

**cef.up working paper
2015-02**

**DEVIATING FROM THE BENCHMARKS: HUMAN
CAPITAL INPUTS AND THE SURVIVAL OF NEW
STARTUPS**

**Vera Rocha
Mirjam van Praag
Anabela Carneiro**

Deviating from the benchmarks: Human capital inputs and the survival of new startups*

Vera Rocha[†]

Copenhagen Business School-INO

Mirjam van Praag[‡]

Copenhagen Business School-INO

Anabela Carneiro[§]

Universidade do Porto and CEF.UP

DRAFT VERSION, PLEASE DO NOT QUOTE

May 29, 2015

Abstract

This paper studies three related questions: To what extent otherwise similar startups employ different quantities and qualities of human capital at the moment of entry? How persistent are initial human capital choices over time? And how does deviating from human capital benchmarks influence firm survival? The analysis is based on a matched employer-employee dataset and covers about 17,500 startups in manufacturing and services. We adopt a new procedure to estimate individual benchmarks for the quantity and quality of initial human resources, acknowledging correlations between hiring decisions, founders' human capital, and the ownership structure of startups (solo entrepreneurs versus entrepreneurial teams). We then study the survival implications of exogenous deviations from these benchmarks, based on spline models for survival data. Our results indicate that (especially negative) deviations from the benchmark can be substantial, are persistent over time, and hinder the survival of firms. The implications may, however, vary according to the sector and the ownership structure at entry. Given the stickiness of initial choices, wrong human capital decisions at entry turn out to be a close to irreversible matter with significant survival penalties.

Keywords: human resources, human capital, startup conditions, new ventures, firm survival, entrepreneurs, intra-industry dynamics

*The authors acknowledge GEP-MSESS (Gabinete de Estratégia e Planeamento – Ministério da Solidariedade, do Emprego e da Segurança Social) for allowing the use of Quadros de Pessoal dataset.

[†]vr.ino@cbs.dk

[‡]mvp.ino@cbs.dk

[§]anacar@fep.up.pt

1 Introduction

This paper addresses three related questions: To what extent do otherwise similar startups employ different quantities and qualities of human capital at the moment of entry? How likely are deviating initial human capital choices adapted at a later stage? And how does deviating from the *benchmark* influence firm survival? The main objective of our study is to measure the causal impact of entering with different levels of quality and quantity of human resources (HR) compared to a benchmark of identical firms, with identical owners. To this end, we have to obtain reliable estimates of individual benchmark levels of the quality and quantity of personnel at startup, which are determined by observed and unobserved individual, firm, and industry characteristics of each startup. Deviations from these benchmarks are considered exogenous and may relate to forecasting errors or random factors, allowing us to estimate the effect of interest.¹

The quantity and quality of HR at entry, namely startup size and (the owner’s and workers’) human capital, has been widely shown to matter for the performance of new ventures (Bates, 1990; Feeder and Willard, 1990; Cooper et al, 1994; Agarwal and Audretsch, 2001; Geroski et al., 2010). There is, however, a great heterogeneity in the initial workforce employed by new firms, even within narrowly defined industries, regions, and startup years.²

Some firms enter with a larger and/or more skilled workforce than close competitors, possibly due to entrepreneurs’ relatively higher skills, better financial conditions, overconfidence, or risk-taking preferences (Camerer and Lovallo, 1999; Koellinger et al., 2007; Katila et al., 2012; Åstebro et al., 2014). In contrast, many entrepreneurs are not able to match the competitors’ HR choices, either due to financial and/or ability constraints, forecasting errors regarding the expected market demand and competition, or risk aversion (e.g., Holtz-Eakin et al., 1994; Cressy, 1996; Cabral and Mata, 2003; Hyytinen et al., 2014), just to name a few potential reasons. Alternative explanations rely on the seminal models of learning by doing of Jovanovic (1982) and Pakes and Ericson (1998). Many firms may actually decide to enter at sub-optimal positions and to experience a (hopefully temporary) comparative disadvantage, in order to have the opportunity to learn about themselves and expand later, if successful (Audretsch and Acs, 1990; Audretsch et al., 1999; Santerelli and Vivarelli, 2007).

The long-term associations between startup conditions and the survival of new ventures (e.g., Geroski et al., 2010), and the stickiness of most initial decisions – including those related to HR (Hamermash and Pfann, 1996; Gilbert, 2005; Agarwal et al, 2009) – suggest,

¹Random variations in capital constraints, such as winning the lottery, may also be included in such random factors.

²See, for instance, Mata and Machado (1996), Cabral and Mata (2003), and Agarwal and Audretsch (2001) for evidence on startup size heterogeneity, and Haltiwanger et al. (1999, 2007) for evidence on the heterogeneity in workforce composition, even within narrow groups of firms.

however, that sub-optimal positions may be really difficult to correct later on. If that is the case, learning attempts by firms deliberately entering with inferior positions may be costly.

Though we find in the literature valuable discussions about how important is for new firms to match the so-called minimum efficient scale of the industry (e.g., Audretsh and Acs, 1990; Audretsch et al., 1999, 2000), the available evidence is restricted to startup size decisions in the manufacturing industry. A comparison between manufacturing and services firms may provide relevant insights about the importance of deviating from human capital benchmarks in industries with different entry barriers and capital requirements. We still lack comprehensive analyses of both human capital *quantity* and *quality* decisions at entry, comparing startups in both manufacturing and services.

Few studies have attempted to measure a causal impact of these decisions at entry on survival, which is obviously no “sine cure”. Most studies have disregarded the heterogeneity of the human capital of the entrepreneurs themselves founding these new ventures. They may be heterogeneous in their education levels, experience, and (unobserved) skills, which may of course influence their HR decisions at entry. Moreover not all entrepreneurs run their business alone. Some of them enter in teams, which may also have implications on human capital choices.

In view of that, using the human capital choices of competitors in the same industry as a benchmark may be limiting, since it neglects the heterogeneity caused by the size of the founding team (solo or team) and the human capital of the entrepreneurs themselves. Because we want to mimic, as much as possible, random variations in startup size and workforce quality to obtain unbiased impact estimates on survival, we use a new approach to estimate the human capital benchmarks for each new venture under analysis. We estimate benchmarks for both the number of employees at startup and the employees’ average skill level at entry. We then analyze how startups deviate from these individual benchmarks, and how that matters for their survival prospects. Thus, our approach takes into account that HR choices at entry are endogenous and possibly co-determined by observed and unobserved characteristics of the entrepreneur.

The study is based on matched employer-employee data for Portugal, and covers 17,579 startups entering the private sector during the period 1992-2007. All ventures are started by people who had been employees just before becoming business owners (BOs), and each venture employs at least one employee at entry. In this way, we try to study opportunity entrepreneurship of BOs for whom decisions about HR management are relevant. The analysis furthermore compares different startups according to the sector they enter (manufacturing versus services) and their initial ownership structure (solo entrepreneurs versus entrepreneurial teams).

The paper is structured as follows. Next section briefly frames the paper into the litera-

ture. Section 3 starts by presenting the data and the sample. It then explains the methods used to measure individuals' skills and to estimate human capital benchmarks. Section 4 provides some descriptive statistics on startup conditions and deviations from the estimated benchmarks by groups of firms. It also demonstrates how persistent initial decisions about the size and quality of the workforce are. Section 5 discusses the empirical results for the relationship between deviations from human capital benchmarks and firm hazard. Section 6 concludes.

2 Human capital choices at entry and the dynamics of new ventures

Relative positions are often more of a concern among individuals than absolute positions (Solnick and Hemenway, 1998; Alpizar et al., 2005). The same may be true among firms. The relative position of startup firms in terms of their initial human capital quantity and quality may encompass strategic decisions and/or signal entrepreneurs' constraints, risk preferences, or biased expectations.

Startup size was early understood as a key strategic decision for entrepreneurs (Birley and Westhead, 1994; Audretsch et al., 1999). Part of this decision could simply be understood as a choice between a higher risk/reward larger scale startup versus a lower risk/reward smaller scale startup. Entry size choices may also be explained by founders' characteristics, namely their ability or entrepreneurial talent and, partly as a consequence, cash constraints (Cabral and Mata, 2003; Parker and van Praag, 2006). More skilled entrepreneurs may suffer fewer financial restrictions and may have higher levels of self-confidence. They may be, therefore, more prone to take risks, more able to run a larger venture, and more likely to be able to pay the costs necessary to start at a larger scale (Mata and Machado, 1996; Cabral and Mata, 2003; Colombo et al., 2004). Entrepreneurs may, independent from their ability levels, also be overconfident in their own abilities and future chances of success. This may also lead to larger startups, all else equal (Bolger et al., 2008; Katila et al., 2012; Åstebro et al., 2014; Hyytinen et al., 2014). Other behavioral biases, such as overoptimism (Camerer and Lovallo, 1999; Hmieleski and Baron, 2009), or lower levels of loss aversion (Koudstaal et al., 2015) might have a similar impact on startup size.

On the contrary, those who do not reach the size benchmarks are more likely to correspond to financially constrained (Cabral and Mata, 2003) and/or less talented entrepreneurs (Lucas, 1978), risk averse entrepreneurs (e.g., Hvide and Panos, 2014), or people who have a strong preference for autonomy (Benz and Frey, 2008). Another explanation for startups entering at a sub-optimal scale is the entrepreneurs' expectations of learning by doing (Jo-

vanovic, 1982; Pakes and Ericson, 1998). Many entrepreneurs are actually uncertain about their ability and efficiency, so they may decide to enter at a small scale, relying on the expectation that they will be able to correct their entry decision later on and grow, as they update their beliefs about their ability and efficiency (Audretsch and Acs, 1990; Audretsch et al., 1999, 2000).

