed in terms of number of plants instead of employment. Like the estimator proposed by Ellison and Glaeser this estimator for γ is also, by construction, unbiased. Most notably it has a much smaller variance. To see this note that:

$$\frac{V\left(\hat{\gamma}_{EG}\right)}{V\left(\hat{\gamma}_A\right)} = \left(\frac{n-1}{n(1-H_E)} \right)^2 \frac{V\left(G_E\right)}{V\left(G_F\right)}. \tag{9}$$

An heuristic argument suffices to justify the better efficiency of this estimator. If all plants had the same dimension, the indexes would be identical (H_E would be $1/n$). As the Herfindhal index increases, the first term of the product in the RHS of (9) increases. One would also expect the second term (the ratio of the variances) to be larger with increases in the Herfindhal index. Thus, we argue that a more precise estimate for γ is obtained if we ignore the confounding influence of plant size (employment) and work directly with counts of plants. From another perspective, Holmes & Stevens (2002) provide additional evidence against the use of an index based on employment plant size. These authors found evidence that plants located in areas where an industry concentrates (as measured by the EG index) are larger, on average, than plants in the same industry outside the same area, thus suggesting that the EG index will tend to overstate the extent of localization economies.

A clear disadvantage of the EG index is that it does not provide an indication of statistical significance. In Appendix B we show how one can construct and implement a test for the null hypothesis that $\gamma = 0$ for the $\hat{\gamma}_A$ statistic.

3.3. An Alternative Method for Measuring Localization

An implicit assumption in the work of Ellison-Glaeser is that in the absence of natural advantages (or spillover effects) all individual industries would be faced with the same location probabilities, $p_j\left(=x_j\right)$. If these $p_j s$ are obtained from the RUM framework, as is claimed, then this amounts to the underlying assumption that all industries would have identical profit functions. But, what drives the location of a chemical plant may be very different from what drives the location of a textile or food processing plant. In other words, we claim that if natural advantages (or spillovers) were inexistent then one would still expect to find different patterns of location across industries, simply because industries value regional characteristics differently. For example, wages may be an important component of the profit function for the apparel industry but may not be a determinant factor in the locational decision of an automotive plant. To incorporate this dimension into the framework laid out by Ellison and Glaeser, we take a different route – we explicitly model the location decision process of firms and measure concentration in excess of that which would result if all industries were influenced by the same set of (observed) locational factors. That is, instead of approximating the "attractiveness" of a region by its share of manufacturing employment[2], we let each industry have a different valuation for the "attractiveness" of a region based on the particular combination of factors that are relevant for that industry.

Hence, we admit that the profit function faced by firm i in our reference industry, if it decides to locate in region j, may be written as,

$$\log \pi_{ij} = \boldsymbol{\theta}' \mathbf{y}_j + \eta_j + \varepsilon_{ij} \,, \tag{10}$$

[2] At this point it should be noted that Ellison and Glaeser report the use of other alternatives to manufacturing employment such as the area and the population.

where, the $\mathbf{y}_j$ are those regional characteristics that affect the location decisions of firms in all industries (e.g. wages, land costs, market accessibility and transportation costs), $\boldsymbol{\theta}$ is a vector of parameters, and η_j is a (regional) random effect that picks the unobservable locational advantages of that region for a particular industry. The other random term, ε_{ij}, is as defined earlier. Now, conditional on the $\eta_j s$ and again drawing on result we can write,

$$p_{j/\boldsymbol{\eta}} = \frac{\exp(\boldsymbol{\theta}'\mathbf{y}_j + \eta_j)}{\sum_{j=1}^{J}\exp(\boldsymbol{\theta}'\mathbf{y}_j + \eta_j)} = \frac{\exp(\eta_j)\lambda_j}{\sum_{j=1}^{J}\exp(\eta_j)\lambda_j}. \tag{11}$$

The likelihood function (conditional on the $\eta_j s$) implied by the above expression is that of a conditional logit model:

$$L(n_1, n_2, ..., n_J / n, \boldsymbol{\eta}) = \prod_{j=1}^{J} p_{j/\boldsymbol{\eta}}^{n_j}. \tag{12}$$

which in turn is the kernel of a multinomial distribution with parameters $(p_{1/\boldsymbol{\eta}}, p_{2/\boldsymbol{\eta}}, ..., p_{J/\boldsymbol{\eta}}, n)$,

$$L(n_1, n_2, ..., n_J / n, \boldsymbol{\eta}) \propto n! \prod_{j=1}^{J} \frac{p_{j/\boldsymbol{\eta}}^{n_j}}{n_j!}. \tag{13}$$

Now, if we assume that the $\exp(\eta_j)s$ are iid. gamma distributed with parameters $(\delta^{-1}, \delta^{-1})$ then $\exp(\eta_j)\lambda_j$ follows a gamma distribution with parameters $(\delta^{-1}\lambda_j, \delta^{-1})$. It follows from that the $(p_1, p_2, ..., p_J)$ are Dirichlet distributed with parameters $(\delta^{-1}\lambda_1, \delta^{-1}\lambda_2, ..., \delta^{-1}\lambda_J)$ .Therefore the unconditional likelihood function may be written as,

$$L(n_1, n_2, ..., n_J / n) = n! \int \prod_{j=1}^{J} \frac{p_j^{n_j}}{n_j!} g(p_1, p_2, ..., p_J) dp_1 dp_2, ..., dp_J \ . \tag{14}$$

The above integral has a closed form, whose solution is known as the Dirichlet-Multinomial distribution [Mosimann1962]:

$$L(n_1, n_2, ..., n_J / n) = \frac{n!\Gamma(\delta^{-1}\lambda_\bullet)}{\Gamma(\delta^{-1}\lambda_\bullet + n)} \prod_{j=1}^{J} \frac{\Gamma(\delta^{-1}\lambda_j + n_j)}{\Gamma(\delta^{-1}\lambda_j) n_j!} \tag{15}$$

where $\lambda_\bullet = \sum_{j=1}^{J}\lambda_j$. The resulting likelihood function offers no particular challenge and can be easily implemented. But the interesting feature of this approach is that now,

$$E(p_j) = \frac{\lambda_j}{\lambda_\bullet} \tag{16}$$

and

$$V(p_j) = \frac{\lambda_j}{\lambda_\bullet}\cdot\left(1-\frac{\lambda_j}{\lambda_\bullet}\right)\cdot\frac{1}{\delta^{-1}\lambda_\bullet+1} \tag{17}$$

and by analogy with (4) and the approach of Ellison and Glaeser we can interpret

$$\tilde{\gamma} = \frac{1}{\delta^{-1}\lambda_\bullet+1} = \frac{\delta}{(\lambda_\bullet+\delta)} \tag{18}$$

as an index of excessive spatial concentration for that industry, that is, an alternative estimator for the EG parameter. As δ (the variance of the region specific random error) increases, so does $\tilde{\gamma}$ and in the limit when δ tends to infinity, $\tilde{\gamma}$ will tend to 1. On the other hand, $\tilde{\gamma}$ will approach zero as δ tends to zero.[3] Because in this latter situation, the Dirichlet-Multinomial distribution collapses to a standard multinomial distribution we can use a likelihood ratio test to test the hypothesis that the industry is more concentrated than what we would expect ($\delta = 0$).[4] To implement this model, we wrote the likelihood function in Stata (version 7) using that package's standard numerical maximization routine (a modified Newthon-Raphson algorithm). To obtain starting values, we first estimated a Poisson regression- which in this context produces the same estimates for the variable coefficients as the conditional logit model [Guimarães, Figueiredo & Woodward (2003)]. Convergence was fast with a very small number of iterations (less than 10 for most cases).[5]

4. Localization of Portuguese Manufacturing Industries

4.1. Data and Variables

Our main source of data was the Quadros do Pessoal database for 1999, the most recent available year. The *Quadros do Pessoal* are a yearly survey collected by the Ministry of Employment for all the existing companies operating in Portugal (except family businesses without wage earning employees) and cover 45,350 plants for the year of 1999.[6] Using this source, we tallied the number of plants as well as employment for each *concelho*

---

3 Unlike the EG index which often produces negative estimates, our estimator will always generate estimates that belong to the unit interval.

4 Because we are testing a value which is in the boundary of the set of admissible values for δ, we follow the suggestion in Cameron & Trivedi (1988) and adjust the level of significance of the chi-square statistic accordingly. Also, we should note that to apply the likelihood ratio test, we need to rescale the likelihood function of the Conditional Logit model as in (13).

5 There were however, a few cases (see below) when the model was unable to converge. We took it as evidence that the data were not overdispersed enough, although this procedure may be questionable. For these industries, we assumed that $\tilde{\gamma}$ was zero.

6 For a thorough description of this database see, for example, Mata, Portugal & Guimarães (1995). Unless otherwise noted the "Quadros do Pessoal" was the source for all the information used in this paper.

in continental Portugal.[7] We rely on the 3-digit (103 industries) classification of the Portuguese Standard Industrial Classification system (CAE).[8] Using the 275 Portuguese *concelhos* as the spatial choice set, we estimated a location regression for each industry (the Dirichlet-Multinomial model), as well as the corresponding measure of excessive concentration (localization) given by (18).

The choice of regressors for our location model was dictated by location theory. Location theory distinguishes three different sets of factors driving the firm's location decision problem: external economies, costs of production factors, and accessibility (transportation costs) to input and final demand markets. External economies can arise from two different sources. Localization economies are those external economies that result from the spatial concentration of firms of a particular industry in a given region and that are internalized by firms of that particular industry. This is what we are indirectly measuring through $\tilde{\gamma}$ (along with natural advantages of the regions). The other externality, urbanization economies, accrues from the clustering of general economic activity in a given area and benefits all plants locating in that particular area. Urbanization economies are proxied in our model by the "concelho" density of service and manufacturing establishments per square kilometre in 1999.

To control for the impact of factor prices, we obtained information on the cost of labour and land. Labour costs are measured by an index of the *concelho*'s average manufacturing base wage rate in 1999.[9] Since industrial and residential users compete for land, one may argue that when modelling with small areas and controlling for urbanization, as in our case, land costs can be proxied by population density. Consequently, following the suggestion of Bartik (1985), we use population density to approximate land costs.[10] We did not consider the cost of capital because it is practically invariant across alternatives. Interest rates do not differ regionally, and despite some minor differences in municipal taxes, the overall tax burden on manufacturing activity comes mostly from taxes set at the national level.

To account for market accessibility at a given location (and transportation costs) we enter two variables in the model. The drive time distance from each *concelho* to the Oporto-Lisbon corridor (the more urbanized coastal side of the country) measures large-scale accessibility, i.e. access to the largest markets. Small-scale accessibility, i.e., access to

---

7 The *concelho* is an administrative region in Portugal. In recent years some new *concelhos* have been created by the incorporation of parts of existing *concelhos*. To maintain data compatibility, we use the spatial breakdown of 275 *concelhos* that was still valid in 1997. These have an average area of 322.5 Km^2.

8 Revision 2 of the CAE.

9 Because we are not using real wages, a higher average manufacturing base wage rate in a given *concelho* can also indicate the presence of a highly-skilled workforce. If investors are willing to pay higher wages for more qualified workers, the coefficient of this variable is expected to be positive.

10 We used population for the year of 1996, taken from the National Institute of Statistics (INE).

regional markets, is proxied by the distance in time by road from each *concelho* to the administrative centre (the capital) of the related *distrito*.[11]

4.2. Results

We computed the localization index $\tilde{\gamma}$ for each of the 3-digit SIC industries at the *concelho* level.[12] As indicated before, for a small number of industries (8) the model did not converge and we assumed that $\tilde{\gamma}$ was zero in these cases. Additionally, for 17 industries, the $\tilde{\gamma}$ index was not statistically different from zero at the 95 per cent level of significance. For the remaining 75 industries (75 per cent of the 100 industries analyzed) we find evidence of excess of concentration ($\tilde{\gamma} > 0$). Therefore, a high percentage of Portuguese manufacturing industries appear to be localized, a result that corroborates similar analyses based on the EG index but performed for others countries.[13]

As previously observed for other countries as well, the degree of localization varies significantly across Portuguese industries. In Figure 1 we show a histogram of $\tilde{\gamma}$ at the *concelho* level for the 100 3-digit SIC industries. As can be seen, the localization index displays a very skewed distribution, the majority of industries showing slight levels of localization. Indeed, of the 75 industries whose $\tilde{\gamma}$ is significantly above 0 we find that 64 per cent display a degree of localization below the mean value of 0.013.

In Table 1 we list the 27 3-digit SIC industries that have a $\tilde{\gamma}$ index above the industry average. Among them we find a large number of traditional industries for which localization is determined by the historical specialization of particular regions. This is the case of the motorcycles and bicycles industry (SIC 354), the tannery industry (SIC 191), the jewellery industry (SIC 362), the textile industries (SICs 172, 173, 171, 176 and 177), the footwear industry (SIC 193), the ceramic industries (SICs 264 and 262), and the cork industry (SIC 205). Again, this pattern coincides with evidence for other countries that suggests that typically traditional industries are highly localized.[14]

Table 1 also shows that several more technologically advanced industries (such as fabrication of material for radio and television apparatus, the automobile industry, several industries that produce machinery and equipments, some primary metal industries, and the pharmaceutical industry), exhibit higher than average levels of localization.

---

11 The *distrito* is a higher administrative region level composed of several adjacent *concelhos*. Continental Portugal is divided in 18 *distritos*. The time distance variables report to the year of 1996. They were constructed using an algorithm that selected the shortest time route between locations, using as parameters the average travelling speed for the particular type of road as well as a road network compiled from road maps (ACP 1998/9; Michelin 1999) and detailed information from the Portuguese Road Institute (Instituto Português de Estradas). We thank Adelheid Holl for making this unpublished data available for the present study.

12 Our dataset contains information for 100 3-digit SIC industries. For SICs 231, 233, and 300, the *Quadros do Pessoal* dataset did not report any plant in 1999.

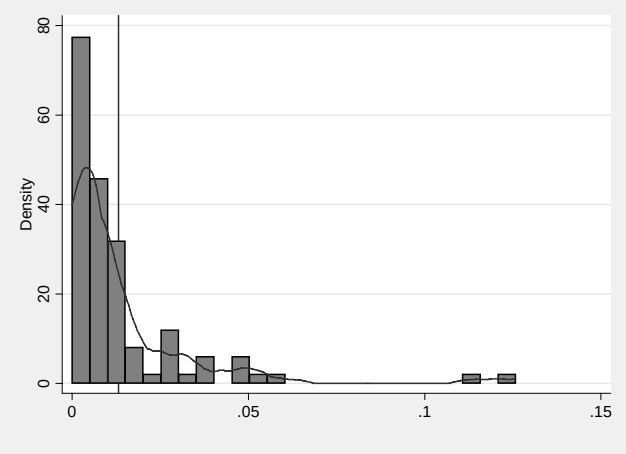
13 Ellison & Glaeser (1997) found that 446 out of 459 4-digit SIC industries in the United States were localized. Maurel & Sedillot (1999) found that 77% of the 273 4-digit French industries display excess of concentration. Similar results were found for the UK by Devereux et al. (2003).

14 See Ellison & Glaeser (1997), Table 4, Maurel & Sedillot (1999), Tables 1 and 2, Devereux et al. (2003), Tables 4 and 6, and Krugman (1991), Appendix D.

As could be expected, shipbuilding and several industries that process sea products are also among the most localized industries. Access to a natural resource, the sea, is crucial to explain firms' location decisions in this latter type of industries.

As these three different groups of localized industries suggest, high levels of localization can correspond to different location strategies. Past static externalities dominate in the case of the industrial clustering of firms in the traditional sectors. Current dynamic knowledge spillovers can explain the location strategy of the more technologically advanced industries. Natural advantages of the regions prevail in industries such as shipbuilding.

Figure 1
Histogram $\tilde{\gamma}$ at the *concelho* level



[insert Tables 1 and 2 about here]

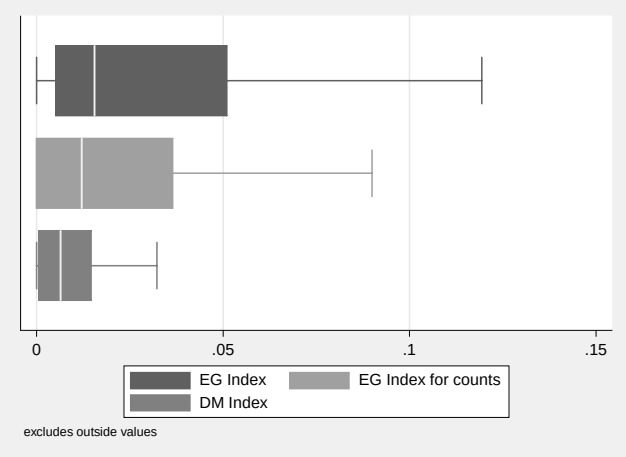
Table 2 displays the group of 25 3-digit SIC sectors that are non localized industries according to our results ($\tilde{\gamma} = 0$). As can be seen, the majority of these industries are capital-intensive industries for which returns to scale are a crucial factor. For this last group it is important to distinguish our measure of localization from a simple measure of geographic concentration. While some of these sectors (such as tobacco, petroleum refining or aircraft and space vehicles fabrication) can be highly concentrated in space, this concentration is almost entirely explained by industrial concentration, and thus by returns to scale rather than natural advantages or external economies associated with firms' clustering.

4.3. Comparison with Results Based on the EG Index

We now compare our estimates of localization with those provided by the EG index ($\hat{\gamma}_{EG}$) and the alternative EG index based on counts of plants ($\hat{\gamma}_{A}$). If we look at the extent of localization across 3-digits sectors, we find very similar results for the three measures. 68, 75, and 80 per cent of the industries exhibit excess of concentration

according to the modified EG index, our index, and the EG index, respectively. In Figure 2 we display the box-whisker plots computed across the 100 3-digit SIC industries for the three measures. To increase readability the graph omits a few extreme (high) values for each one of the distributions. As can be seen, all distributions show the same pattern of skewness with increasing interquartile ranges. Nevertheless, as we anticipated, our proposed measure of localization (labelled as DM index in the figure) exhibits much smaller values when compared with the EG index ($\hat{\gamma}_{EG}$) and the modified EG index ($\hat{\gamma}_{A}$). We take these results as confirmatory evidence that the original EG index tends to overstate the degree of localization of industry.

Figure 2
Box-Wisher plots for the three localization indexes



The three measures also encompass substantial differences when we look at the hierarchy of individual industries. The Spearman rank correlation coefficients between $\tilde{\gamma}$ and $\hat{\gamma}_{EG}$ (0.35) and between $\tilde{\gamma}$ and $\hat{\gamma}_{A}$ (0.49) are small, in spite of the fact that we reject independence between any of three statistics. Furthermore, as a quick inspection of Table 1 will reveal, among the top 27 most localized industries according to $\tilde{\gamma}$ we only find 14 (18) of the industries for a similar ranking based on $\hat{\gamma}_{EG}$ ($\hat{\gamma}_{A}$).