All these arguments can be used to explain the heterogeneity found in workforce size even among startups in the same industry, region, or cohorts. The same arguments will apply when discussing heterogeneity in the composition of the workforce in terms of the quality of employees (e.g., their education, experience, and skills) (Haltiwanger et al., 1999, 2007). In addition, for quality, there is also evidence of a positive hierarchical sorting by ability in firms, which means that there might be complementarities between workers' skills and those of their superiors (Garicano and Hubbard, 2005; Garicano and Rossi-Hansberg, 2006) – who, in micro and small startup firms, often correspond to business founders. More recent evidence based on matched employer-employee data for Portugal also confirms that more skilled entrepreneurs are more likely to attract more skilled workers and to pay higher wages on average (Baptista et al., 2013). Thus, the more skilled entrepreneurs may be more likely to match – or even surpass – the average human capital quality hired by closer competitors, while the less skilled (who may also correspond to the most financially constrained) may have a harder time matching human capital quality benchmarks.

We have argued (and will test) that entrepreneurs may make a variety of decisions about the quantity and quality of personnel hired at entry. Our second argument (and test) will be that initial positions may be sticky and hard to correct over time, for instance due to firm structural inertia (Gilbert, 2005), labor market rigidities, and consequent adjustment costs (Hamermesh and Pfann, 1996). Therefore, deviating from the benchmark quantity and quality of human capital at the moment of entry may be rather persistent. Even if some adjustments in initial hiring decisions can be made later on, startup conditions may become imprinted in the firm and have even stronger effects on survival than current conditions (e.g., Geroski et al. 2010). The stickiness of initial positions might enlarge the consequences of deviations from the benchmark in terms of later outcomes, such as firm survival. Our third argument (and test) concerns the potential survival consequences of initial personnel decisions in startups.

The non-randomness of human capital decisions hinders the unbiased measurement of the effect of (deviating) initial HR choices on firm survival. In view of this, and to measure this effect, we pull all these observed and unobserved heterogeneities across firms and entrepreneurs and their combinations together in a benchmark that is defined for each individual entrepreneur. The remaining deviation from this benchmark can be considered random, allowing us to measure the unbiased effect of HR decisions on survival. We must, however, be

aware that benchmark human capital levels are not necessarily the *optimal* levels, but rather the expected levels according to all these aforementioned heterogeneities and correlations.³

The relationship between human capital deviations and firm hazard may be quite complex and potentially different for particular groups of startups. For instance, economies of scale are undoubtedly more important in manufacturing industries than in services, so the (in)ability to (reach) exceed the startup size benchmark may have different implications for firm survival according to the sector. Deviations from the benchmark quality of the workers hired at entry may also have different consequences according to the degree of complementarity between firms' capital intensity and human capital quality. Positive and negative deviations may have distinct consequences. Deviations from benchmark size and skill levels at the moment of entry may have heterogeneous effects on new venture survival according to several (yet unknown) factors. While answering our three questions we distinguish between solo entrepreneurs and entrepreneurial teams operating in manufacturing and services.

3 Data and Methods

3.1 Data and sample

Our data come from *Quadros de Pessoal* (hereafter, QP), a large longitudinal linked employer-employee dataset collected by the Portuguese Ministry of Employment. QP covers all firms operating in the Portuguese private sector and employing at least one wage earner. Available information at the firm-level includes employment, sales, industry, ownership, location, among others. At the individual-level, QP reports information about each worker's age, education, gender, qualifications, wages, occupational category, tenure, number of hours worked, and type of contract. All firms, establishments and workers are identified with a unique identification number, so they can be followed and matched over time. Raw QP files are available for the period 1986-2009.⁴

Entries of new firms are identified by the first year a firm is recorded in QP files. Our analysis is based on startup firms entering in t , either in manufacturing industries or in

³Though the literature often refers to sub-optimal positions (and namely sub-optimal scales), it is hard (if not impossible) to measure the *optimal* quantity and quality of human capital. It varies according to the industry, the macroeconomic environment, the characteristics of firms and entrepreneurs active in the market, among other factors. However, by studying the consequences of deviating from the benchmark, we may get some insights about how close might be these benchmarks to the so-called optimal levels. If firms minimize their hazard by entering with the benchmark quantity and quality of human capital, then the benchmarks may be close to the optimal choices. If, instead, firms can get some survival bonus by deviating above (below) the benchmark, it may suggest that benchmark human capital choices are sub-optimal (over-optimal).

⁴There is a gap for the particular years of 1990 and 2001 in the worker-level files, for which no information was gathered at the individual-level.

services, whose founder(s) was/were in paid employment in $t - 1$ or $t - 2$. We exclude those startup firms founded by individuals who left their job in $t - 1$ or $t - 2$ in a firm that either closed down or suffered a massive downsizing in the same year, which would be closer to necessity-driven startup firms. Initial HR choices may be more constrained for this group of individuals. For the same reason, we also exclude startup firms whose founder(s) was (were) never observed in the QP files before, or who were absent for long periods of time (i.e., three or more years), as in these cases we are not able to accurately identify the reason for their absence in the files – they may have been unemployed, self-employed without employees, inactive, or temporarily in the public sector. As their previous status in the labor market (which is unobserved) may be correlated with their startup conditions in terms of human capital, we leave them out of the current analysis.

Hence, our final sample is composed of all the startup firms entering between 1992 and 2007 (excluding 2001), employing at least one wage earner at the moment of startup, and for which we can identify the business owner(s) (BOs) at entry, who must come from paid employment.⁵ In other words, we base our analysis on entrepreneurs that are potentially meaningful, i.e., entrepreneurs who do not enter to escape from unemployment (Reynolds et al., 2002) and who do not have intentions to always remain own account workers. Instead, these entrepreneurs are likely to have higher ambition levels and to have high impact in the economy (Levine and Rubinstein, 2013; Henrekson and Sanandaji, 2014). They are, therefore, relatively unconstrained and are more likely to have a choice about the quantity and quality of human capital they wish to employ at entry.

Data for the years 1986-1991 were only used to trace and characterize the experience of these BOs in the labor market. A total of 17,579 startup firms with complete information on the key variables of interest fulfill these conditions. About 22% of these firms operate in manufacturing industries, and more than half of them are founded by a solo entrepreneur. Table A.I in the Appendix presents the distribution of solo entrepreneurs' and teams' startups across 2-digit industries.

We analyzed the survival of these firms until the end of the period observed or until the moment of an eventual ownership change (depending on which of the options occurs first).⁶ The analysis stops in 2007, the last year for which we can accurately identify the exit of firms. Firm exit is identified by the moment when a firm ceases to answer the survey. Following previous studies that also use QP dataset (e.g., Mata and Portugal, 2002; Geroski

⁵Due to the missing data at the worker-level for 2001, we are not able to identify the founder of firms entering during that year, so entries occurring in 2001 had to be excluded.

⁶About 40% of the firms in our sample suffer changes in the entrepreneurial team founding the firm. After ownership changes, the business is no longer the same, as the entrepreneur-firm match may also change. It is not however our aim in this paper to study what explains these processes of ownership change, neither their impacts. For this reason, we censor the spells at the point of ownership change. We conduct some robustness checks to confirm that our results are not affected by this procedure.

et al., 2010), we have required an absence of the firm from the files larger or equal to two years in order to identify its definite exit. For this reason, data for 2008 and 2009 were only used to check the presence or absence of firms in QP files.

3.2 Measuring individuals' skills

Previous studies have recognized that finding the right measure of skills is not easy (Iranzo et al., 2008). Most of the existing studies constructed human capital proxies based on *observed* dimensions such as workers' educational attainment, age, earnings, or gender (e.g., Haltiwanger et al., 1999, 2007; Ilmakunnas et al., 2004), thereby disregarding *unobserved* differences such as innate ability or informal skills (see also Iranzo et al., 2008; Martins, 2008). However, Abowd et al. (1999) show that observed levels of workers' skill heterogeneity imperfectly reflect the *true* level of heterogeneity, which also includes an unobserved component at the individual-level.

For this reason, we use the multi-dimensional skill index developed by Portela (2001) to measure workers' and BOs' skills. Following most of the earlier literature, this measure is originally based on individuals' education (namely the number of schooling years attained by each person at each year). Besides this human capital dimension, the skill index for each individual is then adjusted in order to take into account her experience (age) and unobserved permanent heterogeneity (which includes, among others, innate ability, informal skills, and education quality) relative to other comparable individuals in the population.⁷ Individuals' unobserved skills are measured using fixed effects estimation.

Formally, the skill index of each individual i in each year t is computed as follows:

$$S_{it} = mschool * a_{school} * a_{experience} * a_{unobserved\ ability} \quad (1)$$

where $mschool$ is the average schooling years in the economy in each year. In other words, we first assume that each individual enters the labor market with the average education of the active population in that year, to then correct the index in order to take into account the actual position of each individual, in each year, in the schooling distribution. This correction is performed by multiplying $mschool$ by the following correction factor:

$$a_{school} = 0.5 + \frac{\exp((school_i - mschool)/sschool)}{1 + \exp((school_i - mschool)/sschool)} \quad (2)$$

⁷This multi-dimensional skill index has also inspired other authors to construct new and better quality measures in other contexts. Sá et al. (2004, 2012) use a similar index inspired on Portela's skill index to construct a composite measure of universities' education quality in The Netherlands. This index allows them to take into account different university attributes, as well as the relative position of each university in each attribute.

where $school_i$ is the schooling level (in years) of worker i and $sschool$ represents the standard deviation of schooling in the population. This correction factor takes values between 0.5 and 1.5, which intuitively means that individuals more (less) educated than the average will have a multiplicative correction factor larger (smaller) than 1.