4.4. Extensions

To compute our localization index we had to explicitly model the location decision process of firms. This, in turn, allows for a much richer analysis than that obtained using the methodology of Ellison and Glaeser. In this section, we extend our research into two distinct directions. First, we measure the overall levels of localization by regions and try to identify which regions are associated with localized industries. Next, we set out to gain some understanding into the working of the mechanisms that influence the location decisions of firms.

4.4.1. Regional Analysis

The standardized residuals from the Dirichlet-Multinomial regressions (Pearson residuals) provide useful information because they may be regarded as an indirect estimate of the level of localization economies (and natural advantages) for each region and for each specific industry. The analysis of these residuals allows us to link industries with regions. This enables us to find the regions that are associated with the more localized industries. We can go one step further and construct an index of overall localization for each region by aggregating the residuals for all sectors.

Regions Associated with the Most Localized Industries.In this subsection we use the standardized residuals from the Dirichlet-Multinomial regressions to map out those regions that are associated with the 27 most localized industries shown on Table 1. To do this, we identify, for each of these industries, all residuals that exceed 1.96.[15] A summary of the information is presented in Table 3. For the group of most localized traditional industries, we found that using the residuals to analyze excess of concentration resulting from localization economies makes sense. Indeed, these residuals show, for each of the localized traditional industries, the *concelhos* that one would expect to find based on knowledge about the historical patterns of specialization of Portuguese regions. As expected, the highest residuals for the motorcycles and bicycles industry (SIC 354) are found for the Águeda and Anadia *concelhos*; the tannery industry (SIC 191) is localized at Alcanena and the largest spatial concentration of firms in the jewelry industry (SIC 362) is found for Gondomar; the textile industry (SICs 172, 173, 171, 176 and 177) is localized, as expected, in the *concelhos* of the Ave Valley (Guimarães, Santo Tirso, and VN. Famalicão), in the region around Serra da Estrela (Covilhã, Gouveia, Seia, Guarda, and Castelo Branco), and in the Barcelos *concelho*; the highest residuals for the footwear industry (SIC 193) are found for Oliveira de Azeméis, Felgueiras, Santa Maria da Feira and Guimarães, well known regions of clustering of firms for the Portuguese footwear industry; this is also the case of the *concelhos* of Barcelos and Alcobaça, for the ceramic industries (SIC262), and Santa Maria da Feira and Montijo, for the cork industry (SIC 205).

[insert Table 3 about here]

As indicated before, some of the more technologically advanced industries also present levels of localization above the mean. Table 3 indicates for these industries the associated *concelhos*. We found significant clusters of firms in the automobile industries (SICs 341 and 343) in three different regions: in the north side of the country, around the cities of Porto and Aveiro (VN. Gaia, Maia, Oliveira de Azeméis, Ovar, and Águeda); in the central part of the country, around the cities of Viseu and Coimbra (Mangualde and Cantanhede); and, more to the south, around Lisbon (Rio Maior, Abrantes, Palmela, and Loures). The pharmaceutical industry (SIC 244) follows a similar pattern with "excessive concentration" of firms around Porto (Penafiel and VN. Famalicão); around Coimbra and Viseu (Penela, Coimbra, Condeixa-a-Nova, Oliveira de Frades, Tondela, and Mortágua); and around the city of Lisbon (Sintra). The largest significant clusters for SIC 323 (Fabrication of material for radio and television apparatus) are found for Mafra and

[15] The justification for this approach lies in the fact that the distribution of Pearson residuals for a Poisson distribution with moderately sized counts is known to approximate the standardized normal distribution.

Proença-a-Nova. For industries that produce machinery and equipments (SICs 291 and 295), excess of concentration resulting from localization economies is particularly important in Marinha Grande, Braga, Leiria, Oliveira de Azeméis, and Seixal.

Table 3 also makes patent the importance of natural advantages for the shipbuilding industry (SIC 351), as well as for several industries that process sea products (SIC 152). These are localized, as expected, in *concelhos* pertaining to the coastal side of the country.

Overall Levels of Localization by Region. The standardized residuals from the Dirichlet-Multinomial regressions can also be used to measure the overall level of localization by regions. To gain some insight into this question we computed a new variable as the sum (by *concelho*) of all positive standardized residuals across 3-digit SIC industries. In figure 3 we show these data for the 275 Portuguese *concelhos*. The *concelhos* are classified according to the quartiles of the variable, with darker intensity corresponding with increasing values. As can be seen, excess of concentration is particularly important in the coastal corridor between Porto and Lisbon. More precisely, in this axis we found three different areas where localization economies (natural advantages of the regions) are important: a large stretch of territory (consisting of 23 adjacent *concelhos*) is found around the cities of Porto and Aveiro; a second large area is centered around the cities of Coimbra and Leiria; and a third but small one (grouping 8 *concelhos*) can be detected around the city of Lisbon.

Figure 3
Overall levels of localization by *concelhos*

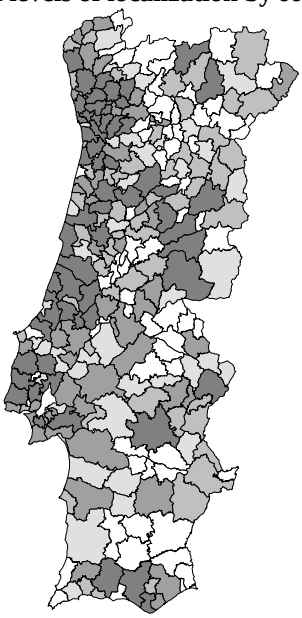


Figure 2 also shows that levels of spatial concentration explained by natural advantages or external economies associated with firms' clustering are high for a large area in the central hinterland of the country. This large area (around Serra da Estrela)

encompasses the *concelhos* of Viseu, Tondela, Mangualde, Guarda, Covilhã, Fundão, Seia, and Castelo Branco.

Beyond the Oporto-Lisbon coastal corridor and the central hinterland of the country around Serra da Estrela, there are a few isolated *concelhos* which also stand out. This is the case of the *concelhos* of Chaves and Mirandela (in the northeast part of the country), Elvas and Évora (in the Alentejo province), and Loulé, Silves and Faro (in the southern part of the country).

4.4.2. Evaluating the Impact of Changes on Location Factors

Given that we explicitly model the location decision process of firms, we are able to perform exercises of comparative statics to determine what changes would be obtained under an alternative scenario for the allocation of the regional resources. The results of this analysis will shed some light into the effect of potential policies aimed at influencing the location decisions of firms. Consider first the situation where there is a 1 per cent increase in the level of one of the variables entering the profit function, say variable k, in region j, everything else remaining constant.[16] The impact of that change on the probability of locating in that region for a given industry amounts to $\theta_k p_{j/\eta}(1-p_{j/\eta})$. Two obvious implications follow. First, the impact of the change is higher for the variable with the highest profit elasticity (θ) in absolute value. The second implication is that that impact will be higher for the more attractive regions (i.e. regions with higher $p_{j/\eta}$)[17] and even higher in regions that benefit from localization economies or natural advantages.[18] It is straightforward from here to infer that an across the board increase of 1 per cent in the level of a given factor will impact differently across regions and thus will change the spatial distribution of that industry.

But are we allowed to conclude that the level of localization (i.e. excessive concentration) of that industry will increase? To answer that question we need to measure the impact on the localization index $\tilde{\gamma}$ resulting from a change in one of the variables entering in the firms' profit function. It is obvious that anything that will directly increase industry profits will reduce the weight that localization economies (or natural advantages) have on driving the location decisions of firms. This can be ascertained by looking at expression (18). If we compute the elasticity of $\tilde{\gamma}$ with respect to one of the variables entering the profit equation, say variable k, we obtain $-\theta_k(1-\tilde{\gamma})$ (again we are taking into account the fact that all explanatory variables are already entered in logarithmic form). Thus, those variables that are more capable of affecting profits (with higher profit elasticities) are precisely the ones that offer the highest potential to counterbalance the effects of local spillovers and natural advantages of the regions. This means that if wages

---

16 Note that in our specification all variables are in logarithmic form.

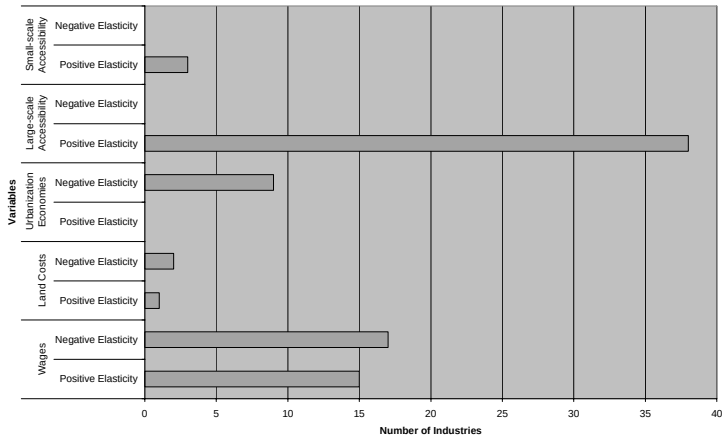
17 Unless $p_{j/\eta} > 0.5$, a very unlikely situation.

18 To get an idea about the extent of this difference across regions we evaluated, for a non-localized industry (SIC 311), the above expression for all the 275 *concelhos*. Next, we computed the ratio of the average value between the upper quartile (the more attractive regions) and the lower quartile of the distribution (the least attractive regions). We find a ratio of 48 indicating that a change of 1% in a given variable for the more attractive regions will have the same redistributive effect as a change of 48% in the least attractive regions. For localized industries that value will be much higher.

have the highest profit elasticity (assumed negative) then a decrease of 1 per cent in the average cost of the workforce across regions will increase profits everywhere and will diminish the relative importance of localization economies and natural advantages leading to a smaller level of excessive concentration, more than an equivalent percentual change in any of the other factors affecting profits. But, on the other hand, we can see in the above expression for the elasticity of $\tilde{\gamma}$, that the impact of any change is smaller for those industries that are more localized.

With that in mind, it becomes relevant to identify for each of the 3-digit SIC industries those factors with the highest profit elasticities.[19] In Figure 4 we summarize the results of our calculations. We find that wages (with a positive sign and thus capturing the cost of the workforce) have the highest profit elasticity for 15 industries (out of 92) while wages (with a negative sign and thus more likely to proxy the quality of the workforce) have the largest significant elasticity for 17 industries. Land costs and urbanization economies are the variables with higher significant impact for 3 and 9 industries, respectively. On the other hand, large-scale accessibility has the largest (positive and significant) impact for 38 industries in contrast with small-scale accessibility which is more relevant for only 3 industries. Thus, it seems fair to conclude that accessibility to the Oporto-Lisbon coastal corridor (and transportation costs) is the factor with the highest potential to reshape the Portuguese industrial landscape.

Figure 4
Highest profit-elasticities by 3-digit SIC industries

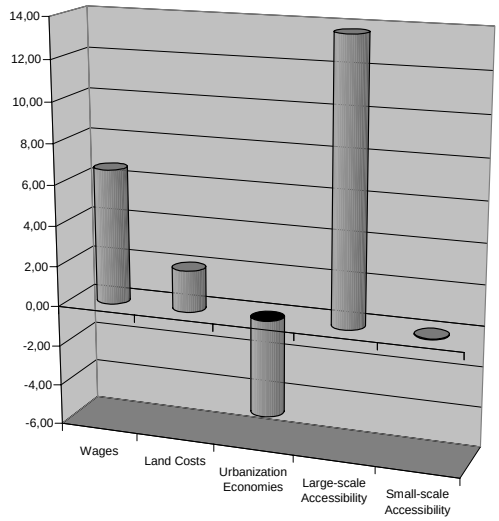


To gain additional insight into this matter we computed the mean impact on $\tilde{\gamma}$ across 3-digit SIC industries resulting from a 10 per cent increase (across regions) for each one of the variables entering in the profit equation. The results are shown in Figure 5. As can be seen, large-scale accessibility is the variable with the highest average impact. A 10

19 Elasticities were computed for 92 3-digit SIC industries. As indicated before, for a small number of industries (8) the model did not converge.

per cent increase in this variable across regions results in a 13,73 per cent average increase of $\tilde{\gamma}$ across 3-digit SIC industries, while the same averages for wages, land costs, urbanization economies and small-scale accessibility are 6,64, 2,05, 4,78, and 0,04 per cent, respectively. These results have relevant policy implications. Abstracting from issues pertaining to the cost of alternative policies, it seems that if public authorities want to reshape the industrial landscape and spread industries across the country, they are more likely to succeed if they invest in transportation infrastructures in order to reduce travel time distance (and improve accessibility) from *concelhos* to the Oporto-Lisbon coastal corridor.

Figure 5
Mean impact on $\tilde{\gamma}$ across 3-digit SIC industries



5. Conclusion

In this paper we have offered an empirical investigation into the level of localization of Portuguese manufacturing industries. We take a critical view of the existing methods for measuring localization and argue that despite its popularity, the EG index has some serious pitfalls that prevent it from being a satisfactory approach. Building on the McFadden's Random Utility (Profit) Maximization framework, we develop an alternative index which is more consistent with the theoretical construct underlying the work of Ellison and Glaeser. With our approach, we are able to simultaneously compute the locational probabilities and the localization index.

When we applied our methodology to Portuguese data, we found evidence that a high percentage of industries are localized, a result that corroborates previous analyses based on the EG index for others countries. But even though Portuguese industries are in

general excessively concentrated, the majority of industries exhibit modest localization levels and only a few sectors display a strong tendency to cluster. These sectors are mostly dominated by traditional industries for which localization is determined by the historical specialization of particular regions in Portugal. Past static externalities may justify the clustering of firms in these sectors. We also find that some more technologically advanced industries exhibit high or moderate localization levels. This result indicates that current dynamic knowledge spillovers may also be important in explaining the clustering strategies of Portuguese firms. As expected, some resource oriented industries are found to be highly localized.

We have also precisely identified regions associated with the most localized industries and measured levels of localization by regions. According to our results, spatial concentration induced by external economies and natural advantages of the regions are particularly important in the Oporto-Lisbon coastal corridor and in the central hinterland of the country around Serra da Estrela. Our model also allowed us to draw some conclusions regarding economic policy. First, we should note that the stimulation of economic activity through firm creation requires a much larger effort for those regions with a less favourable set of locational factors. Second, we note that localized industries are less likely to respond to policies aimed at redistributing economic activity. And finally, we find that accessibility to the more urbanized coastal side of the country (and transportation costs) is the variable with the highest potential to reshape the Portuguese industrial landscape, counterbalancing the weight that localization economies and natural advantages of the regions have on driving the location decisions of firms. Thus, if public authorities want to disperse industries across the country and attract private investment to new localities, they are more likely to succeed if they direct their efforts toward the improvement of transportation infrastructures aimed at reducing the travel time distance (and improve accessibility) from *concelhos* to the Oporto-Lisbon coastal corridor.

Appendix A

In the context of the EG model, the number of investments of a given industry in region j, conditional on the total number of investments in the industry, and on the vector of locational probabilities $(\mathbf{p} = p_1, p_2, ..., p_J)$, follows a binomial law with parameters:

$$E(n_j / \mathbf{p}) = np_j$$
$$V(n_j / \mathbf{p}) = np_j(1 - p_j)$$

We now define a raw index of concentration as:

$$G_F = \sum_{j=1}^{J} \left(\frac{n_j}{n} - x_j \right)^2$$

Expanding terms we obtain

$$G_F = \frac{1}{n^2} \sum_{j=1}^{J} n_j^2 + \sum_{j=1}^{J} x_j{}^2 - \frac{2}{n} \sum_{j=1}^{J} n_j x_j$$

and the expected value of the above equation gives:

$$\begin{aligned} E(G_F / \mathbf{p}) &= \frac{1}{n^2} E\left(\sum_{j=1}^{J} n_j^2 \right) + \sum_{j=1}^{J} x_j^2 - \frac{2}{n} E\left(\sum_{j=1}^{J} n_j x_j \right) \\ &= \frac{1}{n^2} \left(\sum_{j=1}^{J} np_j - np_j^2 + n^2 p_j^2 \right) + \sum_{j=1}^{J} x_j{}^2 - 2 \sum_{j=1}^{J} p_j x_j \\ &= \frac{1}{n} + \frac{(n-1)}{n} \sum_{j=1}^{J} p_j^2 + \sum_{j=1}^{J} x_j^2 - 2 \sum_{j=1}^{J} p_j x_j \end{aligned}$$

Applying the law of iterated expectations we get,

$$\begin{aligned} E(G_F) &= \frac{1}{n} + \frac{(n-1)}{n} E\left(\sum_{j=1}^{J} p_j^2 \right) + \sum_{j=1}^{J} x_j^2 - 2E\left(\sum_{j=1}^{J} p_j x_j \right) \\ &= \frac{1}{n} + \frac{(n-1)}{n} \sum_{j=1}^{J} \left(\gamma x_j - \gamma x_j^2 + x_j^2 \right) + \sum_{j=1}^{J} x_j^2 - 2 \sum_{j=1}^{J} x_j^2 \\ &= \frac{1 + \gamma(n-1)}{n} \left(1 - \sum_{j=1}^{J} x_j^2 \right) \end{aligned}$$

and, as in Ellison & Glaeser (1997), the estimator for γ is obtained by replacing the $E(G_F)$ by the observed value of G_F and solving for γ. The proposed estimator is:

$$\hat{\gamma}_A = \frac{nG_F - \left(1 - \sum_{j=1}^{J} x_j^2\right)}{(n-1)\left(1 - \sum_{j=1}^{J} x_j^2\right)} \ .$$

Appendix B

Under the null hypothesis that $\gamma = 0$, the $p_j = x_j$ for all j and the observed spatial distribution of the investments for the particular industry follows a multinomial distribution,

$$P(n_1, n_2, ..., n_J) = n! \prod_{j=1}^{J} \frac{x_j^{n_j}}{n_j!}$$

Because we can associate a probability of occurrence to each possible distribution of the n investments we may also construct a distribution for the estimator of $\hat{\gamma}_A$ under the null hypothesis that $\gamma = 0$. To do this, we may simply enumerate all possible values of the multinomial distribution. A simple example will help understand the argument. Suppose that we have 3 regions and 4 investments. Admit for the moment that ($x_1 = x_2 = x_3 = 1/3$). The next table lists all possible spatial distributions of these investments, the associated probability and the estimated concentration index ($\hat{\gamma}_A$):

Table A.1
Distribution of Investments by Regions

n_1	n_2	n_3	$\hat{\gamma}_A$	$P(n_1, n_2, n_3)$
4	0	0	1.00	1.2346%
3	1	0	0.25	4.9383%
3	0	1	0.25	4.9383%
2	2	0	0.00	14.8148%
2	1	1	-0.25	7.4074%
2	0	2	0.00	14.8148%
1	3	0	0.25	4.9383%
1	2	1	-0.25	7.4074%
1	1	2	-0.25	7.4074%
1	0	3	0.25	4.9383%
0	4	0	1.00	1.2346%
0	3	1	0.25	4.9383%
0	2	2	0.00	14.8148%
0	1	3	0.25	4.9383%
0	0	4	1.00	1.2346%

This information can be used to construct the distribution for $\hat{\gamma}_A$ which simply aggregates all common estimates and their probability. Thus, the distribution of $\hat{\gamma}_A$ given $x_1 = x_2 = x_3 = 1/3$, $n = 4$, and $\gamma = 0$ is:

Table A.2
Statistical Distribution of the Estimator

$\hat{\gamma}_A$	$f(\hat{\gamma}_A)$	$F(\hat{\gamma}_A)$
-0.25	44.44%	44.44%
0.00	22.22%	66.67%
0.25	29.63%	96.30%
1.00	3.70%	100.00%

From this simple example, we can see that if we had obtained an estimate of 1 for γ we could be fairly confident that $\gamma > 0$, given that the probability of that happening was only 3.7 per cent. But any other estimate would be a plausible outcome if the true value of γ were 0. Using this approach, we can test the probability that $\gamma = 0$ for any given number of investments and vector of locational probabilities.