At this stage, we have a standard human capital measure at the individual-level simply based on their education. However, even if we compare individuals with the same education, we still find a great heterogeneity among them – for instance, in terms of experience. In order to correct for this source of observed heterogeneity, we multiply the skill index of each individual by this second correction factor $a_{\text{experience}}$, which allows to adjust the skill measure of each individual either upwards or downwards, according to her relative position in the age distribution of the individuals with similar education attainment. This correction term is computed as follows:

$$a_{\text{experience}} = 0.5 + \frac{\exp((age_i - mage|school_i)/(sage|school_i))}{1 + \exp((age_i - mage|school_i)/(sage|school_i))} \quad (3)$$

where age_i is the age (in years) of worker i , $mage|school_i$ is the average age of the population within schooling level $school_i$, and $sage|school_i$ is its standard deviation. Again, individuals more (less) experienced than the average of individuals with the same level of education will have a multiplicative correction factor larger (smaller) than 1.⁸

As we are aware that two individuals with precisely the same education level and age may still be very different in their skills and productivity potential, we furthermore correct this skill index in order to take into account individuals' fixed effect, as a proxy for their unobserved and permanent productivity differential. In order to estimate the person fixed effect, we separately estimated a two high-dimensional fixed-effects wage equation using the procedure described in Guimarães and Portugal (2010), based on all the history we have for each individual in wage employment. The dependent variable was defined as the real hourly earnings (in logs).⁹ This wage equation controlled for individual's age (and its square), tenure (and its square), education, qualifications, year dummies, and, following Abowd et al. (1999), both worker and firm unobserved (permanent) heterogeneity. This allowed us to estimate the worker-specific effect (FE_i), which basically reflects the income the individual earns on top of or below what is expected, based on all her observed characteristics and taking into account the fixed effect of the respective firm. This was then introduced in the final correction factor of the skill index, computed as follows:

⁸The several human capital dimensions enter this index multiplicatively because they are assumed to be complementary.

⁹Hourly earnings correspond to the ratio between total regular payroll (base wages and regular benefits) and the total number of normal hours worked in the reference period. Earnings were deflated using the Consumer Price Index. Outliers (i.e., the 1% with highest and lowest real hourly log earnings in each year) were removed from the estimations.

$$a_{\text{unobserved ability}} = 0.5 + \frac{\exp((FE_i - mFE|school_i, age_i)/(sFE|school_i, age_i))}{1 + \exp((FE_i - mFE|school_i, age_i)/(sFE|school_i, age_i))} \quad (4)$$

where $mFE|school_i, age_i$ denotes the average of worker fixed effects for individuals with the same schooling and age, and $sFE|school_i, age_i$ is the standard deviation of those effects.

We now have a skill measure that allows two individuals with the same education level and age to be treated as potentially different in terms of skills, as long as their unobserved permanent skills are different.

After computing the skill index for each individual i in each year t , we were able to construct firm-level measures of workers' and BOs' skills. For each startup firm identified in the data, we computed the average skill index of the workers hired at entry, as well as the (average) skill index of the BO(s) founding the firm.¹⁰ The interpretation of this skill measure is also intuitive, as it provides information on the average education level of workers and BOs, adjusted in order to take into account the heterogeneity of individuals (in terms of experience and unobserved permanent skill dimensions), even when we compare groups of persons with the same education level.

3.3 Modeling human capital choices and deviations from benchmarks

Studies concerned with sub-optimal entry positions and some kind of deviations from the *benchmark* – namely in startup size – normally use the average or the median choice of the industry as benchmark (e.g., Mata and Portugal, 1994; Audretsch et al., 1999, 2000). However, there is also a great heterogeneity at the firm and entrepreneur levels that may make such deviations more likely for some firms/entrepreneurs than others. It is likely that different entrepreneurs with different resources and skill levels enter with different human resources. For this reason, simply comparing the decisions of each firm with those in the same industry may be quite limited, by assigning the same benchmark to all startups entering in the same industry in the same year.

In view of that, we use a new approach to first estimate the startup size and the level of skills that shall be used as benchmark for each firm, at each point in time over their lifecycle. The estimation is based on a system of recursive simultaneous equations, where we allow the key HR decisions at the firm-level to be jointly determined, and to be correlated with

¹⁰Given that the person-specific effect is estimated based on individuals' spells in wage employment, the average skill index of the BO(s) was only possible to compute for those BOs with a record of at least two years in paid employment.

the human capital of the BOs and the ownership structure of the startup firm. The system below summarizes the four equations of interest, which were estimated using data on all active firms during the period 1992-2007 (i.e., both entrants and incumbents), in order to take into account all the competition faced by the 17,579 startups firms in our sample.¹¹ All equations include dummy variables for firm age (η_t), year (δ_y), 2-digit industry (γ_j), and Nuts III region (α_r).

$$\left\{ \begin{array}{l} Solo_{it} = \mathbf{BOHC}'_{it}\boldsymbol{\theta}_{11} + \mathbf{BOExp}'_{it}\boldsymbol{\theta}_{12} + \theta_{13} Si ze_{jt} + \theta_{14} WSkills_{jt} + \\ \quad + \eta_t + \delta_y + \gamma_j + \alpha_r + \varepsilon_{1it} \\ BOSkills_{it} = \theta_{21} Si ze_{jt} + \theta_{22} WSkills_{jt} + \theta_{23} BOSkills_{jt} + \theta_{24} Solo_{it} + \\ \quad + \eta_t + \delta_y + \gamma_j + \alpha_r + \varepsilon_{2it} \\ WSkills_{it} = \mathbf{BOExp}'_{it}\boldsymbol{\theta}_{31} + \theta_{32} WSkills_{jt} + \theta_{33} BOSkills_{jt} + \\ \quad + \eta_t + \delta_y + \gamma_j + \alpha_r + \varepsilon_{3it} \\ Si ze_{it} = \mathbf{BOExp}'_{it}\boldsymbol{\theta}_{41} + \theta_{42} Si ze_{jt} + \theta_{43} WSkills_{jt} + \theta_{44} Solo_{it} + \\ \quad + \theta_{45} BOSkills_{it} + \eta_t + \delta_y + \gamma_j + \alpha_r + \varepsilon_{4it} \end{array} \right. \quad (5)$$

The first equation in the system corresponds to the decision of entering (or staying in) the business alone or in a team. We expect this choice to be closely related to BOs' general and specific human capital (measured by a set of variables in vectors \mathbf{BOHC} and \mathbf{BOExp}), as well as to the size and quality of the workforce to be employed. However, as these two last variables refer to endogenous choices in the system, and the equations must be designed recursively, we use the average size and the average skill index of the workers employed by firms operating in the same industry j and year t ($Si ze_{jt}$ and $WSkills_{jt}$, respectively) as exogenous proxies for the quantity and quality of HR to be hired by the firm.

As discussed above, the decision between running a business alone and sharing the ownership with other(s) may have implications for the overall skill level of the entrepreneurial team (Ucbasaran et al., 2003; Forbes et al., 2006). In this sense, the skill level of the BOs is no longer exogenous. It may be the result of some strategic combination of team members in order to benefit from several kinds of complementarities – including skill complementarities. Additionally, we also expect that entrepreneurs hiring a larger and more skilled workforce at entry are more skilled on average. Again, these two variables cannot be introduced yet as independent variables. Therefore, we include the average size and skills of the workers employed by close competitors instead. Last but not least, we may expect that BOs might want – or need – to be at least as skilled as their competitors, so we also include the average

¹¹We have to restrict the analysis to firms whose BO(s) can be identified in QP files. We are able to identify and track over time a total of 194,357 firms with known BO(s).

BOs' skill level of competitors (those in the same industry, in the same year) in the set of independent variables of this second equation.

The final equations of the system correspond to the quality and quantity of workers to employ each year over firms' lifecycle. From the above discussions, we expect a positive association between workers' and BOs' skills, and also a strong influence from competitors' human capital choices. The size of the workforce to be employed over firms' lifecycles is furthermore believed to be related to the average size of the firms operating in the same industry, and to the decision of entering alone or in teams.

These inter-related entrepreneurial decisions were estimated using the method developed by Roodman (2011). This forms the basis for the estimation of individual benchmarks for human capital quantity and quality (and deviations therefrom). His technique extends the logic of previous seemingly unrelated regression models by allowing the consistent and efficient estimation of fully observed recursive equation systems. This is relevant when endogenous variables in the system influence each other, as in our case. Besides, it allows the combination of a broad panoply of models that were previously hard to estimate within the same system of equations (e.g., probit models, linear regressions, truncated regressions). This is also crucial in our case. Table 1 summarizes the results.

Insert Table 1 here

The estimation results indicate that older, more educated, and more experienced individuals (both in the industry and in management positions) tend to be more prone to start a business alone. In contrast, a longer experience as a BO is apparently negatively associated with the probability of entering alone, maybe because those who were entrepreneurs in the past are more aware of the risks of founding a startup alone. The results in the first column also suggest a positive association between the average skill level of the workers employed by competitors and the propensity of entering alone. This may suggest that solo entrepreneurs in our sample are more risk-takers and more skilled on average than those entering in teams.

The results obtained for the second equation confirm that the level of skills exhibited by BOs is significantly and positively related to their decision of running the firm alone. As expected, BOs' skill levels are also positively aligned with the average skill level of competitors in the same industry.

Finally, the results confirm the existence of a strong and positive relation between workers' and BOs' skills (third equation). More skilled and more experienced BOs seem to hire not only more skilled workers, but also larger workforces on average. Moreover, the quantity (quality) of workers hired over firms' lifecycle are found to be significantly and positively associated with the average quantity (quality) of workers employed by competitors in the same industry.

Having estimated the interrelated human capital decisions of startup firms through this system of equations, we are now able to compute the so-called deviations from the *benchmark* HR choices. This proceeds in two steps. First, based on the results of Table 1, we predict the (correlated) quantity and quality of human capital to be employed at the moment of entry, given a number of characteristics of firms and their BOs, the industry and the region where the firm operates, as well as the macroeconomic environment, and the average choices of the competitors operating in the same industry. Second, we measure the extent to which each firm actually deviates from their close competitors (i.e., entrepreneurs with active firms in the same year, region, and 2-digit industry, and with similar levels of general and specific human capital, including the fixed unobserved component), by calculating the percentage difference between each firm’s observed human capital inputs and the *predicted* values from the above system of equations.

4 Deviations from human capital benchmarks and stickiness of entry choices

Table 2 provides summary statistics for various dimensions that characterize startup conditions – including the quantity and quality of HR at entry – and that are expected to correlate with the survival prospects of the firms under analysis. Given their different capital intensity, we distinguish between startups in manufacturing and services. We furthermore compare solo entrepreneurs with teams in each sector, as they may differ in their initial capital constraints and, therefore, HR choices.

Insert Table 2 here

These statistics confirm that startups in services enter at a smaller scale than those in manufacturing, and that solo entrepreneurs hire a smaller number of workers at entry than teams. Workers hired at entry in services are more skilled on average than those in manufacturing, and so are the entrepreneurs founding these firms. Solo entrepreneurs and teams also differ in their average skill levels, with the former being more skilled on average. Finally, BOs in manufacturing – and especially solo entrepreneurs – present a longer industry-specific experience, while BOs in teams have more entrepreneurial experience from previous businesses than those entering alone. All these differences are statistically significant at the 1% level.

Though these first statistics show that there is a great heterogeneity in startup conditions, they are not informative about the extent to which these startups are matching or

deviating from the benchmarks predicted for them in the previous section. Figure 1 illustrates the kernel density of deviations between firms' *observed* startup size and the *benchmark* size predicted for them at entry, according to the estimations in Table 1. We find that most startups enter below the estimated startup size. Negative deviations are found to be more severe in manufacturing industries, and especially among startups founded by solo entrepreneurs, who may face larger financial constraints, especially in these industries.

Insert Figure 1 here

Figure 2 illustrates the deviations from the predicted levels of workers' skills at the moment of entry. In this case, deviations are more moderate and centered at zero, and differences between solo entrepreneurs and teams are not so evident. However, we still find larger deviations in workers' skills among firms established in services than in manufacturing industries.

Insert Figure 2 here

Our data also suggest that entry decisions related to the quantity and quality of initial human capital are rather sticky, as initial relative positions seem to be difficult to change afterwards. Table 3 reports the observed probabilities of moving along the distribution of size and skill deviations during the first three years of activity.

Insert Table 3 here

These statistics confirm that firms entering smaller than expected tend to remain below the benchmark size in the long run. The great majority of firms whose startup size is smaller than the estimated benchmark by more than 50% remain in the same deviating group three years later. About 32% of them reduce this gap, but not enough to match or surpass the estimated benchmark for size at that age.

A similar pattern is found for skill deviations. The bold values in the main diagonal precisely show that there is a great share of firms staying in the same interval three years later. Actually, most of the firms either stay in the same deviating interval, or move to adjacent positions. This confirms that it is not easy to adjust the initial HR choices of entrepreneurs. This persistence applies to all four sub-groups of firms described above.

5 Deviations from human capital benchmarks and firm survival

So far we have shown that initial HR decisions vary to a rather large extent across firms, even if we look at a *micro cosmos* of firms operating in the same sector and region, founded in the same year, and by BOs with similar levels of skills and experience. Moreover, deviations from the expected levels of skills and numbers of employees are quite sticky over time. Hence, initial HR decisions may have long term consequences. In this section, we assess to what extent these initial decisions are related to firm survival.

We use duration models to study how deviations in the quantity and the quality of initial human capital may influence firm survival. As our data come from an annual survey, durations are grouped into yearly intervals. Those firms that are still operating at the end of the period under observation are right-censored observations. This sampling plan is properly accommodated in the framework of discrete time duration models. We use a standard semiparametric discrete time proportional hazard model and control for unobserved heterogeneity at the firm/BO-level by incorporating a Gamma-distributed multiplicative term (Lancaster, 1990) in the hazard equation. Formally, the hazard rate of each firm i exiting at discrete time t_j , $j = 1, 2, \dots$, given survival until then, may be written as follows:

$$h_{ij} = 1 - \exp\{-\exp[\gamma(t) + X_i(t)\beta + \log(\varepsilon_i)]\} \quad (6)$$

where $\gamma(t)$ is a set of indicator variables for different duration intervals, thus flexibly describing the pattern of duration dependence in hazard rates, $X_i(t)$ is a vector of variables that are expected to be associated with firm survival, β is a vector of unknown parameters to be estimated, and ε_i is the Gamma-distributed random term describing firm/BO unobserved permanent heterogeneity.

Vector X includes the key variables of interest – i.e., deviations from estimated benchmarks for HR quantity and quality – as well as a number of controls that may also influence firm survival, namely BOs’ specific human capital measures, location in urban regions, and a set of industry variables that are typically included in firm survival studies (concentration, minimum efficient scale, industry employment growth rate, industry agglomeration, and entry rates) (e.g., Mata and Portugal, 2002; Geroski et al., 2010).

5.1 Empirical results using linear spline regressions

We may expect that startups deviating negatively (positively) from the benchmark quantity and quality of initial HR enter with a comparative disadvantage (advantage). However,

imposing a linear relationship between size/skill deviations and firm hazard may be quite limiting, especially if negative and positive deviations reflect different, though random, conditions. For this reason, the marginal effects of positive and negative deviations are not necessarily symmetric.

In view of that, we start by estimating linear spline regressions for the relationship between deviations from benchmark size/skills at entry and firm hazard rate. As a starting point, we impose the cut-off (*knot*) to be at zero. This allows us to test whether the estimated relationship has a different (linear) slope according to the sign of deviations.

Table 4 summarizes the results and compares the four groups of startups. The coefficient of *Size (Skill) deviations < 0%* gives us the estimated slope for the (linear) relationship between deviations from the benchmark size (skills) and hazard rate, for those firms entering *below the benchmarks*. Then, the coefficient of *Size/Skill deviations > 0%* measures the *marginal change in the slope* of that (linear) relationship for firms entering *above the benchmarks*. Whenever this second coefficient is statistically significant, it means that the relationship between size/skill deviations and firm hazard significantly changes its slope at the benchmark (i.e., at size/skill deviations=0) – so the marginal effects of positive and negative deviations are not symmetric.

Insert Table 4 here

Solo entrepreneurs entering in manufacturing industries at a scale smaller than the benchmark seem to suffer significant survival penalties. On the other hand, entering larger than the benchmark size barely affects their survival¹², but it significantly improves the survival of entrepreneurial teams in those industries. For startups in services, size deviations do not seem to be significantly related to firm exit risk. These results confirm the existence of larger scale economies in manufacturing than in services, besides corroborating the so-called *liability of smallness* (Brüderl et al., 1992; Mata and Portugal, 1994), which seems to be particularly important for solo entrepreneurs in manufacturing.

The results obtained for skill deviations are rather mixed. Hiring a more skilled workforce than expected is actually found to increase the exit risk of solo entrepreneurs operating in services. This may suggest that they may be hiring over-qualified workers, which may result in increased costs without a significant increase in productivity – if we take into account that most of these firms are either small shops for wholesale and retail trade, or restaurants, cafes and bakeries (see Table A.I in the Appendix). However, for entrepreneurial teams, we find that the survival penalties may be quite severe for those entering with an initial

¹²For positive deviations, the slope of the linear relationship is still negative, but close to zero: $-1.0641 + 0.9704 = -0.0937$.

workforce whose skill level is below the estimated benchmark, both in manufacturing and services.

5.2 Non-linear effects of deviations from human capital benchmarks

The previous results confirmed that positive and negative deviations from HR benchmarks are not symmetrically associated with firm hazard. However, the cut-off point at zero was somehow ad-hoc imposed, relying on the assumption that the estimated benchmarks for startup size and initial workers' skills correspond to the *optimal* values. As previously discussed, this may actually not be the case. Furthermore, the relationship between human capital deviations and exit rate may be non-linear. For these reasons, we estimate alternative specifications using a quadratic approximation for the relationship between human capital deviations and firm hazards. Table 5 summarizes the results.

Insert Table 5 here

These specifications indicate that there is a non-linear quadratic association between size deviations and firm hazard in manufacturing. A non-linear relationship between skill deviations and exit rates is also found for startups established in services. However, in both cases, the shape of the relationship differs according to the ownership structure at entry (solo entrepreneurs versus teams). Also, the inflexion point is not necessarily at zero, which indicates that the benchmark for startup size and workers' skills at entry may not be optimal – at least for firm survival.

Given that the non-linear relationship between human capital deviations and new ventures exit may not be really quadratic, we further extend our analysis by following a more flexible approach that allows us to find out the parametric function of size and skill deviations that better fits our data. The method is proposed by Royston and Sauerbrei (2007) and it is based on multivariable regression spline models. Their algorithm automatically selects the regression spline model (which is not necessarily linear) and the respective knots position that best predicts the outcome variable from each independent variable. We use it as a final check for the estimated relationships between deviations from human capital benchmarks and new ventures exit risk.