However, it is not always feasible to construct the distribution of $\hat{\gamma}_A$ by numerically evaluating all possible distributions of investments by regions (as we did in Table A.1). The number of terms that will need to be computed amounts to

$$\binom{n+J-1}{J-1}.$$

If, for example, $n = 20$ and $J = 10$, then we get 10,015,005 different cases. If n is increased to 40 we will have 2,054,455,634 different cases. In this case, instead of computing the exact distribution we will randomly sample from this known distribution and generate an empirical cumulative distribution function for $\hat{\gamma}_A$. Thus, in our application we will test our hypothesis for each sector by generating a large number of draws (say 10,000) from a multinomial distribution with parameters $(n; x_1, x_2, ..., x_J)$. For each one of these samples, we will compute an estimate of γ and the value reported for our test will be the value of the empirical cumulative distribution evaluated at the observed value for $\hat{\gamma}_A$.

Table 1

Geographic Concentration, by Most Localized Industries According to $\tilde{\gamma}$

3-digit SIC Industry (Portuguese CAE-Rev2)	$\tilde{\gamma}$	**Number of Plants**	**Rank**		
			$\tilde{\gamma}$	$\hat{\gamma}_A$	$\hat{\gamma}_{EG}$
354- Fab. de motociclos e bicicletas	0.126	45	1	4	1
191- Curtimenta e acabamento de peles sem pêlo	0.115	110	2	1	5
362- Fab. de joalharia. ourivesaria e artigos similares	0.060	561	3	2	7
172- Tecelagem de têxteis	0.052	256	4	9	11
173- Acabamento de têxteis	0.048	275	5	16	21
193- Indústria do Calçado	0.048	1932	6	7	15
171- Preparação e fiação de fibras têxteis	0.048	226	7	17	22
351- Construção e reparação naval	0.038	155	8	19	31
323- Fab. de aparelhos receptores e material de rádio e de televisão. etc.	0.037	28	9	69	3
176- Fab. de tecidos de malha	0.036	284	10	14	14
335- Fab. de relógios e de material de relojoaria	0.032	15	11	69	94
247- Fab. de fibras sintéticas ou artificiais	0.029	12	12	69	9
152- Ind. transformadora da pesca e da aquacultura	0.029	106	13	22	23
192- Fab. de artigos de viagem e de uso pessoal	0.029	244	14	23	28
341- Fab. de veículos automóveis	0.027	15	15	69	80
177- Fab. de artigos de malha	0.026	715	16	15	25
275- Fundição de metais ferrosos e não-ferrosos	0.025	154	17	28	48
313- Fab. de fios e cabos isolados	0.021	31	18	69	64
205- Fab. de outras obras de madeira e ind. da cortiça	0.020	1197	19	3	6
264- Fab. de tijolos. telhas e outros	0.019	197	20	31	38
262- Fab. de produtos cerâmicos diversos	0.018	685	21	13	33
332- Fab. de aparelhos de medida. verificação. controlo. navegação e outros	0.018	29	22	24	89
343- Fab. de componentes e acessórios para veículos auto e seus motores	0.015	182	23	45	53
291- Fab. de máquinas e equipamentos diversos	0.015	117	24	35	68
244- Fab. de produtos farmacêuticos	0.015	101	25	10	26
202- Fab. de folheados. contraplacados. painéis. etc.	0.014	37	26	20	88
295- Fab. de outras máquinas e equipamento para uso específico	0.014	831	27	26	29

Table 2

Geographic Concentration, by Non Localized Industries According to $\tilde{\gamma}$

3-digit SIC Industry (Portuguese CAE-Rev2)	$\tilde{\gamma}$	**Number of Plants**	**Rank**		
			$\tilde{\gamma}$	$\hat{\gamma}_A$	$\hat{\gamma}_{EG}$
160 - Indústria do tabaco	0.000	2	76	69	100
355 - Fab. de outro material de transporte, ne.	0.000	4	76	69	99
296 - Fab. de armas e munições	0.000	7	76	69	97
363 - Fab. de instrumentos musicais	0.000	8	76	5	4
283 - Fab. de geradores de vapor (excepto caldeiras para aquecimento central)	0.000	8	76	69	13
242 - Fab. de pesticidas e de outros produtos agro-químicos	0.000	10	76	69	18
314 - Fab. de acumuladores e de pilhas eléctricas	0.000	11	76	69	39
271 - Siderurgia e fabricação de ferro	0.000	13	76	69	2
232 - Fab. de produtos petrolíferos refinados	0.000	13	76	8	16
364 - Fab. de artigos de desporto	0.000	13	76	69	90
353 - Fab. de aeronaves e de veículos espaciais	0.000	13	76	69	92
272 - Fab. de tubos	0.000	15	76	69	91
333 - Fab. de equipamentos de controlo de processos industriais	0.000	15	76	69	93
223 - Reprodução de suportes gravados	0.000	16	76	27	43
352 - Fab. e reparação de material ferroviário	0.000	20	76	69	45
322 - Fab. de aparelhos de rádio, televisão, telefonia e telegrafia	0.000	23	76	69	82
183 - Preparação e fabricação de artigos de peles com pêlo	0.000	27	76	69	52
334 - Fab. de material óptico, fotográfico e cinematográfico	0.000	28	76	21	98
268 - Fab. de outros produtos minerais não metálicos	0.000	29	76	69	17
365 - Fab. de jogos e brinquedos	0.000	29	76	69	41
273 - Outras actividades da primeira transformação de ferro e do aço	0.000	30	76	69	95
371 - Reciclagem de sucata e de desperdícios metálicos	0.000	37	76	69	84
265 - Fab. de cimento, cal e gesso	0.000	56	76	38	50
263 - Fab. de azulejos, ladrilhos, etc.	0.000	57	76	30	30
311 - Fab. de motores, geradores e transformadores eléctricos	0.000	83	76	69	65

Table 3 (cont.)
"*Concelhos*" Associated with the 27 Most Localized Industries

3-digit SIC Industry (Portuguese CAE-Rev2)	*Concelhos* and Standardized Residuals
354 - Fab. de motociclos e bicicletas	Águeda (17.61); Anadia (4.01)
191 - Curtimenta e acabamento de peles sem pêlo	Alcanena (19.46)
362 - Fab. de joalharia, ourivesaria e artigos similares	Gondomar (16.92); Póvoa de Lanhoso (2.72)
172 - Tecelagem de têxteis	Guimarães (10.65); Covilhã (6.20); V. N. Famalicão (4.18); Castelo Branco (2.71)
173 - Acabamento de têxteis	Guimarães (7.05); Santo Tirso (4.90); Barcelos (4.41); V. N. Famalicão (3.69)
193 - Indústria do Calçado	Oliveira de Azeméis (13.01); Felgueiras (9.96); S. M. da Feira (5.50); Guimarães (2.92)
171 - Preparação e fiação de fibras têxteis	Guimarães (8.21); V. N. Famalicão (4.88); Santo Tirso (4.84); Covilhã (4.24); Barcelos (3.63); Gouveia (3.01); Seia (2.21); Guarda (2.20)
351 - Construção e reparação naval	Seixal (4.53); Almada (4.46); Aveiro (3.31); Ílhavo (2.57); Palmela (2.54); Setúbal (2.35); Peniche (2.35); Olhão (2.09)
323 - Fab. de aparelhos receptores e material de rádio e de televisão, etc.	Mafra (6.85); Proença-a-Nova (4.85); Viseu (3.22); Sintra (2.96); Ovar (2.69); Sesimbra (2.61); Braga (2.47); Valongo (2.44); Alcobaça (2.23)
176 - Fab. De tecidos de malha	Barcelos (11.23); Guimarães (6.43); V. N. Famalicão (3.37); Marco de Canaveses (2.39)
335 - Fab. de relógios e de material de relojoaria	Alcobaça (8.79); Cantanhede (7.52); Guimarães (4.69); Mira (3.43); Braga (2.07); V. N. Famalicão (2.02)
247 - Fab. de fibras sintéticas ou artificiais	Portalegre (19.75); Vale de Cambra (3.79); Espinho (3.56); V. N. de Ourém (2.62); Ovar (2.61); Barreiro (2.05)
152 - Ind. transformadora da pesca e da aquacultura	Ílhavo (9.65); Peniche (6.91); Matosinhos (5.03); Vila do Conde (4.59); Olhão (3.43); Figueira da Foz (2.35)
192 - Fab. de artigos de viagem e de uso pessoal	Alcobaça (10.40); Gondomar (6.70); Braga (4.94); S. M. da Feira (2.71); V. N. de Gaia (2.49)

Table 3 (cont.)
"*Concelhos*" Associated with the 27 Most Localized Industries

3-digit SIC Industry (Portuguese CAE-Rev2)	*Concelhos* and Standardized Residuals
341 - Fab. de veículos automóveis	Mangualde (14.70); Ovar (7.19); Rio Maior (2.76); Abrantes (2.48)
177 - Fab. de artigos de malha	Barcelos (13.83); Guimarães (5.22); Fafe (5.11); Alcanena (4.67); V. N. Famalicão (2.68); Santo Tirso (2.20)
275 - Fundição de metais ferrosos e não-ferrosos	Águeda (8.70); Gondomar (5.58); Braga (4.07); V. N. de Gaia (3.65); Chaves (2.27)
313 - Fab. de fios e cabos isolados	Elvas (4.39); Ovar (4.31); Castelo Branco(4.05); Sintra (4.03); V. N. de Gaia (3.77); Guarda (3.53); Esposende (2.14)
205 - Fab. de outras obras de madeira e ind. da cortiça	S. M. da Feira (43.18); Montijo (3.74)
264 - Fab. de tijolos, telhas e outros	Tavira (8.43); Santarém (5.40); Torres Vedras (4.53); Porto de Mós (4.39); Águeda (4.36); Pombal (3.29); Chaves (2.29); Oliveira do Bairro (2.03)
262 - Fab. de produtos cerâmicos diversos	Barcelos (20.32); Alcobaça (10.24); Reguengos de Monsarás (5.15); Porto de Mós (3.73); Caldas das Rainha (3.15); Nazaré (2.37); Mafra (2.20); Redondo (2.00)
332 - Fab. de aparelhos de medida, verificação, controlo, navegação e outros	Mirandela (7.03); V. N. Famalicão (4.87); Sintra (4.04); Loures (3.74); Albergaria-a-Velha (3.08); Lagoa (2.92); Coimbra (2.30)
343 - Fab. de componentes e acessórios para veículos auto e seus motores	Águeda (6.58); V. N. de Gaia (5.31); Palmela (4.23); Oliveira de Azeméis (3.99); Loures (3.36); Cantanhede (2.33); V. N. de Cerveira (2.30); Maia (2.17)
291 - Fab. de máquinas e equipamentos diversos	Braga (6.79); Seixal (3.70); Fundão (2.84); V. N. de Gaia (2.78); Figueira da Foz (2.60)
244 - Fab. de produtos farmacêuticos	Tondela (5.03); Penafiel (4.19); V. N. Famalicão (3.01); Mortágua (2.97); Penela (2.88); Sintra (2.71); Condeixa-a-Nova (2.35); Coimbra (2.30); Oliveira de Frades (2.13)
202 - Fab. de folheados, contraplacados, painéis, etc.	Paredes (6.61); Mangualde (6.32); Viana do Castelo (4.16); Proença-a-Nova (3.50); Setúbal (3.38); Chaves (3.35); Nelas (2.99); Monção (2.39); Oliveira do Hospital (2.15); Castelo de Paiva (1.96)
295 - Fab. de outras máquinas e equipamento para uso específico	Marinha Grande (12.79); Leiria (6.14); Oliveira de Azeméis (4.24); Sintra (3.46); Alcobaça (2.97); V. N. de Gaia (2.52)

References

Bartik, T. (1985), 'Business location decisions in the United States: Estimates of the effects of unionization, taxes, and other characteristics of states', *Journal of Business and Economic Statistics* **3**(1), 14–22.

Bertinelli, L. & Decrop, J. (2002), Geographical agglomeration: The case of belgian manufacturing industry. Federal Planning Bureau of Belgium, Working Paper n.14.

Callejón, M. (1997), 'Concentratión geográphica de la industria y economias de aglomeratión', *Economia Industrial* (5), 61–68.

Cameron, A. C. & Trivedi, P. K. (1998), *Regression Analysis of Count Data*, Econometric Society Monographs, Cambridge University Press, Cambridge, UK.

Carlton, D. (1983), 'The location and employment choices of new firms: An econometric model with discrete and continuous endogenous variables', *The Review of Economic and Statistics* **65**(3), 440–449.

Coughlin, C., Terza, J. & Arromdee, V. (1991), 'State characteristics and the location of foreign direct investment within the United States', *The Review of Economic and Statistics* **73**(4), 675–683.

Devereux, M. P., Griffith, R. & Simpson, H. (2003), 'The geographic distribution of production activity in the UK', *Regional Science and Urban Economics* (WP 26/99). forthcoming.

Dumais, G., Ellison, G. & Glaeser, E. (2002), 'Geographic concentration as dynamic process', *The Review of Economics and Statistics* **84**(2), 193–204.

Duranton, G. & Overman, H. (2002), Testing for localisation using micro-geographic data. Centre for Economic Policy Research, Discussion Paper n. 3379.

Ellison, G. & Glaeser, E. (1997), 'Geographic concentration in u.s. manufacturing industries: A dartboard approach', *Journal of Political Economy* **105**, 889–927.

Figueiredo, O., Guimarães, P. & Woodward, D. (2002), 'Home field advantage: Location decisions of portuguese entrepreneurs', *Journal of Urban Economics* **52**(2), 341–361.

Friedman, J., Gerlowski, D. & Silberman, J. (1992), 'What attracts foreign multinational corporations? evidence from branch plant location in the United States', *Journal of Regional Science* **32**(4), 403–418.

Guimarães, P., Figueiredo, O. & Woodward, D. (2000), 'Agglomeration and the location of foreign direct investment in Portugal', *Journal of Urban Economics* **47**(1), 115–135.

Guimarães, P., Figueiredo, O. & Woodward, D. (2003), 'A tractable approach to the 3rm location decision problem', *The Review of Economic and Statistics* **85**(1), 201–204.

Hansen, E. (1987), 'Industrial location choice in Sao Paulo,Brazil: A nested logit model', *Regional Science and Urban Economics* **17**(1), 89–108.

Head, K., Ries, J. & Swenson, D. (1995), 'Agglomeration benefits and location choice: Evidence from japanese manufacturing investments in the United States', (3-4), *Journal of International Economics* **38**, 223–247.

Holmes, T. & Stevens, J. (2002), 'Geographic concentration and establishment scale', *Review of Economics and Statistics* **84**, 682–690.

Houdebine, M. (1999), 'Concentration géographique des activités et spécialisation des départments français', *Economie et Statistique* **326-327**, (6-7), 189–204.

Krugman, P. (1991), *Geography and* Trade, M.I.T. Press, Cambrige, MA.

Louviere, J., Hensher, D. & Swait, J. (2000), *Stated Choice Methods: Analysis and Application*, Cambridge University Press, Cambridge, UK.

Luger, M. & Shetty, S. (1985), 'Determinants of foreign plant start-ups in the United States: Lessons for policymakers in the Southeast', *Vanderbuilt Journal of Transnational Law* **18**, 223–245.

Mata, J., Portugal, P. & Guimarães, P. (1995), 'The survival of new-plants: Start-up conditions and post-entry evolution', *International Journal of Industrial Organization* **13**, 459–481.

Maurel, F. & Sedillot, B. (1999), 'A measure of the geographic concentration in french manufacturing industries', *Regional Science and Urban Economics* **29**, 575–604.

McFadden, D. (1974), Conditional logit analysis of qualitative choice behavior, P. Zarembka, ed., 'Frontiers in Econometrics', Academic Press, NY, chapter 4, pp. 105–142.

Mosimann, J. (1962), 'On the compound multinomial distribution, the multivariate beta-distribution, and correlations among proportions', *Biometrika* **49**, 65–82.

Schmenner, R., Huber, J. & Cook, R. (1987), 'Geographic differences and the location of new manufacturing facilities', *Journal of Urban Economics* **21**, (1), 83–104.

Woodward, D. (1992), 'Locational determinants of japanese manufacturing start-ups in the United States', *Southern Economic Journal* **58**, (3), 690–708.

REGULATION OF PROFESSIONS IN PORTUGAL: A CASE STUDY IN RENT-SEEKING*

Nuno Garoupa#
Universidade Nova de Lisboa
CEPR, London

July 2004

Abstract

In this paper it is analyzed the regulation of access to and activity of certain professions, namely lawyers and physicians. A quick review of the economic theory of regulation of professions, namely (a) Why regulate, (b) How to regulate, and (c) What to regulate is presented. An application to the regulation of professions in Portugal is developed, including recent evolution for the legal and medical professions, and we draw comparisons with other European countries plus EU law. We suggest an index to measure the quality of regulatory restrictions (hence exposing rent-seeking) in Portuguese professional activities. Some possible lines for institutional reform are detailed in the conclusion.

Keywords: Regulation, Rent-Seeking, Lawyers, Physicians

JEL classification: I18, J44, K29, L43, L51, L84.