5.3 Multivariate spline model with cubic regression splines

We estimate the Royston and Sauerbrei's multivariable regression spline model for survival time data, separately for the four groups of startups under analysis. We then produce the

estimated values and the respective confidence intervals for each group. Figure 3 presents the results obtained for deviations from benchmark startup size. Figure 4 provides comparable figures for the relationship between deviations from benchmark workforce skills and the log of relative hazards.¹³

The results confirm the non-linear association between size deviations and firm hazards in manufacturing. Consistent with the previous results, we find that solo entrepreneurs in these industries suffer higher hazards if they enter below the estimated benchmark for startup size. This is in line with prior studies suggesting that firms not achieving the so-called minimum efficient scale – which is the common benchmark for firm size – face higher exit rates (e.g., Audretsch and Acs, 1990; Mata and Portugal, 2002).

Insert Figure 3 here

Our results further show that firms entering manufacturing industries at a larger scale than the benchmark can still reduce their exit risk comparatively to those entering at the predicted benchmark size (i.e., with deviations from predicted startup size equal to zero). This indicates that manufacturing startups in our sample are, on average, entering at sub-optimal scales. Whether or not this is a common pattern among startups in manufacturing industries is a topic that deserves further research, using data for other countries.

For solo entrepreneurs in services, we do not find any statistically significant association between size deviations and exit risk. Given the micro-sized nature of these startup firms and the industries where they tend to be concentrated, differences of one or two workers often correspond to great percentage deviations from the estimated benchmark, which in practice do not seem to affect the survival prospects of these businesses.¹⁴ For teams, the model instead finds a linear negative association between size deviations and firm hazards, suggesting that entering below (above) the benchmark startup size results in significant survival penalties (gains).

Regarding deviations from the benchmark workforce skills at the moment of entry, the results confirm the mixed evidence already documented in the previous sections. A more skilled workforce seems to increase the exit risk for startups in services, especially those founded by solo entrepreneurs.¹⁵ First, contrary to manufacturing industries, there might

¹³We do not present the estimation results because they are not so informative and easy to interpret as the plots obtained after estimations, given the large number of knots and the existence on non-linear relationships linking some of the knots.

¹⁴Given the heterogeneity in services, we rerun our estimations separately for the sectors 61-63 (Wholesale Trade, Retail Trade, Restaurants and Hotels) and for the remaining services startups. The pattern illustrated in Figure 3 for solo entrepreneurs in services mainly reflects the results found for sectors 61-63. For solo entrepreneurs in other services, we find a more similar pattern to that found for manufacturing industries. The separate figures for these two groups of services industries are available from the authors upon request.

¹⁵We found no great difference between solo and teams in sectors 61-63 and those entering other services.

be no productivity gains by hiring such a highly skilled workforce in these sectors. Second, for solo entrepreneurs in particular – who may be more financially constrained than teams, on average – it may result in higher costs (namely a higher wage bill) that also reduce the viability of their business.

Insert Figure 4 here

For teams, however, we find that entering with a less skilled workforce may result in considerable survival penalties. Actually, the results further indicate that deviating above the estimated benchmark for workers’ skills may still reduce their hazard up to a certain point. This may suggest that entrepreneurial teams in our data enter with sub-optimal levels of workers’ quality compared to their close competitors – they can reduce their hazards by entering with a workforce more skilled than the benchmark. Given that, on average, they are less skilled than solo entrepreneurs, and that skill complementarities may be achieved between entrepreneurs and workers (e.g., Baptista et al., 2013), they become subject to larger survival penalties when they enter with a less qualified workforce. In summary, these different results found for deviations in human capital quality among solo entrepreneurs and teams may mostly reflect the heterogeneity in BOs’ quality, more than different ownership structures per se. Random deviations from the benchmark workforce quality may have different implications according to the quality of entrepreneurs themselves.

As a final robustness check, we repeated the analysis for i) the subsample of startups that never change their BOs under the period under analysis, and for ii) the same sample used above, but without imposing the right-censoring at the point of ownership change (i.e., also including the spells under the ownership of a new entrepreneur or a new team). The estimated relationships between size/skill deviations and firm hazard remained consistent with the patterns illustrated above in Figures 3 and 4. All these additional results may be available upon request.

6 Concluding Remarks

Despite the many discussions about comparative advantages in key resources (including HR), it is hard to find thorough analyses on the role of (relative) positions of startup firms in terms of the quantity and quality of their personnel, and the implications of deviating from benchmarks. We study the survival implications of startups’ relative positions in the quantity and quality of human resources at entry, taking into account the correlation between hiring decisions at entry and founders’ human capital. Our analysis compares startups in industries with very different entry barriers and capital requirements. Firms

founded by solo entrepreneurs and entrepreneurial teams are furthermore compared, given their potential differences regarding financial constraints and BOs' human capital levels.

Our results suggest that deviations from benchmark startup size may be particularly relevant in manufacturing industries, where economies of scale may be of greater importance. We find that many startups enter at sub-optimal scales, and that such inferior positions result in large survival penalties. In view of that, sub-optimal entry strategies based on the intention to learn about the firm itself and the market may be really costly, especially for solo entrepreneurs entering more capital-intensive industries.

In services, our results point out that there might be some risk of hiring overqualified workers, though it seems to be somewhat conditional on the skill levels of entrepreneurs themselves. Contrary to manufacturing industries, there might be no great productivity gains from entering with a highly skilled workforce in services, especially if the industry is not so knowledge-intensive. It may, instead, result in higher costs or even labor turnover later on, which reduces the sustainability of the business. This result deserves further research.

Our analysis also confirms that relative positions at entry are quite difficult to change during the first years of activity in the market, which imply that wrong initial choices might turn out to be a close to irreversible matter, with large penalties for survival. The evidence we find may offer insights and implications for both policy makers and entrepreneurs. By confirming the importance of certain relative positions at the moment of entry and the potential stickiness of initial HR choices, this paper reinforces the need for policies targeting entrepreneurs and newborn firms at a very early stage. Policy intervention aiming at reducing forecasting errors, information asymmetries, or entrepreneurs' ability and financial constraints seem to be of high value.

For prospective entrepreneurs, a similar caveat applies. Given the difficulty in reversing some initial decisions and the likely survival penalties associated with certain inferior and superior positions at the moment of startup, new and forthcoming BOs are encouraged to balance the potential effects of (un)intended deviations relative to closer competitors' choices when they decide to enter the market.

Further research on these questions, using comparable data for other countries with different labor market rigidities and different profiles of entrepreneurs, will be certainly appreciated by scholars, policy makers, and practitioners.

References

- [1] Abowd, J. M., Kramarz, F., Margolis, D. (1999b), "High wage workers and high wage firms", *Econometrica*, 67(2), 251-334.

- [2] Agarwal, R., Audretsch, D. B. (2001), “Does entry size matter? The impact of the life cycle and technology on firm survival”, *The Journal of Industrial Economics*, 49(1), 21-43.
- [3] Agarwal, R., Barney, J. B., Foss, N. J., Klein, P. G. (2009), “Heterogeneous resources and the financial crisis: implications of strategic management theory”, *Strategic Organization*, 7(4), 467-484.
- [4] Alpizar, F., Carlsson, F., Johansson-Stenman, O. (2005), “How much do we care about absolute versus relative income and consumption?”, *Journal of Economic Behavior & Organization*, 56(3), 405-421.
- [5] Åstebro, T., Herz, H., Nanda, R., Weber, R. A. (2014), “Seeking the roots of entrepreneurship: Insights from behavioral economics”, *Journal of Economic Perspectives*, 28(3), 49-70.
- [6] Audretsch, D. B., Acs, Z. J. (1990), “The entrepreneurial regime, learning, and industry turbulence”, *Small Business Economics*, 2(2), 119-128.
- [7] Audretsch, D. B., Houweling, P., Thurik, A. R. (2000), “Firm survival in the Netherlands”, *Review of Industrial Organization*, 16(1), 1-11.
- [8] Audretsch, D. B., Santarelli, E., Vivarelli, M. (1999), “Start-up size and industrial dynamics: some evidence from Italian manufacturing”, *International Journal of Industrial Organization*, 17(7), 965-983.
- [9] Baptista, R., Lima, F., Preto, M. T. (2013), “Entrepreneurial skills and workers’ wages in small firms”, *Small Business Economics*, 40(2), 309-323.
- [10] Bates, T. (1990), “Entrepreneur human capital inputs and small business longevity”, *The Review of Economics and Statistics*, 72(4), 551-559.
- [11] Benz, M., Frey, B. (2008), “Being independent is a great thing: Subjective evaluations of self-employment and hierarchy”, *Economica*, 75(298), 362-383.
- [12] Birley, S., Westhead, P. (1994), “A taxonomy of business start-up reasons and their impact on firm growth and size”, *Journal of Business Venturing*, 9(1), 7-31.
- [13] Bolger, F., Pulford, B. D., Colman, A. M. (2008), “Market entry decisions: Effects of absolute and relative confidence”, *Experimental Psychology*, 55(2), 113-120.
- [14] Brüderl, J., Preisendörfer, P., Ziegler, R. (1992), “Survival chances of newly founded business organizations”, *American Sociological Review*, 57(2), 227-242.
- [15] Cabral, L., Mata, J. (2003), “On the evolution of the firm size distribution: facts and theory”, *The American Economic Review*, 93(4), 1075-1090.

- [16] Camerer, C., Lovo, D. (1999), "Overconfidence and excess entry: An experimental approach", *The American Economic Review*, 89(1), 306-318.
- [17] Colombo, M. G., Delmastro, M., Grilli, L. (2004), "Entrepreneurs' human capital and the startup size of new technology-based firms", *International Journal of Industrial Organization*, 22(8-9), 1183-1211.
- [18] Cooper, A. C., Gimeno-Gascon, F. J., Woo, C. Y. (1994), "Initial human and financial capital as predictors of new venture performance", *Journal of Business Venturing*, 9(5), 371-395.
- [19] Cressy, R. (1996), "Are business startups debt-rationed?", *The Economic Journal*, 106(438), 1253-1270.
- [20] Feeser, H. R., Willard, G. E. (1990), "Founding strategy and performance: A comparison of high and low growth high tech firms", *Strategic Management Journal*, 11(2), 87-98.
- [21] Forbes, D. P., Borchert, P. S., Zellmer-Bruhn, M. E., Sapienza, H. J. (2006), "Entrepreneurial team formation: An exploration of new member addition", *Entrepreneurship Theory & Practice*, 30(2), 225-248.
- [22] Garicano, L., Hubbard, T. N. (2005), "Hierarchical sorting and learning costs: Theory and evidence from the law", *Journal of Economic Behavior & Organization*, 58(2), 349-369.
- [23] Garicano, L., Rossi-Hansberg, E. (2006), "Organization and inequality in a knowledge economy", *The Quarterly Journal of Economics*, 121(4), 1383-1435.
- [24] Geroski, P. A., Mata, J., Portugal, P. (2010), "Founding conditions and the survival of new firms", *Strategic Management Journal*, 31(5), 510-529.
- [25] Gilbert, C. G. (2005), "Unbundling the structure of inertia: resource versus routine rigidity", *Academy of Management Journal*, 48(5), 741-763.
- [26] Guimarães, P., Portugal, P. (2010), "A simple feasible alternative procedure to estimate models with high-dimensional fixed effects", *The Stata Journal*, 10(4), 628-649.
- [27] Haltiwanger, J. C., Lane, J. I., Spletzer, J. R. (1999), "Productivity differences across employers: The roles of employer size, age, and human capital", *The American Economic Review*, 89(2), 94-98.
- [28] Haltiwanger, J. C., Lane, J. I., Spletzer, J. R. (2007), "Wages, productivity and the dynamic interaction of businesses and workers", *Labour Economics*, 14(3), 575-602.
- [29] Hamermesh, D. S., Pfann, G. A. (1996), "Adjustment costs in factor demand", *Journal of Economic Literature*, 34(3), 1264-1292.

- [30] Henrekson, M., Sanandaji, T. (2014), “Small business activity does not measure entrepreneurship”, *Proceedings of the National Academy of Sciences of the United States of America*, 111(5), 1760-1765.
- [31] Hmieleski, K., Baron, R. (2009), “Entrepreneurs’ optimism and new venture performance: A social cognitive perspective”, *Academy of Management Journal*, 52(3), 473-488.
- [32] Holtz-Eakin, D., Joulfaian, D., Rosen, H. S. (1994), “Sticking in out: Entrepreneurial survival and liquidity constraints”, *Journal of Political Economy*, 102(1), 53-75.
- [33] Hvide, H., Panos, G. (2014), “Risk tolerance and entrepreneurship”, *Journal of Financial Economics*, 111(1), 200-223.
- [34] Hyytinen, A., Lahtonen, J., Pajarinen, M. (2014), “Forecasting errors of new venture survival”, *Strategic Entrepreneurship Journal*, 8(4), 283-302.
- [35] Ilmakunnas, P., Maliranta, M., Vainioma, J. (2004), “The roles of employer and employee characteristics for plant productivity”, *Journal of Productivity Analysis*, 21(3), 249-276.
- [36] Iranzo, S., Shivardi, F., Tosetti, E. (2008), “Skill dispersion and firm productivity: An analysis with employer-employee matched data”, *Journal of Labor Economics*, 26(2), 247-285.
- [37] Jovanovic, B. (1982), “Selection and evolution of the industry”, *Econometrica*, 50(3), 649-670.
- [38] Katila, R., Chen, E. L., Piezunka, H. (2012), “All the right moves: How entrepreneurial firms compete effectively”, *Strategic Entrepreneurship Journal*, 6(2), 116-132.
- [39] Koellinger, P., Minniti, M., Schade, C. (2007), “‘I think I can, I think I can’: Overconfidence and entrepreneurial behavior”, *Journal of Economic Psychology*, 28(4), 502-527.
- [40] Koudstal, M. Sloof, R., van Praag, M. (2015), “Risk, uncertainty, and entrepreneurship: Evidence from a lab-in-the-field experiment”, Forthcoming in *Management Science*.
- [41] Lancaster, T. (1990), *The Econometric Analysis of Transition Data*, Econometric Society Monographs, 17, Cambridge University Press.
- [42] Levine, R., Rubinstein, Y. (2013), “Smart and illicit: Who becomes an entrepreneur and do they earn more?”, NBER Working Paper No. 19276.
- [43] Lucas, R. (1978), “On the size distribution of business firms”, *The Bell Journal of Economics*, 9(2), 508-523.

- [44] Martins, P. (2008), “Dispersion in wage premiums and firm performance”, *Economics Letters*, 101(1), 63-65.
- [45] Mata, J., Machado, J. A. F. (1996), “Firm startup size: A conditional quantile approach”, *European Economic Review*, 40(6), 1305-1323.
- [46] Mata, J., Portugal, P. (1994), “Life duration of new firms”, *The Journal of Industrial Economics*, 42(3), 227-245.
- [47] Mata, J., Portugal, P. (2002), “The survival of new domestic and foreign-owned firms”, *Strategic Management Journal*, 23(4), 323-343.
- [48] Pakes, A., Ericson, R. (1998), “Empirical implications of alternative models of firm dynamics”, *Journal of Economic Theory*, 79(1), 1-45.
- [49] Parker, S. C., van Praag, M. (2006), “Schooling, capital constraints, and entrepreneurial performance: the endogenous triangle”, *Journal of Business & Economics Statistics*, 24(4), 416-431.
- [50] Portela, M. (2001), “Measuring skill: a multi-dimensional index”, *Economics Letters*, 72(1), 27-32.
- [51] Reynolds, P. D., Bygrave, W. D., Autio, E., Cox, L. W., Hay, M. (2002), Global Entrepreneurship Monitor 2002, Executive Report. Babson College, Ewing Marion Kauffman Foundation and London Business School.
- [52] Roodman, D. (2011), “Fitting fully observed recursive mixed-process models with cmp”, *The Stata Journal*, 11(2), 159-206.
- [53] Royston, P., Sauerbrei, W. (2007), “Multivariable modeling with cubic regression splines: A principled approach”, *The Stata Journal*, 7(1), 45-70.
- [54] Sá, C., Florax, R. J. G. M., Rietveld, P. (2004), “Determinants of the regional demand for higher education in The Netherlands: A gravity model approach”, *Regional Studies*, 38(4), 375-392.
- [55] Sá, C., Florax, R. J. G. M., Rietveld, P. (2012), “Living arrangement and university choice of Dutch prospective students”, *Regional Studies*, 46(5), 651-667.
- [56] Santarelli, E., Vivarelli, M. (2007), “Entrepreneurship and the process of firms’ entry, survival and growth”, *Industrial and Corporate Change*, 16(3), 455-488.
- [57] Solnick, S. J., Hemenway, D. (1998), “Is more always better? A survey on positional concerns”, *Journal of Economic Behavior & Organization*, 37(3), 373-383.
- [58] Ucbasaran, D., Lockett, A., Wright, M., Westhead, P. (2003), “Entrepreneurial founder teams: factors associated with member entry and exit”, *Entrepreneurship Theory & Practice*, 28(2), 107-128.

TABLES

Table 1. Recursive mixed-process model for entrepreneurs' key HR decisions
(All firms, entrants and incumbents, 1992-2007, Portugal)

	Solo entrepreneur	BOs' skills	Workers' skills	Workforce size
BOs' age	0.0325*** (0.0002)			
BOs' education	0.1101*** (0.0004)			
BOs' industry-specific (2d) experience	0.0019*** (0.0006)		0.0132*** (0.0014)	0.1045*** (0.0079)
BOs' entrepreneurial experience	-0.0119*** (0.0011)		0.0211*** (0.0026)	0.2495*** (0.0156)
BOs' experience in management positions	0.0458*** (0.0011)		0.0089*** (0.0029)	0.2103*** (0.0164)
Average firm size in the same industry (2d)	-0.0048 (0.0043)	0.0197 (0.0127)		0.8824*** (0.0445)
Average workers' skill in the same industry (2d)	0.0427*** (0.0193)	0.0374 (0.0655)	0.6157*** (0.0351)	-0.4771* (0.2557)
Average BOs' skills in the same industry (2d)		0.7370*** (0.0357)		
Solo entrepreneur		4.9312*** (0.0117)		-3.3677*** (1.1431)
BOs' skills			0.2688*** (0.0029)	0.4847** (0.2306)
Year dummies	Yes	Yes	Yes	Yes
Firm age dummies	Yes	Yes	Yes	Yes
Industry (2d) dummies	Yes	Yes	Yes	Yes
Region (Nuts III) dummies	Yes	Yes	Yes	Yes
Number of observations			488,302	
Log likelihood			2,740,379.7	

Notes: ***, **, and *, mean statistically significant at the 1%, 5%, and 10% levels respectively. Values in parentheses are standard errors. The estimation was performed with the cmp command written for Stata by Roodman (2011). The first equation corresponds to a probit model for the binary decision of entering/staying alone as a BO, or sharing the ownership of the firm with other(s). The last equation for workforce size was estimated with a truncated regression, given the lower bound of 1 employee in our dataset.