* I am grateful to the Scientific Committee of the Bank of Portugal's Second Conference on the Economic Development of Portugal for helpful comments on earlier drafts as well as to Fernando Araújo, Fernando Branco, Anthony Ogus, and Frank Stephen for useful discussions and suggestions. The usual disclaimer applies

Faculdade de Economia, Universidade Nova de Lisboa, Campus de Campolide, P-1099-032 Lisbon, Portugal. Phone: 351-21-3801600. Fax: 351-21-3870933. Email: ngaroupa@fe.unl.pt

1. Introduction

A profession can be defined as an occupation with the following characteristics: specialized skills, that skill is partially or fully acquired by intellectual training, the service calls for a high degree of integrity, and it involves direct or fiduciary relations with clients (Ogus, 1994, page 216).

In this paper it is analyzed the regulation of access to and activity of certain professions, namely lawyers and physicians. The legal and medical professions (also notaries, pharmacists, and accountancy, less so architects and engineers, and even less economists and journalists) appear to be relatively highly regulated according to the European Commission. However, there are important disparities in levels of regulation across European countries. Austria, Germany, Greece, Italy, Luxembourg, and Portugal appear to be quite rigid whereas France, Spain, and Belgium have a relatively less strict approach to regulation of a significant number of professions. By contrast, Denmark, Finland, Sweden, Ireland, the Netherlands, and the United Kingdom have developed a more flexible regulatory framework for the professions (Stocktaking Exercise on Regulation of Professional Services, Overview of Regulation in the EU Member States, 2003).

Even though many economists think that professional regulatory activities are mostly explained by rent-seeking motivation, we find very different institutional arrangements across countries. Whilst ultimately governed by law and oversighted by some public official (judge, bureaucrat or legislator), these regulations are somehow delimited and enforced by the profession itself. Thus, it is of importance to assess which arrangements are more prone to private capture and suggest ways of reforming regulatory institutions.

The present paper has two major parts. In the first part, we present a quick review of the economic theory of regulation of professions, namely (a) Why regulate, (b) How to regulate, and (c) What to regulate.

In the second part, an application to the regulation of professions in Portugal is developed. We briefly analyze current regulation and recent evolution for the legal and medical professions, and we draw comparisons with other European countries plus EU law. A very quick look at the US is presented. We suggest an index to measure the quality of regulatory restrictions (hence exposing rent-seeking) in Portuguese professional activities. Some possible lines for institutional reform are detailed in the conclusion.

2. Theories of Regulation of Professions –Why

We can categorize the different theories in three groups:
2.1. Market Failure (Asymmetry of Information)
2.2. Public Interest (Apart from Market Failure)
2.3. Private Interest (Rent-Seeking)

2.1. Market Failure

The view that regulation pursues public interest in correcting for market failure (Posner, 1975; Noll, 1989) relies on the inefficiency of the market equilibrium. The main market failure that applies to professional markets is information asymmetry (Stephen and

Love, 1999). For most clients and consumers, professional services are credence goods (Darby and Karni, 1973). The consumer is less informed about the nature and quality of the service, and often relies on the expertise of the professional in order to assess (agency function) and implement the adequate strategy (service function). There is a potentially severe problem encompassing some kind of supplier-induced demand. Under these conditions the market usually fails to produce the socially optimal quantity and quality of the professional service. Some protection for the consumer of professional services is necessary to guarantee quality and mitigate inefficiencies. Protection of consumers frequently takes the form of regulation of the profession and its markets. Nevertheless we should have in mind that the costs generated by asymmetry of information must be balanced against the benefits of labour specialization. A reduction in information asymmetry might not be efficient if it also implies a substantial loss of benefits from labor specialization. For example, it is important to emphasize that the information asymmetry does not apply to all consumers. Repeat purchasers in the market for professional services are able to acquire experience and knowledge of the market which reduces the asymmetry of information (e.g., corporate clients in the market for legal services). Professionals must also take note of reputational effects which may arise from social networks even when most consumers are not repeat purchasers. Furthermore, when the service function is provided separately from the agency function, there is scope for revelation of information that limits opportunism (e.g., medical diagnosis and treatment by different medical doctors) (Emons, 1997).

Besides the moral hazard problem we have so far described, there is of course adverse selection since consumers cannot judge the quality of professionals. The "lemons problem" may arise, thus the need for some kind of licensing or an equivalent mechanism (Leland, 1979). Competition among professionals does not solve the problem due to the fact that good professionals may be driven out of the market by bad professionals given the inability of the market to pay for quality.

Another information problem may occur in the market for professional services, namely bounded rationality or rational ignorance. Consumers use simplified rules to process information rather than complex rational analysis. They also usually lack the education level, or even the intellectual ability, to be able to understand all available information on services in a correct way. Regulation is justified if the regulatory body has more information and expertise at its disposal than average consumers (Maks and Philipsen, 2002).

Legal professionals usually stress the need for self-regulation, arguing that severe losses would occur if poorly trained lawyers were allowed to perform services. This loss is particularly significant in the health sector, where injuries to the body and life represent substantial and eventually under-compensated damages. The consequences of medical maltreatment and legal misrepresentation go beyond the direct customer and generate serious negative externalities for the general public. Good health standards and the quality of the legal system are positively related to the quality of physicians and lawyers (Rubin and Bailey, 1994).

Finally a fifth form of market failure that justifies regulation is the existence of public goods. Information concerning the quality of professional services satisfies the conditions of non-rivalry and non-exclusivity in consumption. Therefore, there is the possibility that private provision (by professionals) of information is not efficient. This may

well justify mandatory information disclosure with respect to professional quality (Maks and Philipsen, 2002).

Regulation of the professional services can improve the market equilibrium. Asymmetric information causes moral hazard and adverse selection and eventually negative externalities for the general public thus precluding an efficient level of health and legal safety from being achieved by the market. The benefits of regulation include a decrease of search costs, improvements in service quality and more adequate supply of information concerning quality of professional services. Also, and very important, a reduction in risk is to be expected. In fact, due to the asymmetry of information, regulation could be the most adequate substitute for insurance (Zerbe and Urban, 1988).

Notice that the case for regulation in a public interest perspective is not controversial among economists, however it remains unclear which form of regulation should take place. If severe limitations to entry, prohibitions of advertising and regulation of fees are justified under a theory of public interest, it is still much of an open question. What seems clear is that in a market for professional services, where quality is uncertain, confidence and trust in the professionals is important for efficiency. After a couple of visits to a doctor a patient whose health problems have been solved may start trusting the doctor. An attorney who handles cases with care and arranges affairs with success may create a trust relationship with his clients. The problem is of course that most customers are not repeat purchasers, and even if they were, the costs of mistakes in the initial rounds could be very high.

Regulation and legal rules should aim at enhancing the trust relationship by economizing on information costs. There are three reasons why regulation should create a confidence premium (thus rewarding professionals above marginal productivity): (a) The cost of obtaining information is lower for the professional than for the client, (b) The information involved is productive, (c) The provision of true information must be rewarded in order to avoid strategic behaviour or opportunism. At a first observation, these reasons explain the need for minimum quality standards and even some regulation of fees, but severe restrictions on entry and on advertising do not seem justified (Van den Bergh, 1993).

2.2. Other Goals of Public Nature

Regulation of professionals may also pursue other goals of public nature that are not necessarily economic efficiency (i.e., correcting for information asymmetries and externalities). These goals may be explained by some kind of patronizing view of the government or community values, and usually are related to redistribution (Ogus, 1994, pages 218-219).

Confidence, honesty and trust might be values pursued by the government which in turn may actually promote greater social welfare and foster growth. The social willingness to pay for these values may be above its market or economic value, thus justifying government's intervention. A doctor or a lawyer in a small town may have a socially valuable role or function that goes beyond the professional service s/he provides. Redistribution in favour of the professional against the consumer is just a form of paying for these social services.

The problem with this explanation is that it can hardly apply to all professionals. If a doctor or a lawyer enjoys local monopoly power in a small town, then we expect s/he to earn extra profits (marginal revenue above marginal cost) that could be in some ways

justified by these other social services s/he provides. However, why a lawyer in a big city where s/he surely does not provide such social services should enjoy the some extra profits (due to regulation of fees) is hardly justified under a theory of public interest. Furthermore, why consumers of professional services should abstain from revealing their willingness to pay for those social services in a competitive market seems odd and could in fact conflict with an adequate welfare analysis (Kaplow and Shavell, 2002).

2.3. Private Interest

The last theory of regulation relates to private interest and relies on capture and collusion (Posner, 1974). From this perspective the regulation of markets for professional services is seen to arise and be sustained because it is in the interests of the members of the profession. It essentially allows for their cartel-like behaviour (Benham and Benham, 1975). As a result, the capture theory predicts that professional licensure should decrease the supply of professionals below social optimum, increase the prices charged by professionals, and increase existing professionals' incomes beyond marginal productivity, thus generating rents and quasi-rents (Stigler, 1971; McChesney, 1987; Olsen, 1999; Hadfield, 2000; Kleiner and Kudrle, 2000).

The most successful groups in obtaining wealth transfers are likely to be small, usually single issue oriented and extremely well organized. On the other side, those who bear the cost of paying rents are large fractions of the population, difficult to organize and with information problems. When these conditions are met, wealth transfers are expected to take place from the public as a whole to the very well-organized interest groups.

The government should protect the public from these interest groups but incentives to provide public interest legislation can be overcome by pressure by those benefiting from wealth transfers. Moreover, wealth transfers may not be recognized by the public in general and comparisons with other jobs and occupations can be difficult (Van den Bergh, 1993). Just take the case of confidence premium. Comparing figures about the income situation of professionals and other occupations may provide some evidence about how better paid they are, but we can hardly distinguish the confidence premium from pure rents. Unemployment within the profession below average unemployment could be an indication of rent-seeking but could just be that the population requires more professional services than other goods and services on average. Less regional variance with respect to payments could help to identify rent-seeking (payments less subject to local market and business conditions indicate some degree of market power), but at the same time it could be that the willingness to pay for health and legal professional services varies less across regions than for other goods and services. Market concentration indices for professional services can be constructed but are of course subject to the appropriate delimitation of the market (e.g., most large law firms are specialized in certain areas of the law) and the distortions of the public sector (e.g., the national health service is the major provider of medical services in many European countries).

The fact that rent-seeking behaviour is intrinsically difficult to identify, specially when there are sound public interest arguments for regulation to be made, makes rent-seeking and regulatory capture to be more likely. Nevertheless, it is possible to develop legal and political instruments to limit it. Promoting competition, in particular by making use of the internal European market (which should promote a free flow of professional services), auditing professional bodies (including comparative institutional analysis) or

forcing the separation of the service function from the agency function (e.g., medical diagnosis and treatment by different medical doctors) certainly helps to mitigate the problem.

2.4. A Compromise between Theories

In contrast to both pure private and public interest theories, the public and the professionals have an impact on the existent forms and contents of professional regulation. Thus, professionals will sometimes, but not always, are able to use regulations to limit supply and generate rents. On the other hand, public interest will be pursued sometimes, but not always (Peltzman, 1976). In fact, public and private interest theories mirror two distinct historical phases on economic research, emphasizing the corrective and the redistributive roles of regulation. The distinction between these two theories has lost validity even in economic theory due to game theory and institutional research (Hägg, 1997) that combine both.

Different institutional arrangements and regulations are consistent with both theories. In particular, self-regulation is not necessarily a sign of rent-seeking. Professional regulatory bodies are consistent with public interest theory. Identifying rent-seeking requires a more detailed analysis of the legal substance than just the legal form.

3. Institutional Arrangements -- How

There are several possible institutional arrangements to correct for market failure in the market for professionals as well as avoid private capture. We categorize these solutions in three groups:
3.1 Regulation by the Government
3.2 Self-Regulation
3.3 Regulation by Third Parties

3.1. Regulation by the Government

Regulation by the government usually includes quality regulation, certification and licensing. The government could subsidize high quality suppliers to ensure that they remain in the market even if adverse selection persists. Unfortunately it does not guarantee that the higher quality service will actually be supplied due to moral hazard. Second, penalties can be imposed on low quality suppliers and entry to the market could be restricted to some adequate standard (Dingwall and Fenn, 1987). These regulations however require a regulatory agency that must avoid capture and be able to do what consumers cannot: assess quality and signal it to potential clients (Stephen and Love, 1999). Apart from simple mandatory disclosure measures (e.g., professional specialty, professional education) and prohibiting what seems obvious misleading advertising (e.g., saying one is a lawyer or a doctor when one is not), effective quality regulation by the government seems difficult to imagine.

Under certification or licensing, a document (certificate or license) is awarded to an individual who satisfies certain conditions. These conditions may be education or training. The government as well as a private agency may certificate or license

professionals, and regulate professional education, compulsory periods of training, and performance requirements.

The difference between licensing and self-regulation is that while rules are issued by public authorities in both settings (since the professional body is entrusted with public authority), entry and performance are regulated by the state in the first case (eventually delegated to a private agency independent from the profession) and by the profession in the second case. The consequence is that self-regulation promotes strong professional association (as we know with lawyers and doctors) whereas licensing does not. A profession becomes only a real profession if it has the decisive power to fix remuneration; otherwise it is just a form of licensing (just like economists in Portugal and journalists almost everywhere).

The two arguments against licensing and thus making the case for self-regulation are the following: (a) It still does not solve the problem of asymmetric information because neither the government nor a private agency independent from the profession have better knowledge of the quality of the service the profession provides than the profession itself (though they might have better knowledge than the average consumer), (b) It is less flexible (in dynamic markets where innovation is important agencies should be able to change quickly) and generates costs to be borne by the government rather than by the profession itself (Miller, 1985). The second argument nevertheless has serious limitations. First, the profession can regulate fees to cover these costs (hence they will be borne by taxpayers or consumers in both cases). Second, rents created by the exercise of regulatory powers by the professional body can undermine flexibility. For example, rents may be used to successfully resist competition from other regulatory bodies offering more efficient rules (Curran, 1993).

3.2. Self-Regulation

Professional regulators have the necessary information to extract signals in markets for credence goods (the well-known specific knowledge argument by Miller, 1985) but can hardly avoid the ultimate form of regulatory capture. Yet this type of bodies persists in most jurisdictions. One view is that there is a social contract between the profession and the community in order to reduce moral hazard. Naturally safeguards are required in order to ensure the profession does not operate a cartel. Also various watchdogs are necessary (Dingwall and Fenn, 1987). Another view is that the reduction in costs of extracting information by professionals more than compensates for potential losses due to cartel-like behaviour (Ogus, 1995). These potential losses can be mitigated if there is more than one professional body in competition with each other (nevertheless in most jurisdictions professional bodies have a national or local monopoly), a large heterogeneous profession (Shaked and Sutton, 1982), and adequate legal instruments (e.g., efficient tort law) (Danzon, 1985 and 1991; Gravelle, 1990).

Though self-regulation solves the information problem we have discussed before, it is difficult not to expect that professional bodies use their regulatory powers to restrict competition somehow. Such rent-seeking behaviour, alongside other significant costs of administering the regulatory system, causes a significant deadweight loss.

In order to tackle this problem, we should have in mind four specific dilemmas: (a) It will be easier for professionals not to pass their better information and expertise to the users unless of course they have an interest in doing so (this will increase search costs for

the consumers since asymmetric information will not be reduced), (b) Professionals will induce demand of services that clients, if fully informed, would not require (inefficient allocation of resources), (c) Control and enforcement of quality standards will not be very effective due to collusion (hence we should investigate for sanctions for malpractice), (d) Fees will be set above confidence premium.

3.3. Regulation by Private Parties

Alternatives to professional regulation have been proposed, most of them never implemented. One solution could be independent rating agencies designed by repeat purchasers to perform the agency function on behalf of infrequent consumers (Stephen and Love, 1996). Others suggest deregulation via competition that will generate quality signals with adequate liability rules and removal of informational barriers (Leffler 1978; Klein and Leffler, 1981; Carr and Mathewson, 1988; Van den Bergh and Faure, 1991; Miller and Macey, 1995).

There has been a recent trend to relate effective regulation of professional services with litigation. The large scale of litigation in the US allows litigants to use their financial leverage to force changes of a regulatory nature and professionals to limit opportunism. If appropriate regulation does not exist for professional services, litigation can provide an effective substitute when it generates a transfer of wealth from the profession (the injurers) to the consumers (the injured) (Viscusi, 2002). Even so, there are important objections to the use of litigation as a way to stimulate effective regulation: (a) Consumers do not have the appropriate information to make a comprehensive analysis whether or not negligent behaviour, reckless attitudes, or professional malpractices were exercised (thus, litigation will usually be an inferior substitute for regulation), (b) Consumers may be opportunistic when making decisions with respect to filing lawsuits and settling out of court (e.g., nuisance litigation), thus generating too much litigation, (c) Litigation may not create the adequate incentives for efficient levels of professional services since it usually aims at providing compensation, (d) Litigation may fail in achieving efficient risk-sharing (restoring pre-accident levels of utility may not be possible, specially in the context of health effects).

In the context of medical malpractice there is some further controversy concerning the effectiveness and efficiency of litigation. Kessler and McClellan (1996, 1997, 2002a, 2002c) have shown that malpractice liability provides important incentives for medical care. Doctors in areas with greater malpractice pressure tend to use more defensive medicine, better treatment and medical high productivity seems to be positively related to the willing of patients to litigate (Olsen, 1997). However, once the incentives for hospitals and managed care organizations are explicitly taken into account, the empirical results are less striking. In fact, there is some debate among economists over optimal liability rules for physicians and health organizations, though most agree that tort reform and managed care function are substitutes in achieving incentives for adequate performance (Danzon, 1997; Kessler and McClellan, 2002b; Agrawal and Hall, 2003; Arlen and MacLeod, 2003).

4. Regulatory Instruments -- What

Currently, the literature has been focusing on controlling regulatory instruments and reflecting the private interest nature of their use. These instruments are:

4.1. Entry Restrictions

Entry restrictions are justified in order to assure quality of professional services but on the other hand they undermine competition by creating professional monopoly rights (Shaked and Sutton, 1981; Van den Bergh, 1999). These restrictions require candidates to have specialized skills acquired by intellectual education at university (in Europe, after obtaining a university degree; in the US, after completing studies in a professional graduate school) and by training (for a mandatory period). These requirements of education (a specific diploma) and traineeship may be determined both by the government and the professional body. It should be noted, nevertheless, than in Sweden and Finland there are no restrictions on who can provide legal advice and representation while in Spain only a university law degree is required.

Controls over these requirements can be exercised at three levels: (a) By defining the content of intellectual and training requirements, (b) By exercising influence over the organizations that educate and perform training of professionals (Shepherd, 2000), (c) By evaluating candidates after education and training at an exam or other type of screening device (eventually subjecting admission to some kind of *numerus clausus*). From a public interest perspective, we would expect some control over entry requirements but no strong influence over organizations that educate and perform training as well as a strict examination of candidates. Some level of education and training is indeed positive since the relationship between human capital and high quality services is expected to be positive. Moreover, reliance on self-regulation may increase the specificity of human capital investment and individual commitment to the profession (Donabedian, 1995).

Entry restrictions can also apply to para-professionals (e.g., para-medicals or other legal professionals) under the argument they supply an inferior quality service. However, they also do it at lower prices. It turns out that the entry of low quality para-professionals could be welfare improving (Shaked and Sutton, 1981). In other words, restrictions on para-professionals are expected to be undesirable unless the profits of the profession are given a sufficiently high weight in the social welfare (Gehring and Jost, 1995).

From our discussion it is clear that entry restrictions should be more similar to certification rather than a very comprehensive and strict examination of candidates before, during, and after education and training takes place. Notwithstanding, the absence of severe restrictions on entry does not necessarily imply competition. Professional markets tend to be spatially localized (Stephen and Love, 1999). Hence mobility might be seriously undercut and thus promote local monopolies (Pashigian, 1979). For example, in many jurisdictions lawyers may only appear before courts in the local area corresponding to the bar they have been admitted.

In Europe, many of the entry restrictions are in the process of being removed. The implementation of the Establishment Directive means that it is possible for lawyers and doctors qualified in one member state to become full members of the profession in another

member state without further examinations, though for example it does not apply to mobility for the legal profession between UK jurisdictions (Stephen, 2003). In the US, the lack of reciprocity between state bar associations seems to lead to lower number of practicing lawyers and higher incomes, though not to higher prices of legal services (Lueck et. al., 1995).

Entry restrictions can collide with competition law in Europe and anti-trust in the US. For many years, entry regulations issued by professional bodies were not subject to competition authorities. In Europe, the European Court of Justice (ECJ) explicitly recognizes that professionals may be subject to higher standards of conduct, and therefore accepts some restrictions. However, whether or not competition rules apply will depend on whether the professional body could reasonably have considered the restriction adequate for the proper functioning of the profession. Hence simply showing that the restriction itself is not necessary for proper functioning does not suffice for enforcing competition law (Andrews, 2002). As follows from the Wouters case (309/99), the ECJ precludes two ways to regulate professions. Either the government has empowered the professional body to regulate the profession without the government being fully involved, or the government retains the power to adopt professional rules. Regarding the latter, these professional rules will be considered state measures and excluded from the scope of EU competition law. The US case law however seems to point out in a different and more competitive direction by not tolerating outright collusion, for instance on prices, simply because it is the market for a professional service.

Even though entry restrictions are important and significant, entry to legal and medical professions has continued to grow in most jurisdictions. Obviously what is important is the growth in supply relative to demand (Stephen, 2003). Nevertheless, we should notice that empirical evidence points out that economic growth is negatively affected by more lawyers, the explanation being that their professional services do more redistribution than production (Murphy et. al., 1991).

4.2. Restrictions on Advertising

Restrictions on advertising can be justified under a public interest perspective inasmuch as they apply to other markets of goods and services. Advertising is a common method to provide information and, from a social welfare perspective, advertising should be allowed when it is productive, that is, it conveys important and relevant information to consumers concerning professional services. There is no reason to suppose that advertising of professional services should be subject to different regulations than those applied generally to other experience and credence goods and services. This argument conflicts with the claim used by professional bodies that advertising should be prohibited because it threatens the integrity and ethical responsibility of the profession by commercializing it. According to most professional associations, competition would be contrary to the dignity of the profession. However, as we observe in Europe, lawyers seem to be increasingly aware that dignity has a price. When Belgian lawyers seemed to lose business to Dutch and British law firms, the professional association decided to relax constraints on advertising (Faure, 1993).

Two kinds of advertising can be distinguished, price advertising being more controversial than quality advertising. When information about price is easier to obtain than information about quality (which is true for experience and credence goods but not for

search goods), increasing the availability of price advertising might discourage quality competition and encourage price competition, leading to a degradation of the average quality in the market (Cave, 1985). This argument may support some restrictions on price advertising, but not necessarily banning it.

The general conclusions of empirical evidence seem to be that restrictions on advertising increase the price of professional services and that the more advertising exists the lower the price is. However, there are several articles that contradict these findings (Rizo and Zeckhauser, 1992; Love and Stephen, 1996). There is no systematic evidence that distinguished between the effects of the two forms of advertising (Stephen, 2003). Nevertheless, quality advertising is much more common than price advertising (Stephen, Love and Peterson, 1994).

Even more difficult to understand is why physicians are not allowed advertising, but managed care organizations can do it (e.g., Médis in Portugal). They operate in the same market for professional services and there is no economic reason to justify why physicians cannot advertise in price and quality but managed care organizations can.

4.3. Restrictions on Fees

Restrictions on fees can be seen as way of assuring the confidence premium to professionals. Fees can be subject to control by the profession itself, by the courts or by the government by use of mandatory fee schedules. Over time, in most jurisdictions, mandatory scales have been transformed into recommendations. However, in Germany legal fees are still determined by the government. In Belgium and the Netherlands a recommended legal fee schedule is produced by the professional body and in Belgium there is a recommended minimum. Medical fees are set by the government in most public health services (e.g., NHS in the UK or SNS in Portugal) or by managed healthcare organizations (e.g., Médis in Portugal).

Price fixing is very restrictive and not very common. Moreover, it is unclear if it enforces high quality production (it seems it would if quality were either high or low and with homogeneous consumer preferences, Maks and Philipsen, 2002). Recommended fees suggest a more sophisticated approach to cartel-like behaviour. Though we would expect recommended fees to be seen as mandatory by the profession, the evidence provided by Shinnik and Stephen (2000) for conveyancing markets in Scotland and Ireland goes on the opposite direction. The authors nevertheless recognize that these markets satisfy the necessary conditions for successful deviations from collusive agreements. Another possibility is that recommended fees provide a focal point against which professionals discount thus colluding at a lower level (Stephen, 2003).

Limitations on fee contracts (e.g., contingent fee contracts in the market for lawyers is forbidden in Europe) are more difficult to justify on the basis of quality assurance. Moreover, the enforcement of limitations on fee contracts is costly and generates incentives for bargaining on the shadow of the law (e.g., informal contingent fees in Europe). In fact contingent fees for both legal and medical professional services would solve the moral hazard problem. The fundamental argument put against contingent fee contracts in the legal profession is that they conflict with the principle that lawyers should not have a vested interest in the cases they take. There could be a conflict of interest between client and lawyer over if and when to settle. The determination of an appropriate fee if settlement takes place would of course solve the problem. Also, we would expect

well-informed clients to prefer an hourly fee contract (and avoid conflict over settlement) whereas less experienced litigants would prefer contingent fee contracts.

Professional bodies can also manage the subsidies the government supplies to consumers of professional services, usually the national health service for health services and legal aid for legal services. The costs of legal aid and national health services have been growing rapidly. Usually it is caused by the increasing number of cases, rather than by fees paid to lawyers or physicians. Though these fees are usually much lower than normal fees, the profession can use them as a way of attracting consumers. Professionals have no clear incentive to avoid using government subsidies to generate oversupply of services.

4.4. Restrictions on Organizational Forms

Special regulations apply to law and medical firms. Restrictions on organizational forms are difficult to justify by public interest. If some aspects of professional services may favour partnerships rather than incorporation, we should expect the market to solve that, not the professional body.

Common organizational restrictions exclude incorporation (even where incorporation is permitted usually unlimited liability is maintained and the directors of the firm must be professionals) and multidisciplinary partnerships (i.e., involving members of more than one profession) from possible organizational forms. The usual justification for these restrictions is agency costs. Effort in production and quality are difficult to measure by others outside of the profession, thus making sole practitioners or professional partnerships the most likely form of organization where adequate incentives will be less costly to be designed (Carr and Mathewson, 1990; Matthews, 1991). The problem of course is that by banning other organizational forms, specialization of professionals beyond particular aspects of their service (thus lowering the cost of providing services) and economies of scope (by providing a ``one stop shopping'' including lawyers, accountants, surveyors or medical doctors, dentists, and beauty consultants) are lost. For example, in the European countries where multidisciplinary partnerships are permitted, commercial law is increasingly dominated by the legal branch of the major international accounting firms (Stephen, 2002).

A second type of restrictions on organizational form concern the separation between the service function (assess or diagnosis the problem) and the agency function (implement the correct solution). This separation limits opportunism and creates incentives for reveal of information (Emons, 1997). However, it can be seen as prohibition on vertical integration between different stages in production, thus generating costs in terms of technology (economies of scale) and agency costs (hold-up problem). The issue then is whether or not the benefits from formally separating the roles outweigh the costs (Stephen, 2003).

In the UK, as well as in Ireland and most of Australia, the legal profession has two branches: solicitors and barristers. Solicitors provide legal advice to the public and have rights of audience in the lower courts. Barristers have the rights of audience in higher courts and can be commissioned to advise solicitors, and they provide the majority of judges in the higher courts in later stages of their career. A member of one profession cannot become a member of the other. The debate over the efficiency of separating the legal profession in the UK is inconclusive (Bishop, 1989; Ogus, 1993; Bowles, 1994).

4.5. Restrictions on Conduct

The introduction of professional standards and ethics generates a number of costs, including administrative costs (defining, monitoring, and enforcing quality), compliance costs (from fulfilling professional obligations), and opportunity costs (since opportunistic behaviour is restricted) (Ogus, 1994).

Professionals are expected to pursue an agenda to minimize these costs. They will lobby for their own quality level and standards (Hau and Thum, 2000). A standard can be an effective mechanism to protect insiders from competitors by imposing their own quality standard thus reducing to zero compliance costs. On the other hand, a conflict between the government and the professions with respect to accepting and formally observing conduct rules is not likely, because professionals are usually involved in the actual formation of these rules (Maks and Philipsen, 2002).

Administrative costs will depend on how the professional body regulates the conduct of professionals. Many forms of conduct regulation can be found in the professional rules. A code usually describes the tasks and duties of the profession and is often called professional ethics. The professional body also establishes disciplinary procedures in case the restrictions on conduct are violated. These rules usually define under which conditions professionals might be sanctioned and eventually expelled from the profession.

There are two reasons why the enforcement of restrictions on conduct is not expected to be high. First, it is not a problem of controlling entry, but rather of controlling exit. There are clear incentives to avoid conflicts within the profession and make exit too easy. Second, the alternative mechanisms (litigation in court) still rely too much on the profession. By controlling the production of expert witnesses (directly, by providing and managing expert witnesses; indirectly, by training them), the professional body may block any attempt to force physicians and lawyers to leave the profession for violating professional conduct or gross malpractice. Naturally, in most countries, professionals are subject to contractual and extra-contractual liability, however it is difficult for judges to make a decision on medical malpractice or negligence in preparing a lawsuit if expert witnesses are not available.

Some limitations to the discretion professional bodies have in dealing with restrictions on conduct have been emerging out of international professional federations (though these are mostly recommendations) and to some extent by EU directives on professional services (not surprisingly usually perceived by professionals as intrusions into national legal and medical culture). However, evidence points out those most disciplinary actions are taken for lack of dignity or improper behaviour towards other professionals rather than professional malpractice (Faure, 1993; Hellingman, 1993).

In the US, lawsuits for medical negligence are quite too frequent nowadays (some people talk about a medical malpractice crisis), but were very infrequent 50 years ago. Physician liability existing prior to 1960s might actually have been too low, resulting from capture and the consequent use of self-regulation to deny expert witnesses testimony in malpractice cases. However, after the 1960s, it became much easier to obtain expert witnesses due to the erosion of local medical societies in disciplining unethical practices and local rules (Olsen, 1997). The consequence was a blow up of litigation over medical malpractice and thus the current need for tort reform in medical negligence (Miller, 1997; Dauer and Marcus, 1997; Sloan and Hall, 2002; Fine, 2003). Liability for medical

malpractice is also of growing importance in European tort litigation. Contrary to the US experience, the medical malpractice explosion does not seem to have come to an end yet (Faure and Koziol, 2001).

4.6. A Guideline for Research on Rent-Seeking

Table 1 summarizes most of the discussion we have presented. It also suggests some guidelines to identify rent-seeking behaviour from the profession. We will use these results while presenting the well-known methodology developed for European comparative analysis of professions (Faure et. al., 1993).

Table 1

Self-Regulation of Professions

	ENTRY RESTRICTIONS	FEE RESTRICTIONS	ADVERTISING RESTRICTIONS	ORGANIZATION RESTRICTIONS	CONDUCT RESTRICTIONS
PUBLIC INTEREST	MINOR	MINOR	PRICE	NO	MORE ON SUBSTANCE
PRIVATE INTEREST	SEVERE	SEVERE	PRICE QUALITY	YES	MORE FORMAL

5. Portuguese Experience

As far as I know and have been able to look for, there is no previous economic analysis of the market for professionals in Portugal with the possible exception of Amorim and Kipping (1999). In this part of the paper we investigate the current regulatory framework in Portugal with respect to lawyers (Ordem dos Advogados) and medical doctors (Ordem dos Médicos). The choice of these two professions is justified by the fact that, in order to make international comparisons, they are easily defined and delimited in terms of the services they offer. Other professional bodies are more difficult to compare due to less well-specified services (e.g., Ordem dos Economistas). Second, this group of professions has been the focus of several controversies and, to some extent, the Portuguese Government has recently enacted reforms of their regulatory setups. We find evidence of rent-seeking in the way these markets are regulated, though more in the market for medical services than for legal services.

The Portuguese case is presented in a comparative way within the Western world. We make detailed references to Spain since neither Faure et. al. (1993) nor the most recent research report on legal services by Paterson et. al. (2003) present an overview of this country. Details on US, UK, the Netherlands, Belgium and Germany are not presented since they can be found at Faure et. al. (1993) for legal and medical professions. For legal services, a detailed report for Denmark, Italy, France, UK (England and Wales) and Germany is available at Paterson et. al. (2003).

5.1. Lawyers

5.1.1. Entry Restrictions

In Portugal, intending lawyers (advogados) must have a recognized law degree. The organization of professional training after graduation from law school is within the competence of the professional body (Ordem dos Advogados, created by the Government in June 1926). There are six districts (Lisbon, Oporto, Coimbra, Évora, Algarve, Madeira and Azores) with competence to regulate the traineeship. Training follows for a mandatory period of eighteen months. This training period compromises a bar examination and supervised practice by a senior lawyer who must attest the moral and professional capacity of the trainee.

Lawyers have a very wide and nearly exclusive power of representation in courts as well as legal advising. The use of a lawyer is obligatory in court cases in which ordinary appeal is admissible, in legal action in which appeals are always admissible independently of the value of the case, and in appeals and in legal action in superior courts (article 36 of the Civil Procedure Code). Legal consultancy and legal advise are an exclusive power for lawyers with very minor few exceptions (article 53 of DL 49/84, Portuguese Bar Statute, Estatuto da Ordem dos Advogados).

The profession in Spain is organized by the professional body (Consejo General de la Abogacia Española). There are eighty-two districts, though only three, apart from Madrid, have over five thousand resident lawyers. Before taking up pursuit of the profession an oath to observe the constitution and judicial order must be taken. Once inscribed, the lawyer is immediately vested with all the rights and duties of a full member of the profession. There is no requirement for any further training.

Most entry regulations applied in Europe and in the US are similar to Portugal and Spain. Lawyers must belong to their professional association. There are no competing bar associations in the same jurisdiction. They are usually powerful interest groups.

Educational requirements do vary. A law degree is enough for practicing law in Spain and in the US, but not in most countries. Quite extensive mandatory training periods exist followed up by examination. Making licenses dependent on requirements of continuing education is not practiced, but professional associations run courses and seminars in joint ventures with law schools and law firms to help updating knowledge.

In general in Europe lawyers can plead before any court. There are however limitations in UK, Germany and the US. The division between barristers (specializing in advocacy) and solicitors (specializing in advice to clients) is only observed effectively in the UK and Ireland.

The evolution of the legal profession can be observed in Table 2, where we show the number of lawyers per 100,000 in several countries. One can see immediately that jurisdictions with less restrict entry rules (Spain as well as the US) have a substantially higher number of lawyers per capita. Not surprisingly, the Spanish Government pressed by the Spanish Bar is preparing a new law that will make entry more restrictive and stipulate mandatory training. Madrid alone had 40,467 resident lawyers in 2001 (twice as much as Portugal) followed by Barcelona with 14,193 and Valencia with 8,015. The Portuguese Bar has also expressed some concern over the growing number of lawyers (easily explained by the development of numerous private law schools after 1987 and new public law schools in the middle of the 90s), in particular looking at the growth rate between 1980 and 2000 (the

highest of the sample in Table 2). There were 5,134 lawyers in 1980; 11,319 lawyers in 1990; and 18,629 lawyers in 2000 according to figures published by the Portuguese Bar. The figure of 24,000 is expected to have been achieved in 2003. Predictably minor legal consulting services that were not under the obligation of membership in the professional body are now being questioned (Guedes da Costa, 2003, page 139).

European directives (namely Directive 77/249, Directive 89/48, and Directive 98/5) have been implemented. The regulation approved by the Portuguese Bar in 1994 specifies the requirements necessary for the inscription of EU lawyers, by which an exam, written and oral, in Portuguese has to be successfully passed. The Spanish Ministry of Justice has regulated in 1996 the requirements necessary for accession to the profession by EU lawyers; an aptitude test has to be passed. More severe restrictions are applied in Belgium where registration with the local bar is subject to a law degree from Belgium.

Table 2

Number of Lawyers per 100,000

COUNTRY	1983	1990	2000	% 1983-2000
PORTUGAL	54 (1980)	116	188	248% (1980)
SPAIN	135	-	241	79%
UK	100	-	283	183%
GERMANY	70	-	142	103%
BELGIUM	122	137	155	27%
NETHERLANDS	30	57	77	157%
US	250	261	338	35%
ITALY	80	-	160	100%
FRANCE	51	-	68	33%

Source: Faure et. al. (1993), World Bank Legal and Judicial Reform Practice Group, Council of the Bars and Law Societies of the European Union, own calculations.

5.1.2. Other Restrictions

European law bars subscribe to a professional code (the so-called Code of Conduct for Lawyers in the European Union) that provides minimum common standards, though it is recognized (it says in its preamble that it is not possible nor desirable) that a general unified regulatory framework should not be developed. Common standards include: (a) Personal advertising and publicity is forbidden unless explicitly allowed by the local bar; (b) Contingent fees (*pactum de quota litis*) are banned; (c) Multidisciplinary partnerships are restricted since lawyers cannot share honorariums and fees with other professionals unless explicitly allowed by the local bar; (d) Lawyers should not conflict with other lawyers, but if they do, the local bar should be asked to intervene before the case goes for litigation; (e) A lawyer should not accept instructions to represent a client in substitution for another lawyer in relation to a certain matter if the client has not fully paid and reimbursed the first lawyer. The Code also refers to the ``corporate spirit of the profession'' by which a relationship of trust and cooperation should be developed (a

principle regulated under the name of duty of solidarity among lawyers, for example, by article 83 of DL 49/84, Portuguese Bar Statute, and article 34 of Real Decreto 658/2001, Spanish Bar Statute, Estatuto General de la Abogacia Española). Nevertheless, with respect to this matter, in Portugal, rules tend to be stricter, for example, lawyers are forbidden from expressing publicly opinions concerning legal matters taken care by another lawyer unless agreed by the latter (article 86 of DL 49/84, Portuguese Bar Statute).