Table 2. Descriptive statistics (Portugal, startups entering during the period 1992-2007, excluding 2001)

	All startups	Solo entrepreneurs Manufacturing	Solo entrepreneurs Services	Entrepreneurial teams Manufacturing	Entrepreneurial teams Services
Initial workforce size (nr. employees)	4.425 (4.011)	5.370 (5.595)	3.094 (2.055)	7.764 (7.000)	4.944 (3.219)
Initial workforce skills (average skill index)	6.283 (2.022)	5.628 (1.698)	6.483 (2.144)	5.677 (1.544)	6.426 (1.994)
BOs' skills (average skill index)	8.055 (2.754)	7.294 (2.646)	8.513 (2.910)	7.081 (2.234)	7.978 (2.551)
BOs' industry-specific (2d) experience (years)	2.107 (2.890)	3.498 (3.607)	1.936 (2.935)	2.666 (2.754)	1.671 (2.342)
BOs' entrepreneurial experience (years)	1.365 (1.142)	1.257 (1.134)	1.255 (1.059)	1.563 (1.260)	1.503 (1.199)
BOs' management experience (years)	1.231 (1.590)	1.046 (1.535)	1.284 (1.741)	1.131 (1.282)	1.254 (1.456)
Number of firms	17,579	2,021	8,180	1,826	5,552

Notes: Descriptive statistics refer to the entry year. Values in parentheses correspond to standard deviations.

Table 3. Transition probability matrix for size and skill deviations three years after entry (all startups, 1992-2007)^a

		Deviations from the size benchmark 3 years after entry ^b					
		[-100%, -50%[[-50%, -10%[[-10%,10%]]10%,50%]]50%,100%]	>100%
Deviations from the size benchmark at entry	[-100%, -50%[57.8%	31.5%	4.6%	3.0%	1.5%	1.6%
	[-50%, -10%[23.7%	48.5%	12.2%	8.9%	3.6%	3.1%
	[-10%,10%]	7.9%	36.4%	19.6%	19.1%	10.6%	6.4%
]10%,50%]	4.8%	22.2%	18.1%	26.8%	15.1%	13.0%
]50%,100%]	2.9%	12.4%	9.6%	23.5%	20.8%	30.8%
	>100%	1.7%	4.4%	5.1%	9.6%	14.1%	65.1%
		Deviations from the workers' skills benchmark 3 years after entry ^b					
		[-100%, -50%[[-50%, -10%[[-10%,10%]]10%,50%]]50%,100%]	>100%
Deviations from the workers' skills benchmark at entry	[-100%, -50%[28.8%	55.2%	10.4%	5.6%	0.0%	0.0%
	[-50%, -10%[1.3%	58.9%	27.7%	11.3%	0.8%	0.1%
	[-10%,10%]	0.3%	20.9%	54.5%	23.4%	0.8%	0.0%
]10%,50%]	0.1%	9.8%	26.5%	60.4%	3.2%	0.1%
]50%,100%]	0.0%	4.6%	12.9%	42.7%	37.4%	2.5%
	>100%	0.0%	11.8%	23.5%	23.5%	5.9%	35.3%

^a Only for firms that survived at least 3 years. Each row sums 100%. ^b Deviations from the predicted size/skills three years after entry are computed as the percentage difference between each firm *observed* size/skills and the estimated *size/skill benchmark* at the third year of activity. The benchmarks for size/skills three years after entry were obtained from the recursive system of equations reported in Table 1.

Table 4. Discrete-time hazard models with linear splines for deviations from benchmark size and skills (knots at zero)

	SOLO Manufacturing	SOLO Services	TEAMS Manufacturing	TEAMS Services
Size deviations < 0%	-1.0641*** (0.1057)	0.1279 (0.1600)	-0.0608 (0.1875)	-0.3638 (0.2722)
Size deviations > 0%	0.9704*** (0.1215)	-0.1661 (0.1649)	-0.6103*** (0.2193)	0.4486 (0.2757)
Skill deviations < 0%	-0.1753 (0.3402)	0.3992 (0.3287)	-1.4641** (0.6060)	-1.2121** (0.5185)
Skill deviations > 0%	-0.2300 (0.2573)	0.5300** (0.2617)	1.2964*** (0.4607)	0.6373 (0.4138)
Number of observations	6,749	27,316	5,633	15,543
Log likelihood	-4,048.4	-10,046.8	-2,414.2	-6,734.0

*, **, and *** mean significant at the 10%, 5%, and 1% levels, respectively. Values in parentheses are standard errors. All the specifications control for startup size, average workers' skills at entry, founders' skill index and specific human capital measures (2-digit industry experience, entrepreneurial experience, management experience), an indicator variable for firms established in urban areas (Porto or Lisbon), a set of 2-digit industry-level variables (HH index, minimum efficient scale, industry annual growth in terms of employment, industry agglomeration, and entry rate), year dummies, and duration dummies.

Table 5. Discrete-time hazard models with a non-linear relationship between deviations from human capital benchmarks and firm hazard

	SOLO Manufacturing	SOLO Services	TEAMS Manufacturing	TEAMS Services
Size deviations	-0.6380*** (0.0665)	-0.0172 (0.0795)	-0.3498** (0.1535)	-0.5366*** (0.1119)
Size deviations squared	0.2309*** (0.0286)	-0.0050 (0.0314)	-0.1667*** (0.0608)	0.0884 (0.0508)
Skill deviations	-0.2885 (0.3089)	0.5616* (0.2882)	-0.8224 (0.5545)	-0.4021 (0.2621)
Skill deviations squared	-0.1539 (0.1871)	0.4323** (0.1699)	0.4543 (0.3385)	0.7143*** (0.1235)
Number of observations	6,749	27,316	5,633	15,543
Log likelihood	-4,050.3	-16,403.1	-2,416.5	-6,725.3

*, **, and *** mean significant at the 10%, 5%, and 1% levels, respectively. Values in parentheses are standard errors. All the specifications include the same controls as in Table 4.

FIGURES

Fig. 1. Deviations from benchmark startup size

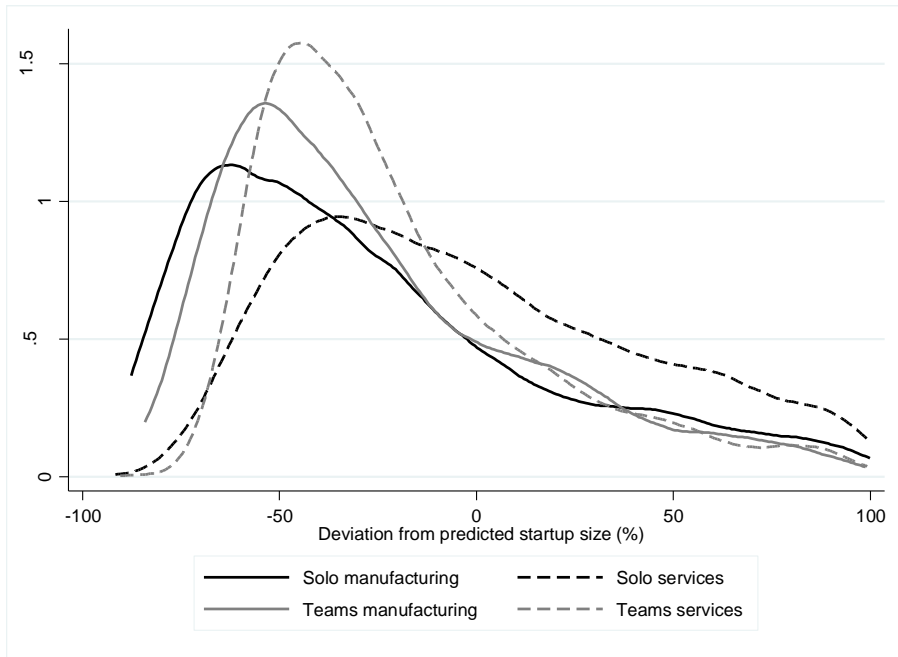


Fig. 2. Deviations from benchmark workforce skills at entry

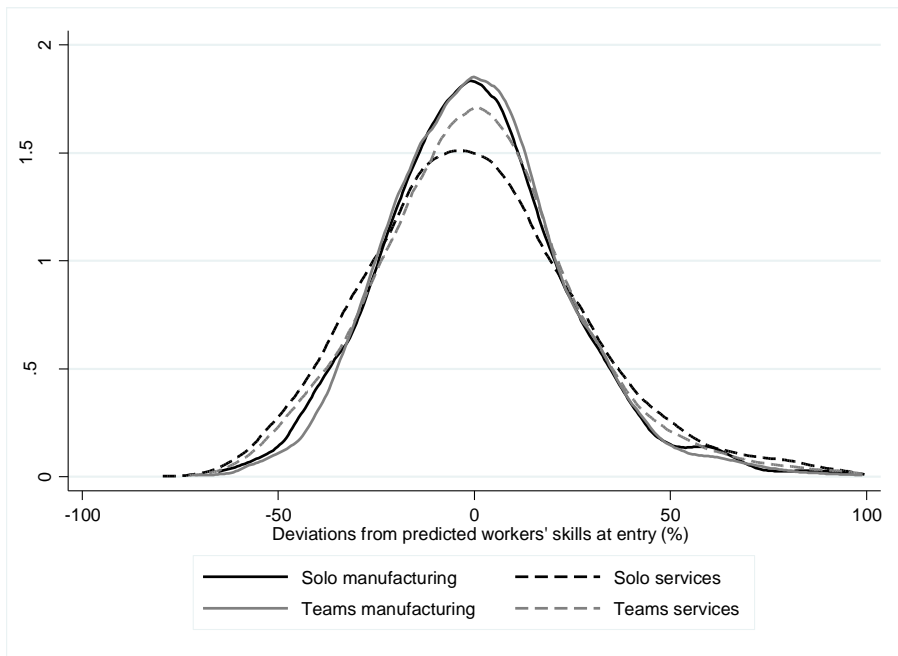


Fig. 3. Estimated relationship between deviations from benchmark startup size and log (relative hazards), by groups of startups (grey lines are 95% pointwise confidence intervals)