More recently, the association of European law bars has emphasized that: (a) Contingent fees (i.e., an agreement between a lawyers and his client by virtue of which the client undertakes to pay the lawyer a share of the result regardless of whether in the form of money or any other benefit) being forbidden is a necessary rule of the profession; (b) Fee sharing with non-lawyers is a consequence of the duty of confidentiality and avoidance of conflicts, thus multidisciplinary partnerships should not be permitted since they offend the core values of the profession; (c) These restrictions cannot be considered a restriction of competition under EU competition law since they are applied in the specific context of a profession; (d) Comparative conclusions with respect to different regulations across Europe should be avoided because they follow from legal and cultural intrinsic differences, and are respected by the jurisprudence of the ECJ (CCBE Response to the European Commission Competition Questionnaire on Regulation in Liberal Professions and its Effects, May 2003, in relation to the report by Paterson et. al., 2003).

Price advertising is banned in most jurisdictions, except the US (though regulated by each state bar), under the cover that comparative advertising is strictly prohibited. Quality advertising is usually allowed for partnership but not for sole practitioners. Competition within the European Union has pushed bars to relax somehow the constraints, a pressure also felt in Portugal (Boletim da Ordem dos Advogados, May 2003). Nevertheless, publicity cannot refer to any particular case or reveal names of current or previous clients (article 80 of DL 49/84, Portuguese Bar Statute). Overall, the regulation of publicity for legal services is still more restrictive in Portugal, Spain and France and much less restrictive in the UK and the Netherlands, Germany and Belgium being intermediate cases with a trend for deregulation (Guedes da Costa, 2003, page 181).

The arguments against deregulation of publicity for legal services used by the professional body are: (a) publicity misleads the public and it has a negative effect on the quality of the profession (untrue of quality advertising and probably true in price advertising), (b) it is very expensive (we do not know since the current market is very thin due to strict regulations), (c) it generates unfair competition because only the large law firms can benefit from advertising (the US experience shows otherwise), and (d) it is against professional ethics by violating the so-called principle of non-commercialization of legal services (Guedes da Costa, 2003, page 182).

With respect to fees, in most countries prices can be freely negotiated and usually more competent lawyers charge higher fees, except in Germany. Recommended fees exist in Belgium, the Netherlands, and to some extent in Portugal (Guedes da Costa, 2003, page 208). Fees are usually based on hours worked, litigation value (except in Belgium), and complexity of the case. Contingent fees are allowed in the US but not in Europe (in Portugal contingent fee or *pactum de quota litis* is strictly forbidden by article 65 of DL 49/84, Portuguese Bar Statute). Usually legal fees take the form of hourly fees or flat fees (in Portugal, the so-called avença). A first exception was developed in UK where a lawyer receives an up-rating on the normal fee if the case is won which is not related to the value of damages (conditional fees). Similar arrangements are now being allowed in many

countries, and under consideration in Portugal (Boletim da Ordem dos Advogados, May 2003).

Legal aid is usually run by independent government funded bodies (Netherlands and US), legal aid boards (Scotland and Spain) or courts (Germany), the exceptions being Belgium and the new system in Portugal where legal aid is funded by the Government but run by the professional body. A new protocol between the Portuguese Ministry of Justice and the Portuguese Bar (March 2003) has just transferred to the professional body the management and control of legal aid by creating the Institute for Access to Justice (Instituto de Acesso ao Direito). Though it has been presented as a way to improve performance and quality of legal aid (under the usual argument that lawyers know best how to evaluate and control quality), our expectation is that an increase in Government's expenditure on legal aid will follow.

Apart from disciplinary action (in Portugal, it is an exclusive power of the professional body, article 90 of DL 49/84, Portuguese Bar Statute, though before the creation of the Portuguese Bar in 1926, courts could engage in disciplinary action), lawyers can be liable for their wrongdoings while serving their clients. Not only tort liability is not well developed for legal services in Portugal (in fact, for professional services in general), but there is no mandatory liability insurance in clear contrast with recommendation by the Code of Conduct for Lawyers in the European Union (apparently it will be introduced after a new law concerning the organization of law firms is approved, see below). The current situation is justified by the professional body under the argument that, while in other countries such as the US and the UK there is a system of competitive legal services (the so-called advocacia livre) necessary under the supervision of courts, in Portugal as well as in Spain, France or Italy, we have a system of collegial legal services (the so-called advocacia colegial) under the supervision of professional bodies with regulatory exclusionary powers (Guedes da Costa, 2003, page 329).

The structure of legal firms in Europe, including Spain and Portugal, has been changing since the 90s. Sole practitioners or small professional partnerships have been growingly replaced by large professional partnership, corporations (where they are allowed, not in Portugal) and multidisciplinary organizations (not allowed in Portugal and Spain). These changes in the structure of legal firms have pushed the Portuguese Government and the Portuguese Bar to elaborate a new framework for law partnerships (first steps were taken with DL 237/2001 but a new law is being prepared by the Portuguese Law Partnerships Institute). Though they explicitly recognize that EU competition in the market for legal services is the main cause of this reform, the new law will still not allow incorporation (due to the so-called principle of non-commercialization of legal services) and multidisciplinary organizations (the argument here is that these organizational forms dilute the control mechanisms of professional quality). Limited liability (subject to mandatory liability insurance) and mandatory rules for promotion within the partnership are the major changes. Mergers of law partnerships are subject to approval by the professional body and cooperation (the so-called consórcio) between them is allowed for limited periods.

The entry of foreigner law firms or partnerships in the Portuguese market for legal services is not helped by current regulations. In contrast with the Spanish case (article 28 of Real Decreto 658/2001, Spanish Bar Statute), where entry regulations have been reformed to allow for the establishment of multinational law firms and partnerships (the use of their original denomination as well as their original organizational form are allowed under

certain conditions), the situation in Portugal is of a more closed market and very strict regulations apply to the use of original denominations and organizational forms.

5.2. Physicians

5.2.1. Entry Restrictions

According to the Portuguese Medical Association (Ordem dos Médicos, created by the Government in 1938), there were around 29,000 medical doctors in Portugal in 2000, more than 21,000 employed by the national health system (SNS), and all registered at one of the three regional sections of the professional body (Norte, Centro and Sul). General practitioners account for around 35 percent and hospital doctors (secondary care) are more than 46 percent. From Table 3, it is clear that there has been a steady increase in the number of physicians in Portugal, though still below the European average (the UK being the exception), at an intermediate growth rate (above UK, US, and Germany but below all the other countries). The most worrying statistics is however that the number of doctors entering the workforce in the period 1992-2000 is the lowest in Europe (European Observatory on Health Care Systems, 2000), 4.1 new doctors per 100,000 in Portugal against 9.1 in the Netherlands or 10.9 in Belgium (see Table 4).

There are currently seven medical schools in Portugal (two open after 2000). All medical training programs are similar. Three years of core basic sciences are followed up by three more years of clinical program oriented to specialization. After graduating, a general internship for 18 months takes place. After successful completion of the internship, a physician is free to practice medicine without supervision. However, if a medical career in the national health system is the objective, further training is required for specialization, from three to six years (there is a proposal by the Government to reform medical internship and further training for specialization, reducing its duration in two years from 2007 on). The Government and the Portuguese Medical Association are jointly responsible for certification of specialist training. The most popular specialty is internal medicine (almost 20 percent) followed closely by gynecology, general surgery and pediatrics (data from European Observatory on Health Care Systems, 2000).

According to international databases (Table 3), Spain has one of the highest relative number of physicians (the second highest after Italy). Doctors are organized in fifty-two provincial associations belonging to the Spanish Medical Association (Consejo General de Colegios Oficiales de Médicos). Many work for the national health system, but the growing number of doctors has pushed for the development of the private sector in the early 90s (Mutuas). In Spain, the postgraduate training of medical specialists and general practitioners is structured in a system (the MIR) of practical work for three to five years. Certification in a certain specialty is governed by a national commission made up of representatives of university professors, scientific societies and the Spanish Medical Association.

Entry regulations are not very different across our sample of countries, with the exception of the Netherlands where registration is not required. As a consequence, a complex insurance system has been developed in the Netherlands to protect consumers. One of the consequences is that now it is actually easier for a doctor registered in a professional body in another country of the European Union to practice medicine there than a Dutch doctor (because the insurance premium is much lower for the former).

European directives (namely Directive 93/16) have been growingly implemented. The medical diplomas and certificates obtained in any state of the European Union are recognized by each member state (Directive 93/16 complemented in details by Directive 97/50, Directive 98/21, Directive 98/63 and Directive 99/46). After registration in the professional body, a physician can practice under the rules of the country (given the recognition by the ECJ of the so-called principle of double deontology). Given the shortage of physicians in Portugal and the high number of doctors in Spain, many Spanish doctors have made use of this European legislation to establish themselves in Portugal.

Table 3

Number of Physicians per 100,000

COUNTRY	1980	1990	2000	% 1980-2000
PORTUGAL	205	285	318	55%
SPAIN	217	225	429	98%
UK	133	140	181	36%
GERMANY	245	298	358	46%
BELGIUM	232	350	419 (2001)	81% (2001)
NETHERLANDS	192	251	328 (2001)	71% (2001)
US	200	245	280 (1999)	40% (1999)
ITALY	335 (1983)	490	567 (1999)	69% (1983-1999)
FRANCE	188 (1983)	265	330 (2001)	76% (1983-2001)

Source: Faure et. al. (1993), WHO Regional Office for Europe Statistics, World Bank Health Indicators, OECD Health Data, own calculations.

Table 4

Number of Physicians per 100,000 Entering the Workforce

COUNTRY	1992	1993	1995	1996	2000
PORTUGAL	-	-	4.1	-	-
SPAIN	-	-	-	-	-
UK	-	-	-	-	-
GERMANY	-	-	15.4	14.8	-
BELGIUM	-	-	10.9	10.1	-
NETHERLANDS	-	-	9.1	9.9	-
US	-	-	5.9	-	5.6
ITALY	15.3	-	-	-	-
FRANCE	-	8.5	-	-	-

Source: European Observatory on Health Care Systems, US National Center for Health Workforce Analysis.

5.2.2. Other Restrictions

Portuguese physicians must comply with a professional code issued by the Portuguese Medical Association, the Portuguese Medical Association Statute (Estatuto da Ordem dos Médicos, DL 282/77) and the Medical Profession Statute (Estatuto do Médico, DL 373/79), which among other things: (a) Establishes that doctors should always act in the defense of the collective interests of the profession; (b) Explicitly forbids doctors to reduce fees in order to compete with other doctors (though doctors can provide medical services for free); (c) Makes clear that doctors are expected to follow the fees recommended by the Medical Association; (d) Forbids doctors from criticizing other doctors without prior consultation with the professional body.

Advertising is regulated in most jurisdictions, US and UK being less restrictive and Portugal being one of the most restrictive. With the exception of announcement of opening or closing practice, listing in the phonebook and the nameplate (and even this one is clearly regulated in dimension and content), advertising is banned. Competitive pressure and publicity in the internet have led the professional body to issue a new document on publicity, General Regulations Applying to Publicity of the Medical Profession (Regulamento Geral sobre Publicidade, June 2000), clarifying the strictness of the rules justified by the so-called principle of non-commercialization of medical services and alleged protection of consumers. In this document, the professional body urges the Government to apply and extend these prohibitions to managed healthcare organizations (which are not under the regulatory jurisdiction of the Portuguese Medical Association), and threatens doctors who cooperate with advertising of these organizations in violation of professional rules with appropriate disciplinary action. Advertising is allowed in Spain as long as it does not convey false information or bad publicity to the medical profession.

With respect to fees, Portugal alongside with Germany and the Netherlands has the least competitive market (in fact, competition is strictly forbidden by the professional body in Portugal). Recommended fees exist and are expected to be observed in Portugal. Fees are flexible in Spain in the private sector though the Spanish professional code points out that medical services should not aim at profits. Nevertheless, Spain as well as the UK have a powerful national health service that effectively restrains fee competition. The same does not happen in the US, where fees can be freely negotiated.

In Portugal, most doctors work for the national health service, but sole practitioners or small professional partnerships exist. Corporations and multidisciplinary organizations are not allowed. Most countries regulate the structure of doctor's firms, usually imposing limitations to incorporation (e.g., in most states of the US professional corporations can only provide services in one profession, or in Belgium unlimited personal liability applies), though Portugal seems more restrictive than average.

Liability for medical negligence not only is underdeveloped in Portugal, but it is also extremely complex. First, it can be contractual (breach of contract in the private sector) or extra-contractual liability (negligence for doctors in the national health system). Whereas for contractual liability, the patient has a period of twenty years to sue the physician after the wrongdoing (article 309 of the Civil Code), for extra-contractual liability, the same patient has three years from the moment s/he knows a wrongdoing took place (article 498 of the Civil Code). A similar liability dichotomy exists in the UK, but the development of expert witnessing and the different structure of the legal system has not produced the chilling effect that is observed in Portugal. Moreover, these liability rules clearly

undermine incentives for private medical services. However, this is not the only odd rule. Whereas for doctors in the private sector, law enforcement is exercised by regular courts, doctors in the national health service are under the jurisdiction of administrative courts. Given that many physicians work for the national health service but practice privately in part-time, conflicts and questions of court jurisdiction usually take place when patients want to sue doctors. Overall, the situation is confusing and difficult to understand even for legal scholars (Boletim da Ordem dos Advogados, December 2002). Not surprisingly, lawsuits for medical negligence are occasional and unlike to succeed in Portugal.

6. Comparative Institutional Analysis

Ranking the different institutional frameworks is a difficult task given the very distinct institutional details. Although modern techniques allow a more rigorous construction of indices, including a factor analysis approach, we provide a more simplified approach. Following the methodology proposed by Faure et. al. (1993), we construct a comparative institutional ranking of the regulations of professional services. The interpretations of the index should be very careful having in mind that it depends crucially on the questions surveyed (which do not cover all institutional details) and the relative importance we give to each set of questions (we try to correct somehow for this problem by presenting weighted averages).

We provide a summary of our own cross-national comparisons (Tables 7 and 8) as well as a detailed analysis of previous research by Faure et. al. (1993) and Paterson et. al. (2003) (Tables 9 and 10). Our index is based on a set of questions (Tables 5 and 6). They are a modified version of Faure et. al. (1993) where questions concerning professional schools, management of legal aid in the case of lawyers, and malpractice litigation have been included. We also eliminated some questions that in our view were duplications.

The process by which we construct a market failure approach index is the following: A country gets a point if the answer to the question complies with the market failure approach and zero otherwise. Complying with the market failure approach means that the answer to the question is consistent with improving market performance (as summarized in Table 1)

Table 5 (cont.)

Cross National Comparison with Respect to Lawyers

QUESTION NUMBER		UK	US	GER	BEL	NET	SPAIN	POR	POINTS YES	POINTS NO
	ENTRY REGULATIONS									
1	A law degree obtained from a recognized law school in the country is required for practice as an attorney?	YES	YES	YES	YES	YES	YES	YES	1/4	0
	Registration is required?	YES	YES	YES	YES	YES	YES	YES	1/4	0
	License is required?	YES	YES	YES	YES	NO	YES	YES	1/4	0
	Membership of professional body is required?	YES	YES	YES	YES	YES	YES	YES	0	1/4
2	Additional training is required? If Yes, how long?	YES (36)	NO	YES (30)	YES (36)	YES (36)	NO	YES (18)	0	1/2
	The additional training ends with an examination?	YES	--	YES	YES	YES	--	YES	0	1/2
3	The right (license) to practice as a lawyer is valid for the rest of one's active life?	YES	YES	YES	YES	YES	YES	YES	1	0
4	Are the law schools very competitive with respect to attracting students and faculty?	YES	YES	NO	YES	YES	NO	YES	1	0
5	Do only attorneys have the right to plead before courts in your country?	YES	NO	NO	YES	NO	YES	YES	0	1/4
	Does the right to plead depend on additional requirements?	YES	YES	NO	NO	NO	NO	NO	0	1/4
	Do attorneys have the right to plead before any court in your country?	NO	NO	NO	YES	YES	YES	YES	1/4	0
	Do only attorneys have the right to provide legal advice?	NO	YES	YES	NO	NO	NO	YES	0	1/4

Table 5 (cont.)

Cross National Comparison with Respect to Lawyers

QUESTION NUMBER		UK	US	GER	BEL	NET	SPAIN	POR	POINTS YES	POINTS NO
	ENTRY REGULATIONS									
6	Are attorneys established in any EU member country allowed to provide legal advice in your country?	YES	YES	YES	YES	YES	YES	YES	1/4	0
	Are attorneys established in any EU member country allowed to plead before your courts?	NO	YES	YES	YES	YES	YES	YES	1/4	0
	Are there any barriers to establishment?	YES	YES	YES	YES	YES	YES	YES	0	1/4
	Are these requirements discriminating against attorneys from other EU member countries?	YES	YES	YES	YES	YES	YES	YES	0	1/4
	STRUCTURE OF LAW FIRMS									
7	Can attorneys enter into partnerships?	YES	YES	YES	YES	YES	YES	YES	1	0
8	Can attorneys enter into multidisciplinary partnerships?	NO	NO	YES	NO	YES	NO	NO	1	0
9	Can attorneys incorporate?	YES	YES	NO	YES	YES	YES	NO	1/2	0
	With respect to incorporation, do any further restrictions apply?	YES	YES	--	YES	YES	YES	--	0	1/2
	PRICE/FEES									
10	Fees payable for legal service are freely negotiated?	YES	YES	NO	YES	YES	YES	YES	1/3	0
	The government sets fees (min, max, or fixed, or recommended)?	NO	NO	YES	NO	NO	NO	NO	0	1/3
	The self-regulatory organization of attorneys sets (min, max, or fixed, or recommended)?	NO	NO	NO	YES	YES	NO	NO	0	1/3

Table 5 (cont.)

Cross National Comparison with Respect to Lawyers

QUESTION NUMBER		UK	US	GER	BEL	NET	SPAIN	POR	POINTS YES	POINTS NO
	ENTRY REGULATIONS									
11	Fees can be based on hours worked?	YES	YES	NO	YES	YES	YES	YES	1/4	0
	Fees can be based on litigation value?	YES	YES	YES	NO	YES	YES	YES	1/4	0
	Fees can be based on the complexity of the case?	YES	YES	NO	YES	YES	YES	YES	1/4	0
	Can attorneys use contingent fees?	NO	YES	NO	NO	NO	NO	NO	1/4	0
12	Can attorneys who are more competent than others charge higher fees?	YES	YES	NO	YES	YES	YES	YES	1	0
13	Are legal aid boards run by the professional association of attorneys?	NO	NO	NO	YES	NO	NO	YES (IAD)	0	1
	ADVERTISING									
14	Advertising is allowed subject to the same constraints as any other services?	NO	NO	NO	NO	NO	NO	NO	0	1/3
	The state restricts the advertising of attorneys relative to other services?	NO	YES	YES	NO	NO	NO	NO	1/3	0
	The self-regulatory body restricts the advertising of attorneys?	YES	NO	YES	YES	YES	YES	YES	0	1/3
15	Advertising is very limited (e.g., phone book and the name plate)?	NO	NO	NO	YES	NO	NO	NO	0	1/6
	Special expertise can be advertised?	YES	YES	YES	NO	YES	YES	YES	1/6	0
	Fee level can be advertised?	YES	YES	NO	NO	YES	NO	YES	1/6	0
	Is comparative advertising possible?	NO	YES	NO	NO	NO	NO	NO	1/6	0
	Co-operation with other attorneys can be	YES	YES	YES	YES	YES	YES	YES	1/6	0

Table 5 (cont.)

Cross National Comparison with Respect to Lawyers

QUESTION NUMBER		UK	US	GER	BEL	NET	SPAIN	POR	POINTS YES	POINTS NO
	ENTRY REGULATIONS									
	advertised?									
	Co-operation with foreign attorneys or partners can be advertised?	YES	YES	YES	YES	YES	YES	YES	1/6	0
	QUALITY STANDARDS AND ENFORCEMENT									
16	The state defines the codes of conduct?	NO	YES	NO	NO	NO	NO	NO	1/2	0
	The self-regulatory body defines the codes of conduct?	YES	NO	YES	YES	YES	YES	YES	0	1/2
17	Is continuing education required as one of the items of the code of conduct?	YES	YES	YES	NO	NO	NO	NO	1	0
18	Are lawyers required to keep records on all details of a case?	YES	YES	YES	YES	NO	YES	YES	0	1/2
	Must the records be supplied to the disciplinary body, when a complaint is filed?	YES	YES	YES	YES	NO	YES	YES	0	1/2
19	Does the law require attorneys to give best advice?	NO	YES	YES	NO	YES	YES	YES	1/2	0
	Does the self-regulatory body require attorneys to give best advice?	NO	YES	YES	NO	YES	YES	YES	1/2	0
20	If the client of an attorney can prove that he did not obtain best advice and that as result he suffered a loss (e.g. lost a case), can the self-regulatory body punish the attorney?	YES	YES	YES	YES	YES	YES	YES	1/2	0

Table 5 (cont.)

Cross National Comparison with Respect to Lawyers

QUESTION NUMBER		UK	US	GER	BEL	NET	SPAIN	POR	POINTS YES	POINTS NO
	ENTRY REGULATIONS									
	Could the sanction be the expulsion from the professional association?	NO	YES	YES	YES	YES	YES	YES	1/2	0
21	Is liability for professional negligence usually applied by courts?	YES	YES	NO	NO	YES	NO	NO	1/2	0
	Is expert witnessing common in professional litigation?	YES	YES	NO	NO	YES	NO	NO	1/2	0

Source: Faure et. al (1993), Interview with the Chairman of the Portuguese Bar, José Miguel Júdice (Portugal), Estatuto General de la Abogacía Española (Spain).

Table 6 (cont.)

Cross National Comparison with Respect to Doctors

QUESTION NUMBER		UK	US	GER	BEL	NET	SPAIN	POR	POINTS YES	POINTS NO
	ENTRY REGULATIONS									
1	A medical degree from a recognized medical school in the country is required for practice as a doctor?	YES	YES	YES	YES	YES	YES	YES	1/4	0
	Registration is required?	YES	YES	YES	YES	NO	YES	YES	1/4	0
	License is required?	NO	YES	YES	YES	NO	YES	YES	1/4	0
	Membership of professional body is required?	NO	NO	YES	YES	NO	YES	YES	0	1/4
2	Additional training is required? If Yes, how long?	YES (60)	YES	YES (60)	NO	YES (12)	YES	YES (18)	0	1/2
	The additional training ends with an examination?	NO	YES	YES	NO	NO	YES	YES	0	1/2
3	The right (license) to practice as a doctor is valid for the rest of one's active life?	YES	YES	YES	YES	YES	YES	YES	1	0
4	Are the medical schools very competitive with respect to attracting students and faculty?	YES	YES	NO	YES	YES	NO	NO	0	1
5	Do most doctors work privately or for private insurance companies?	NO	YES	NO	NO	NO	YES	NO	1/2	0
	Do most doctors work for the national health system?	YES	NO	YES	YES	YES	NO	YES	0	1/2
6	Are doctors established in private practice in any EU member country allowed to provide medical services in your country?	YES	NO	NO	YES	YES	NO	NO	1/4	0

Table 6 (cont.)

Cross National Comparison with Respect to Doctors

QUESTION NUMBER		UK	US	GER	BEL	NET	SPAIN	POR	POINTS YES	POINTS NO
	ENTRY REGULATIONS									
	Does the doctor have to establish himself in your country before he can practice medicine?	NO	YES	YES	NO	NO	YES	YES	0	1/4
	Are there any barriers to establishment?	YES	YES	NO	NO	NO	YES	NO	0	1/4
	Are these requirements discriminating doctors from other EU member countries?	YES	YES	NO	NO	NO	YES	NO	0	1/4
	STRUCTURE OF DOCTOR'S FIRMS									
7	Can doctors enter into partnerships?	YES	YES	YES	YES	YES	YES	YES	1	0
8	Can doctors enter into multidisciplinary partnerships?	YES	YES	YES	YES	YES	YES	NO	1	0
9	Can doctors incorporate?	YES	YES	NO	YES	YES	YES	NO	1/2	0
	Do any further restrictions apply?	NO	YES	--	YES	NO	YES	--	0	1/2
10	Can doctors be employed by professional managers who are not doctors?	YES	YES	NO	NO	NO	NO	NO	1/2	0
	Only by recognized hospitals?	--	--	YES	YES	NO	YES	YES	1/2	0
	PRICE/FEES									
11	In private practice the dominant mode of payment for doctors is fee for service?	YES	YES	YES	YES	NO	YES	YES	1/2	0
	In private practice the dominant mode of payment for doctors is capitation fee?	NO	NO	NO	NO	YES	NO	NO	0	1/2
12	Is the fee freely negotiable between the doctor and the patient?	NO	YES	NO	NO	NO	NO	NO	1/5	0
	Is there a minimum or maximum fee (or fee	NO	NO	YES	YES	NO	NO	NO	0	1/5

Table 6 (cont.)

Cross National Comparison with Respect to Doctors

QUESTION NUMBER		UK	US	GER	BEL	NET	SPAIN	POR	POINTS YES	POINTS NO
	ENTRY REGULATIONS									
	schedule)?									
	Is there a fixed fee (or fee schedule) which is typically applied (say in more than 90% of the cases)?	YES	NO	YES	NO	YES	YES	YES	0	1/5
	The government sets fees (min, max, or fixed, or recommended)?	YES	NO	YES	NO	NO	YES	YES	0	1/5
	The doctors association or some other doctors' organization sets the fees (min, max, or fixed, or recommended)?	YES	NO	YES	YES	YES	NO	YES	0	1/5
13	Can doctors who are more competent than others charge higher fees in the private sector?	YES	YES	YES	YES	NO	YES	YES	1/2	0
	Can doctors who are more competent than others charge higher fees in the public sector?	NO	YES	NO	NO	NO	NO	NO	1/2	0
	ADVERTISING									
14	Advertising is allowed subject to the same constraints as any other services?	NO	NO	NO	NO	NO	NO	NO	0	1/3
	The state restricts the advertising of doctors relative to other services?	NO	YES	NO	YES	NO	NO	NO	1/3	0
	The self-regulatory body restricts the advertising of doctors?	YES	YES	YES	YES	YES	YES	YES	0	1/3
15	Only the academic title and the special expertise can be advertised?	NO	NO	YES	YES	YES	YES	YES	0	1/6

Table 6 (cont.)

Cross National Comparison with Respect to Doctors

QUESTION NUMBER		UK	US	GER	BEL	NET	SPAIN	POR	POINTS YES	POINTS NO
	ENTRY REGULATIONS									
	Advertisements in newspapers can be placed at any time?	YES	YES	NO	NO	NO	YES	NO	1/6	0
	Or only when a practice is opened?	NO	NO	YES	YES	YES	YES	YES	0	1/6
	Advertising is basically limited to the announcement of the opening and closing of a practice, the listing in the phone book and the nameplate?	NO	NO	YES	YES	YES	NO	YES	0	1/6
	Fee level can be advertised?	NO	YES	NO	NO	NO	NO	NO	1/6	0
	Co-operation with other doctors or specialists can be advertised?	NO	YES	YES	NO	NO	YES	NO	1/6	0
	QUALITY STANDARDS AND ENFORCEMENT									
16	The state defines the codes of conduct?	NO	YES	YES	NO	NO	NO	NO	1/2	0
	The self-regulatory body defines the codes of conduct?	YES	NO	YES	YES	YES	YES	YES	0	1/2
17	Is continuing education required as one of the items of the code of conduct?	NO	YES	YES	YES	NO	YES	YES	1	0
18	Are doctors required to keep records on all details of a case?	YES	YES	YES	NO	YES	YES	YES	0	1/2
	Must the records be supplied to the disciplinary body , when a complaint is filed?	YES	YES	YES	YES	YES	YES	YES	0	1/2
19	Does the law require doctors to give best advice	YES	YES	YES	NO	NO	YES	YES	1/2	0

Table 6 (cont.)

Cross National Comparison with Respect to Doctors

QUESTION NUMBER		UK	US	GER	BEL	NET	SPAIN	POR	POINTS YES	POINTS NO
	ENTRY REGULATIONS									
	and therapy?									
	Does the self-regulatory body require doctors to give best advice and therapy?	YES	YES	YES	YES	YES	YES	YES	1/2	0
20	If a patient can prove that he did not obtain best advice or therapy and that as result he suffered a loss (e.g. lost a case), can the self-regulatory body punish the doctor?	YES	YES	YES	YES	YES	YES	YES	1/2	0
	Could the sanction be the expulsion from the professional association?	YES	YES	NO	YES	YES	YES	YES	1/2	0
21	Is liability for professional negligence usually applied by courts?	YES	YES	YES	NO	NO	YES	NO	1/2	0
	Is expert witnessing common in professional litigation?	YES	YES	NO	NO	NO	NO	NO	1/2	0

Source: Faure et. al (1993), Estatuto do Médico, Regulamento sobre Publicidade and Código Deontológico (Portugal), Código de Ética y Deontología Medica (Spain).

Table 7

Comparative Institutional Analysis: Lawyers

	US	BEL	NET	SPAI	E&W	GER	POR	TOTAL
ENTRY	4.5	4.25	3.75	4	3.25	2.75	2.75	6
FEES	4	2.42	3.42	3.75	3.75	1.58	2.75	4
ORGAN.	1.5	2.5	2.5	2.5	1.5	2	1	3
ADVERT.	2	1.67	1.16	1	1.16	1.34	1.16	2
CONDUCT	5	4	4	2	2.5	3	2	6
TOTAL	17	14.84	14.83	13.25	12.16	10.67	9.66	21
SUM OVER TOTAL	80.9% (1)	70.7% (2)	70.6% (3)	63.1% (4)	57.9% (5)	50.8% (6)	46.0% (7)	100%
WEIGHTED AVERAGE	81.7% (1)	73.0% (2)	71.2% (3)	65.4% (4)	59.5% (5)	53.8% (6)	47.9% (7)	100%

Table 8

Comparative Institutional Analysis: Physicians

	US	E&W	BEL	SPAI	GER	NET	POR	TOTAL
ENTRY	4	3.75	4.75	2.75	2.25	4	2.25	6
FEES	3	1.7	1.9	1.9	1.5	0.4	1.7	3
ORGAN.	3.5	4	2.5	2.5	2	3.5	1	4
ADVERT.	1.67	1	0.67	1	0.5	0.33	0.33	2
CONDUCT	5	3	3	3.5	3.5	1.5	3	6
TOTAL	17.17	13.45	12.82	11.65	9.75	9.73	8.28	21
SUM OVER TOTAL	81.8% (1)	64.1% (2)	61.0% (3)	55.5% (4)	46.4% (5)	46.3% (6)	39.4% (7)	100%
WEIGHTED AVERAGE	84.2% (1)	63.8% (2)	57.7% (3)	56.0% (4)	44.2% (5)	41.8% (6)	37.1% (7)	100%

6.1. The Questionnaire

The questions are divided across the five dimensions we have considered in previous sections: entry, organization, price, advertising, and conduct regulations.

Most of the information used to answer this questionnaire (except questions four and twenty-one for doctors and questions four, thirteen, and twenty-one for lawyers) has been made available at Faure et. al. (1993) and Paterson et. al. (2003), the latter only for the legal profession. For the Portuguese and Spanish cases, the questionnaire was mailed to the Portuguese and Spanish law bars (Ordem dos Advogados and Consejo General de la Abogacia Española) and the Portuguese and Spanish medical associations (Ordem dos Médicos and Consejo General de Colegios Oficiales de Médicos). Only the Portuguese law bar replied and the information provided by them was used to compile the answers. For the other three cases, we have used information available by means of codes of professional conduct or other information available at their webpages.

The answers to questions four (differentiation of professional schools), twenty-one (use of professional malpractice), and thirteen for lawyers (legal aid) were based on our own understanding of professional education, professional litigation and the management of legal aid in the sample countries.

The construction of our index is based on the spirit of Table 1. Therefore, a certain number of points is assigned every time the answer to the question means no existence of regulation or promotion of free competition, unless that regulation is clearly consistent with the market failure approach. Some points might be controversial, so we look at them in more detail here:

Question One: Registration and licensing seem to be the most efficient way of regulating the market for legal and medical services (Ogus, 1994, page 221). Insurance is an expensive alternative (e.g., the Dutch medical profession).

Question Two: After obtaining a degree in Law or in Medicine, additional training and further examination controlled by the professional body seems unnecessary.

Question Three: Submitting the right to practice as a doctor or as a lawyer to periodical review would certainly indicate rent-seeking motivation (control of the profession) and can hardly be justified on efficiency grounds.

Question Four: Product differentiation is a signal of competition whereas product homogeneity imposed administratively by the Government or the professional body has no substantive efficiency justification.

Question Seventeen: Continuing education is expected to raise the quality of the professional service. Mandatory continuing education is efficient in the absence of market incentives (due to asymmetry of information).

Question Eighteen: Keeping records and mandatory disclosure of those records to the disciplinary body cannot be justified for efficiency reasons since professionals should be free to decide on what type of information they want to record and eventually disclose. These rules increase production cost (hence prices) with no obvious gain for customers, either legal clients or patients.

6.2. Relations between Indices

Our points do not match exactly the rankings offered at Faure et. al. (1993) for two reasons: (a) They offer three indices (libertarian, efficiency, and consumer protection) that in our view are less compelling, and (b) We average out questions within the survey by relevant item.

In Tables 9 and 10 we present the results for libertarian (Faure a)), efficiency (Faure b)), and consumer protection (Faure c)) as well as their rankings for a sample of five countries (UK, US, Netherlands, Belgium and Germany). The libertarian index measures the absence of restrictive rules, the optimal framework being free competition without any limits. One point is assigned whenever a regulation is not used in a country and zero is assigned whenever the regulation is enforced. The efficient index looks for regulations only for market failures commonly accepted in economics (therefore, this is the index closer in spirit to ours). Finally, the consumer protection index accepts regulations that a country adopts in order to minimize losses of welfare for consumers thought at the expense of freedom of competition.

Paterson et. al. (2003) also provide an index of regulation for different professions based on entry (IAS a)) and conduct (IAS b)) restrictions. They measure how much a given profession is regulated, hence producing a result somehow similar to the libertarian index provided by Faure et. al. (1993). The entry and conduct indices are aggregated in a composite index which we do not present since it is just the sum of the points obtained in each of the regulation indices.

In Table 9 we can see the ranking for the legal profession (the medical profession was excluded from their project though there was the intention of carrying on such study in the original proposal) for a sample of fifteen countries (all current members of the European Union). Their ranking does not match ours because we look at improving market performance given the existence of a market failure. Hence we look at quantitative issues (e.g., number of restrictions), but also at quality and nature of regulatory instruments and constraints.

As mentioned in the context of our own index, we should note that equal weight was given to the questions in both research projects (with the exception of multiple questions relating to similar issues) and therefore the issues covered with more questions carry more weight in the final ranking.

We have made use in our questionnaire of data available at Faure et. al (1993) and Paterson et. al. (2003). In the first project, the data was obtained by direct questionnaire to local experts in the five countries analyzed, in some cases the authors of the respective chapter in the book. In the second project, questionnaires were sent to the national law bars. In some very minor cases, there are inconsisencies between those two sets of information.

Table 9

Comparative Institutional Analysis: Lawyers

	US	**BEL**	**NETH**	**SPAIN**	**UK**	**GER**	**POR**	**SAMPLE**
Faure a)	2.60 (2)	2.33 (3)	3.08 (1)	-	2.30 (4)	1.50 (5)	-	5
Faure b)	3.00 (2)	2.53 (3)	3.28 (1)	-	2.30 (4)	1.90 (5)	-	5
Faure c)	3.10 (1)	2.07 (5)	2.55 (2)	-	2.30 (3)	2.22 (4)	-	5
IAS a)	-	2.50 (6)	2.10 (3)	3.40 (9)	2.90 (8)	3.70 (12)	3.50 (10)	15
IAS b)	-	2.10 (6)	1.80 (5)	3.10 (12)	1.20 (4)	2.80 (10)	2.20 (8)	15
Garoupa a)	8.09 (1)	7.07 (2)	7.06 (3)	6.31 (4)	5.95 (5)	5.38 (6)	4.60 (7)	7
Garoupa b)	8.17 (1)	7.30 (2)	7.12 (3)	6.54 (4)	5.75 (5)	5.08 (6)	4.79 (7)	7

Notes: In brackets, the ranking position.

IAS a) also includes Finland (1), Sweden (2), Denmark (3), Ireland (5), Italy (7), Greece (10), Luxemburg (13), France (14), and Austria (15).

IAS b) also includes Finland (1), Sweden (2), Denmark (3), Ireland (6), France (9), Luxemburg (10), Austria (13), Italy (14), and Greece (15).