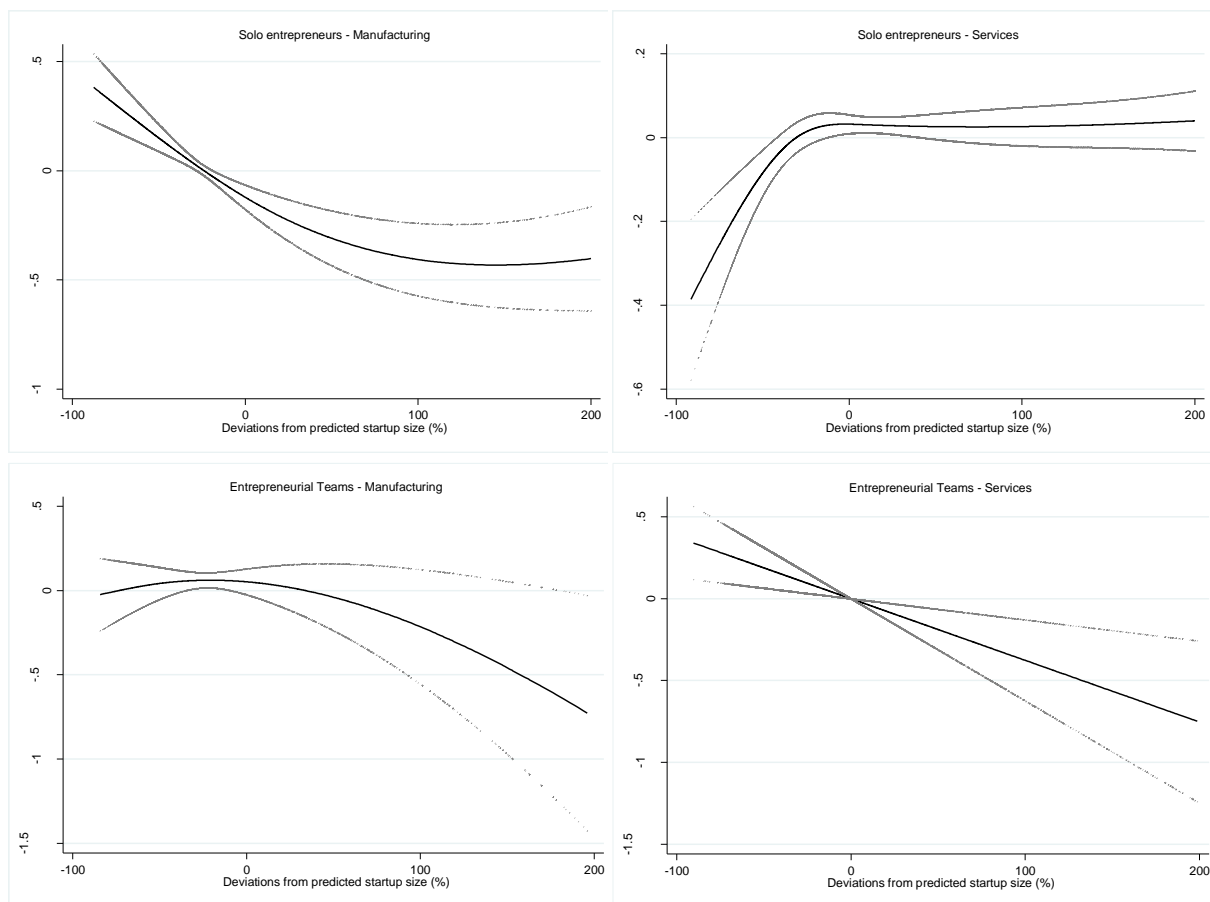
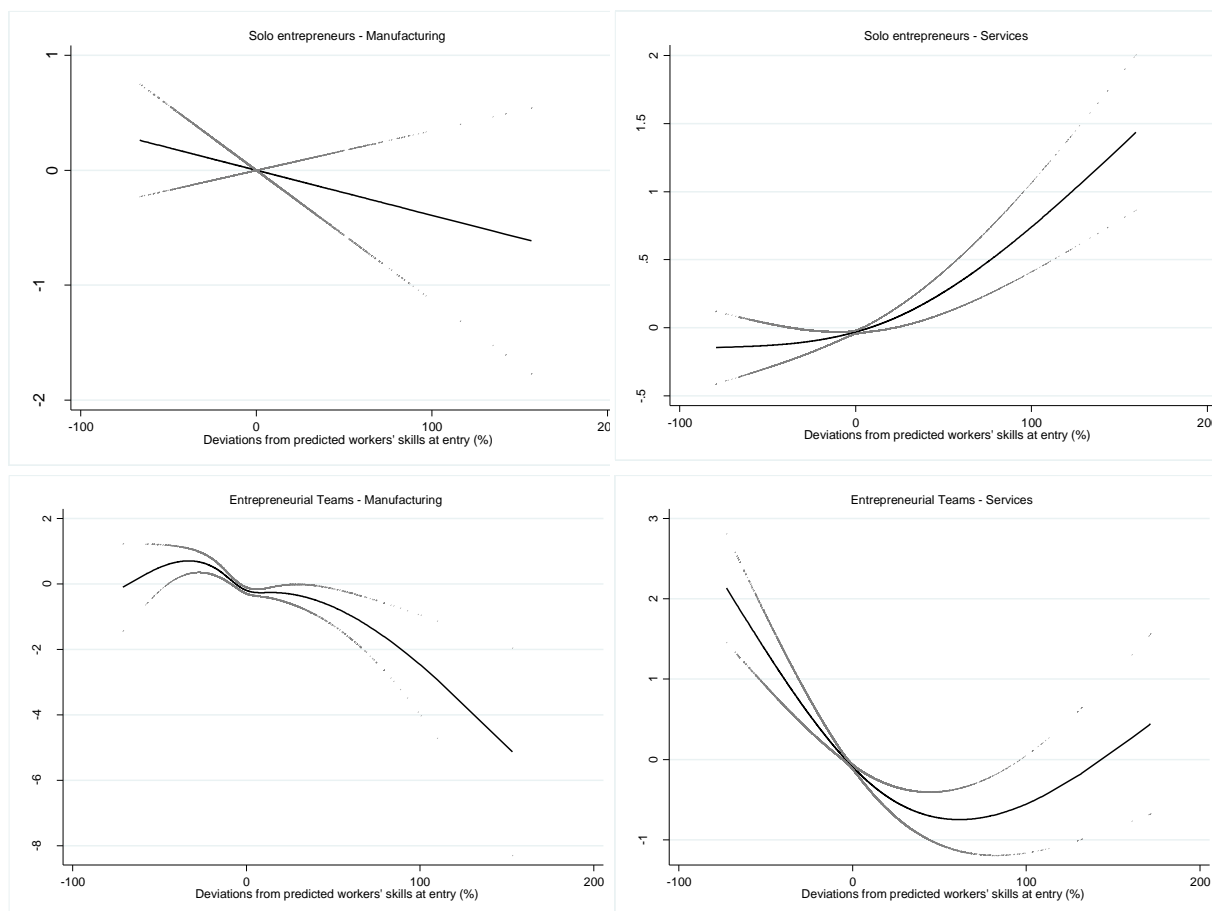


Fig. 4. Estimated relationship between deviations from benchmark workers' skills at entry and log (relative hazards), by groups of startups (grey lines are 95% pointwise confidence intervals)



APPENDIX

Table A.I. Distribution of startups by industries

Manufacturing:	Solo Entrepreneurs	Entrepreneurial Teams
(31) Manufacture of Food, Beverages and Tobacco	10.0%	11.7%
(32) Textile, Wearing Apparel and Leather Industries	37.4%	27.5%
(33) Manufacture of Wood and Wood Products, Including Furniture	7.2%	7.9%
(34) Manufacture of Paper and Paper Products, Printing and Publishing	7.2%	7.6%
(35) Manufacture of Chemicals and Chemical, Petroleum, Coal, Rubber and Plastic Products	2.5%	2.4%
(36) Manufacture of Non-Metallic Mineral Products, except Products of Petroleum and Coal	5.9%	7.0%
(37) Basic Metal Industries	0.4%	0.9%
(38) Manufacture of Fabricated Metal Products, Machinery and Equipment	23.4%	26.9%
(39) Other Manufacturing Industries	6.1%	8.1%
Total	100%	100%
Total number of startups	2,021	1,826
Services:	Solo Entrepreneurs	Entrepreneurial Teams
(61) Wholesale Trade	14.8%	19.2%
(62) Retail Trade	31.7%	28.6%
(63) Restaurants and Hotels	22.1%	20.2%
(71) Transport and Storage	5.3%	5.8%
(72) Communication	0.3%	0.3%
(81) Financial Institutions	0.7%	0.7%
(82) Insurance	0.3%	0.3%
(83) Real State and Business Services	17.6%	17.6%
(91) Public Administration and Defense	3.4%	3.6%
(92) Sanitary and Similar Services	0.3%	0.3%
(93) Social and Related Community Services	0.7%	0.7%
(94) Recreational and Cultural Services	0.7%	0.9%
(95) Personal and Household Services	2.2%	2.0%
Total	100%	100%
Total number of startups	8,180	5,552