Table 10

Comparative Institutional Analysis: Physicians

	US	UK	BEL	SPAIN	GER	NETH	POR	SAMPLE
Faure a)	2.00 (1)	2.00 (1)	1.63 (4)	-	1.10 (5)	1.80 (3)	-	5
Faure b)	2.40 (1)	2.40 (1)	2.03 (3)	-	1.30 (5)	2.00 (4)	-	5
Faure c)	2.90 (1)	2.30 (2)	1.78 (3)	-	1.21 (5)	1.40 (4)	-	5
Garoupa a)	8.18 (1)	6.41 (2)	6.10 (3)	5.55 (4)	4.69 (5)	4.63 (6)	3.94 (7)	7
Garoupa b)	8.42 (1)	6.38 (3)	5.77 (2)	5.60 (4)	4.42 (5)	4.18 (6)	3.71 (7)	7

Note: In brackets, the ranking position.

6.3. Our Findings with Respect to Lawyers

Tables 7 and 8 present the results for each regulatory instrument. We also add a weighted average where the same weight is given to each regulatory instrument to overcome the problem that the number of questions varies for different regulatory instruments.

With respect to lawyers, we can immediately see that the US regulatory framework seems closer to improving market performance for legal services than most European jurisdictions essentially due to the fact that the US is not so much regulated and is more competitive. Within the EU we identify three groups: the Netherlands and Belgium that seem to have a regulatory framework producing a result similar to the US (a result consistent with Faure et. al., 1993), a second group of jurisdictions (UK and Spain) with a performance below the US but clearly above the performance of the third group (Portugal and Germany). Germany's results are somewhat justified by excessive regulation of fees and advertising.

With respect to other available indices, we obtain fairly similar results for most countries. Spain performs better in our ranking than in Paterson et. al. (2003) because we value more the less restrictive entry rules applied in this country.

Looking at the case of Portugal, we identify two main sources of problems with respect to correcting for market failures without running into capture. The more serious problem is of course restrictions on organizational forms that are clearly more restrictive in Portugal than average. The recent developments have been pushed by EU competition, but they are still insufficient by not allowing law firms to become commercial societies. Though data is not available, we do have the feeling that the Portuguese legal market has been able to resist for longer to penetration by UK and American law firms than Spain or Belgium, and multidisciplinary partnerships (lawyers, business consultants and accountants) are actively opposed by the legal profession. For example, various London-based-law firms appeared in the Dutch market after 1999, and forced the Dutch bar to become more business-minded with regard to fees, advertising and professional liability. In Spain, big auditing firms have been associated with well-known partnerships to create large modern and US-style law firms.

The second source of problems concerns restrictions on conduct, namely the inexistence of effective professional litigation and the fact that the codes of conduct rely on the will of the professional body, and not on the Government. Negligence for professional conduct is still not sufficiently developed in Portugal (Boletim da Ordem dos Advogados, December 2002).

With respect to entry restrictions and advertising rules, they are similar to most of other European jurisdictions. The pressure introduced by the sudden increase of law schools in the late 80s has probably done more to relax some of the constraints than actual competition from outside.

The Portuguese market for legal services is quite competitive in terms of fees as most of the European markets (with the exception of Germany), however legal aid being run by the professional body (an institutional design also developed by Belgium) is hardly consistent with improving market performance.

6.4. Our Findings with Respect to Physicians

With respect to physicians, we can see that the US regulatory framework again seems closer to improving market performance for medical services than most European jurisdictions. Within the EU we identify two groups: the first group (UK, Belgium, and Spain slightly below) with a performance below the US but clearly above the performance of the second group (Portugal, the Netherlands, and Germany). Most of these results are consistent with Faure et. al. (1993), with the exception of the Netherlands. This is essentially due to the fact that we mark as negative the Dutch system having no registration (since the market failure approach relies on some degree of consumer protection).

Looking at the case of Portugal, the reasons for performing so badly are very different from the Netherlands, but somehow similar to Germany, the difference being that German regulations are not so restrictive. We find that restrictions on fees, advertising, organizational forms, and conduct are too severe, plus the lack of access to medical schools makes entry very restrictive.

Portugal has the most severe restrictions on advertising and organizational forms of our sample of countries. Though most medical services are provided by the national health service (hence the problem of fees chosen by the Government is similar in Portugal as well as in the UK or Germany), the Portuguese Medical Association is the only one that forbids competition and makes clear that recommended medical fees are to be observed (hence they are in fact mandatory). In fact, just by comparing the Portuguese and Spanish medical professional codes one can immediately detect not only that restrictions are much more severe in Portugal, but also competition between physicians is to be avoided at all costs. Whereas the Spanish professional code emphasizes what doctors should and can do, the Portuguese professional code is overwhelmingly about what doctors cannot and should not do. These differences are somehow reflected in Table 8.

Professional liability is much weaker in Portugal than in any other country of our sample, thus deterring lawsuits and eliminating any possibility for regulation by private parties. The absence of effective medical expert witnessing (chilled out by professional regulations that forbid doctors from criticizing other doctors without the consent of the professional body) helps the dilution of liability for malpractice.

7. Conclusions

In this paper we have presented a systematized summary of the economic literature on regulation of professionals, with a special application to legal and medical services. A case-study of Portuguese medical and legal professional bodies has been developed. An index of quality of the regulatory set-up has been constructed where aspects related to entry, fees, organizational forms, advertising, and conduct restrictions are included. A country getting a higher number of points is interpreted to have a professional regulatory framework more consistent with improving market performance (given the existence of a market failure).

Portugal does not perform well in our study, but evidence suggests that for legal services it is not too far away from the EU average whereas for medical services it is clearly below average. Hence our policy recommendations are quite different with respect to both professions. Nevertheless, a word of cautions is necessary here. The enactment of

some of these reforms should be the subject of a more detailed and focus-oriented cost-benefit analysis.

For regulation of legal services, we have identified some target areas that would need reform: (a) More flexible rules with respect to organizational forms (including the possibility of incorporation); (b) Further development of professional liability for malpractice; (c) Promotion of market penetration by UK-based and US-style law firms; (d) Adoption of legal aid institutions more in the line of legal aid boards (UK-style or Dutch-style) rather than the Belgium system; (e) Monitoring of professional rules and enforcement by the Competition Authority (in compliance with ECJ jurisprudence) treating the Portuguese law bar (Ordem dos Advogados) as one of the many industry-specific regulators.

For regulation of medical services, we recommend a serious and more profound reform of the regulatory framework: (a) Reform of the professional code along the lines for example of the Spanish professional code; (b) More flexible rules concerning advertising and organizational forms; (c) Abolishment of recommend fees by the professional body; (d) Education and training should be less controlled by the professional body; (e) Promotion of competition between medical schools with the aim of increasing significantly the number of doctors entering the workforce; (f) Development of effective professional liability for medical malpractice and independent medical expert witnessing; (g) Monitoring of professional rules and enforcement by the Competition Authority (in compliance with ECJ jurisprudence) treating the Portuguese medical association (Ordem dos Médicos) as one of the many industry-specific regulators.

Not surprisingly, the professional bodies oppose openly to most of these reforms. They argue that these reforms will harm the public interest by downgrading the quality of the service they provide. However, most countries have in some ways relaxed these restrictions, and the quality of medical and legal services can hardly be characterized as being inferior to Portugal.

References

Agrawal, G. B. and Hall, M. A., 2003, What if You Could Sue Your HMO? Managed Care Liability Beyond the ERISA Shield, St. Louis University Law Journal 47.

Amorim, C. and Kipping, M., 1999, Selling Consultancy Services: The Portuguese Case in Historical and Comparative Perspective, Business and Economic History 28, 45-56.

Andrews, P., 2002, Self-Regulation by Professions - The Approach Under EU and US Competition Rules, European Competition Law Review 23.

Arlen, J. and MacLeod, B., 2003, Torts, Expertise and Authority: Liability of Physicians and Managed Care Organizations, New York University Law Review, forthcoming.

Benham, L. and Benham, A., 1975, Regulating Through the Professions: A Perspective on Information Control, Journal of Law and Economics 18, 421-447.

Bishop, W., 1989, Regulating the Market for Legal Services in England: Enforced Separation of Function and Restrictions on Forms of Enterprise, Modern Law Review 52, 326-351.

Bowles, R., 1994, The Structure of the Legal Profession in England and Wales, Oxford Review of Economic Policy 10, 18-33.

Carr, J. and Mathewson, G. F., 1988, Unlimited Liability as a Barrier to Entry, Journal of Political Economy 96, 766-784.

Carr, J. and Mathewson, G. F., 1990, The Economics of Law Firms: A Study in the Legal Organization of Firms, Journal of Law and Economics 33, 307-330.

Cave, M., 1985, Market Models and Consumer Protection, Journal of Consumer Policy, 335-351.

Curran, C., 1993, The American Experience with Self-Regulation in the Medical and Legal Professions, in Regulation of Professions, Antwerpen, Maklu.

Danzon, P. M., 1985, Liability and Liability Insurance for Medical Malpractice, Journal of Health Economics 4, 309-331.

Danzon, P. M., 1991, Liability for Medical Malpractice, Journal of Economic Perspectives 5, 51-69.

Danzon, P. M., 1997, Tort Liability: A Minefield for Managed Care?, Journal of Legal Studies 27, 491-519.

Darby, M. R. and Karni, E., 1973, Free Competition and the Optimal Amount of Fraud, Journal of Law and Economics 16, 111-126.

Dauer, E. A. and Marcus, L. J., 1997, Adapting Mediation to Link Resolution of Medical Malpractice Dispute with Health Care Quality Improvement, Law and Contemporary Problems 60, 185-218.

Dingwall, R. and Fenn, P. T., 1987, A Respectable Profession? Sociological and Economic Perspectives on the Regulation of Professional Services, International Review of Law and Economics 7, 51-64.

Donabedian, B., 1995, Self-Regulation and the Enforcement of Professional Codes, Public Choice 85, 107-118.

Emons, W., 1997, Credence Goods and Fraudulent Experts, RAND Journal of Economics 28, 107-119.

Faure, M., 1993, Regulation of Attorneys in Belgium, in Regulation of Professions, Antwerpen, Maklu.

Faure, M., Finsinger, J., Siegers, J., and Van den Bergh, R., 1993, Regulation of Professions, Antwerpen, Maklu.

Faure, M. and Koziol, H., 2001, Cases on Medical Malpractice in a Comparative Perspective, Tort and Insurance Law, Vienna, Springer-Verlag.

Fine, D. K., 2003, Physician Liability and Managed Care: A Philosophical Perspective, Georgia State University Law Review 19.

Gehrig, T. and Jost, P., 1995, Quacks, Lemons, and Self Regulation: A Welfare Analysis, Journal of Regulatory Economics 7, 309-325.

Gravelle, H., 1990, Medical Negligence: Evaluating Alternative Regimes, Geneva Papers on Risk and Insurance 15, 22-26.

Guedes da Costa, O., 2003, Direito Profissional do Advogado, Lisbon, Almedina Editora.

Hadfield, G., 2000, The Price of Law: How the Market for Lawyers Distorts the Justice System, Michigan Law Review 98, 953-1006.

Hägg, G. T., 1997, Theories on the Economics of Regulation: A Survey of the Literature from a European Perspective, European Journal of Law and Economics 4, 337-370.

Hau, H. and Thum, M., 2000, Lawyers, Legislation and Social Welfare, European Journal of Law and Economics 9, 231-254.

Hellingman, K., 1993, An Economic Analysis of the Regulation of Lawyers in the Netherlands, in Regulation of Professions, Antwerpen, Maklu.

Kaplow, L. and Shavell, S., 2002, Fairness versus Welfare, Cambridge, MA, Harvard University Press.

Kessler, D., and McClellan, M., 1996, Do Doctors Practice Defensive Medicine?, Quarterly Journal of Economics 111, 353-390.

Kessler, D., and McClellan, M., 1997, The Effects of Malpractice Pressure and Liability Reforms on Physicians' Perceptions of Medical Care, Law and Contemporary Problems 60, 81-106.

Kessler, D. and McClellan, M., 2002a, Malpractice Pressure, Managed Care, and Physician Behavior, in Regulation through Litigation, Washington, DC, AEI-Brookings Joint Center for Regulatory Studies.

Kessler, D. and McClellan, M., 2002b, How Liability Law Affects Medical Productivity, Journal of Health Economics 21, 931-955.

Kessler, D. and McClellan, M., 2002c, Malpractice Law and Health Care Reform: Optimal Liability Policy in an Era of Managed Care, Journal of Public Economics.

Klein, B. and Leffler, K. B., 1981, The Role of Market Forces in Assuring Contractual Performance, Journal of Political Economy 89, 615-641.

Kleiner, M. and Kudrle, R. T., 2000, Does Regulation Affect Economic Outcomes? The Case of Dentistry, Journal of Law and Economics 43, 547-582.

Leffler, K. B., 1978, Physician Licensure: Competition and Monopoly in American Medicine, Journal of Law and Economics 21, 165-186.

Leland, H. E., 1979, Quack, Lemons, and Licensing: A Theory of Minimum Quality Standards, Journal of Political Economy 87, 1325-1346.

Love, J. and Stephen, F., 1996, Advertising, Price and Quality in Self-Regulating Professions: A Survey, International Journal of the Economics of Business 3, 227-247.

Lueck, D., Olsen, R. and Ransom, M., 1995, Market and Regulatory Forces in the Pricing of Legal Services, Journal of Regulatory Economics 7, 63-83.

Maks, J. A. H. and Philipsen, N. J., 2002, An Economic Analysis of the Regulation of Professions, in The Regulation of Architects, Antwerpen, Intersentia.
Matthews, R., 1991, The Economics of Professional Ethics: Should the Professions be more like Businesses?, Economic Journal 101, 737-750.
McChesney, F., 1987, Rent Extraction and Rent creation in the Economic Theory of Regulation, Journal of Legal Studies 26, 101-118.
Miller, F. H., 1997, Medical Discipline in the Twenty-First Century: Are Purchasers the Answer?, Law and Contemporary Problems 60, 31-58.
Miller, J., 1985, The FTC and Voluntary Standards: Maximizing the Net Benefits of Self-Regulation, The Cato Journal 4, 897-903.
Miller, G. and Macey, J. R., 1995, Reflections on Professional Responsibility in a Regulatory State, George Washington Law Review 63.
Murphy, K. M., Shleifer, A., and Vishny, R., 1991, The Allocation of Talent: Implications for Growth, Quarterly Journal of Economics 106, 503-530.
Noll, R., 1989, Economic Perspectives on the Politics of Regulation, in Handbook of Industrial Organization II, Amsterdam, North-Holland, 1253-1287.
Ogus, A., 1993, Regulation of the Legal Profession in England and Wales, in Regulation of Professions, Antwerpen, Maklu.
Ogus, A., 1994, Regulation: Legal Form and Economic Theory, Oxford, Oxford University Press.
Ogus, A., 1995, Rethinking Self-regulation, Oxford Journal of Legal Studies 15, 97-108.
Olsen, R. N., 1997, The Efficiency of Medical Malpractice Law: A New Appraisal, Research in Law and Economics 19.
Olsen, R. N., 1999, The Regulation of Medical Professions, in Encyclopedia of Law and Economics, Ghent, University of Ghent, 1018-1053.
Paterson, I., Fink, M., and Ogus, A., 2003, Economic Impact of Regulation in the Field of Liberal Professions in Different Member States, Study for the European Commission, DG Competition, Vienna, Institute for Advanced Studies.
Peltzman, S., 1976, Toward a More General Theory of Regulation, Journal of Law and Economics 19, 211-244.
Pashigian, B. P., 1979, Occupational Licensing and the Interstate Mobility of Professionals, Journal of Law and Economics 22, 1-25.
Posner, R. A., 1974, Theories of Economic Regulation, Bell Journal of Economics and Management Science 5, 335-358.
Posner, R. A., 1975, The Social Costs of Monopoly and Regulation, Journal of Political Economy 83, 807-827.
Rizzo, J. A. and Zeckhauser, R. J., 1992, Advertising and the Price, Quantity and Quality of Primary Physician Services, Journal of Human Resources 28, 381-421.
Rubin, P. H. and Bailey, M. J., 1994, The Role of Lawyers in Changing the Law, Journal of Legal Studies 23, 807-831.
Shaked, A. and Sutton, J., 1981, The Self-Regulating Profession, Review of Economic Studies 47, 217-234.
Shaked, A. and Sutton, J., 1982, Imperfect Information, Perceived Quality, and the Formation of Professional Groups, Journal of Economic Theory 27, 170-181.
Sheperd, G. B., 2000, Cartels and Controls in Legal Training, Antitrust Bulletin 45, 437-466.

Shinnick, E. and Stephen, F., 2000, Professional Cartels and Scale Fees: Chiselling on the Celtic Fringe?, International Review of Law and Economics 20, 407-423.

Sloan, F. A. and Hall, M. A., 2002, Market Failures and the Evolution of State Regulation of Managed Care, Law and Contemporary Problems 65, 169-208.

Stephen, F., 2002, The European Single Market and the Regulation of the Legal Profession: An Economic Analysis, Managerial and Decision Economics 23, 115-125.

Stephen, F., 2003, An Economic Perspective on the Regulation of Legal Service Markets, Evidence submitted to the Justice 1 Committee's Inquiry into the Regulation of the Legal Profession.

Stephen, F. and Love, J., 1996, Deregulation of Legal Services in the UK: Evidence from Conveyancing, Hume Papers on Public Policy 4, 53-66.

Stephen, F. and Love, J., 1999, Regulation of the Legal Profession, in Encyclopedia of Law and Economics, Ghent, University of Ghent, 987-1017.

Stephen, F., Love, J., and Peterson, A., 1994, Deregulation of Conveyancing Markets in England and Wales, Fiscal Studies 15, 102-118.

Stigler, G., 1971, The Theory of Economic Regulation, Bell Journal of Economics and Management Science 2, 3-21.

Van den Bergh, R., 1993, Self-Regulation in the Medical and Legal Professions and the European Internal Market in Progress, in Regulation of Professions, Antwerpen, Maklu.

Van den Bergh, R., 1999, Self-Regulation of the Medical and Legal Professions: Remaining Barriers to Competition and EC Law, in Organized Interests and Self-Regulation: An Economic Approach, Oxford, Oxford University Press.

Van den Bergh, R. and Faure, M., 1991, Self-Regulation of the Professions in Belgium, International Review of Law and Economics 11, 165-182.

Viscusi, W. K., 2002, Overview, in Regulation through Litigation, Washington, DC, AEI-Brookings Joint Center for Regulatory Studies.

Zerbe, R. and Urban, N., 1988, Including Public Interest in Theories of Regulation, Research in Law and Economics 8, 4-5.