Entrepreneurship, the Initial Labor Force, and the Location of New Firms

Cristina Carias, IST and Carnegie Mellon University*

Steven Klepper, Carnegie Mellon University**

Date of this version: February 2010

Print date: 20 September, 2010

* Department of Social and Decision Sciences, Pittsburgh, PA 15213, email address: cdasilva@andrew.cmu.edu.

** Department of Social and Decision Sciences, Pittsburgh, PA 15213, email address: sk3f@andrew.cmu.edu.

Acknowledgements: We thank Serguey Braguinsky and David Greenstreet for helpful comments.

Entrepreneurship, the Initial Labor Force, and the Location of New Firms

Abstract

A model of the location choice of new firms is developed in which founders locate their firms close to their home region in order to hire workers they know about through their prior employment. The model is tested using a matched employer-employee data set for Portugal. Consistent with the model, new Portuguese firms in the same industry as their founder's prior employer were more likely to locate in their founder's home region, to hire workers from the founder's prior employer and other firms in the same region and industry, to employ them longer, and to perform better than other new firms.

[Key words: Location, Home Region, Agglomeration, Entry]

Entrepreneurship, the Initial Labor Force, and the Location of New Firms

I. Introduction

Silicon Valley got its name from the agglomeration of the semiconductor industry in Northern California. A well-known genealogy of semiconductor producers in Silicon Valley compiled by the trade organization SEMI indicates that this was driven by spinoffs of incumbent semiconductor firms. Over a hundred semiconductor firms entered in Silicon Valley between 1955 and 1986, and nearly all of them were founded by employees of semiconductor firms that themselves were located in Silicon Valley. Klepper [2007] proposed that spinoffs of indigenous firms played a similar role in the historical agglomeration of the automobile industry around Detroit, MI. Over 50 spinoffs of automobile firms entered in the Detroit area during the first 30 years of the industry, many of which became leaders of the industry, and nearly all of them were founded by employees of automobile firms that were also located in the Detroit area. Buenstorf and Klepper [2009a] and Berchicci et al. [2010] report a similar tendency for spinoffs in the historical tire and modern disk drive industries respectively to locate close to their geographic roots. More broadly, studies of all the start-up entrants in Portugal (Figueiredo et al. [2002]) and Denmark (Dahl and Sorenson [2008]) in recent years reveal a similar tendency of all types of new firms to locate close to where their founders previously worked and resided.

These findings raise numerous questions. Entrepreneurs have long been thought of as "foot-loose" when choosing where to locate (Pflueger and Suedekum [2008]). Indeed, in modern theories of geography, they are drawn to regions that contain more activity in their industry and overall, reflecting the influence of agglomeration economies (Rosenthal and Strange [2004]). However, the patterns reported above suggest that entrepreneurs do not stray far from where they previously worked and/or lived, which we call their home region. Why? Is this tendency to locate close to the entrepreneur's home region more true of spinoffs than other *de novo* entrants whose founders did not previously work in their chosen industry, and if so, why? If new firms tend to locate close to their entrepreneur's geographic roots, what role do agglomeration economies play in influencing their location?

The main purpose of this paper is to explore these questions using a matched employer-employee data for Portugal that provides information about all new firms and their employees, including their employee-founders. The central idea of the paper is that the location of new firms is heavily influenced by knowledge that founders have about prospective hires based on their prior work experience. A natural source of employees for new firms is their founder's prior employer. We also expect that in their prior employment founders interacted with employees of nearby firms in the same industry, enabling them to identify yet other promising hires. A new firm will be more likely to be able to hire the founder's old colleagues and nearby employees in the founder's prior industry if it locates close to them (Dahl and Sorenson [2010])—i.e., close to its founder's home region. If a new firm also enters the same industry as its founder's prior employer, then these prospective hires will not have to change industries, which should make them more productive hires. Therefore, if firms locate close to their founder's home region to exploit knowledge about prospective hires, they should be especially likely to do so if they enter same industry as their founder's prior employer. This forms the basis for a key test of our theory.

The localized knowledge founders possess might be expected to transcend human capital. Founders may have connections to family and friends (Dahl and Sorenson [2008]) and to sources of capital (Michelacci and Silva [2007]) that could also induce them to locate close to their home region. We attempt to isolate the role that human capital knowledge has on the location of new firms by developing a simple model to draw out of our theory additional implications regarding the location of new firms and the types of workers they hire, the longevity of their hires, and the performance of the firms. Consistent with the model, new Portuguese firms that locate in the same industry as their founder's prior employer are more likely to locate in their founder's home region, to hire workers from the prior employer and other firms in the home region of their founders, to employ these workers longer, and to perform better than other new firms.

The implications of the theory are potentially far reaching. With new firms largely locating close to their founders' geographic roots, particularly the better performing firms, the key determinant of the number of entrants in a region is its supply of potential entrants, what Carlton [1983] calls its birth potential. Recent studies suggest that a

region's birth potential is shaped by its incumbent producers (Klepper [2007, 2010], Buenstorf and Klepper [2009a, 2009b]). This can lead to many of the patterns we often associate with agglomeration economies—entrants locating in agglomerated areas and performing better in such areas—even with agglomeration economies not influencing either firm location or performance. Such forces open up a whole new set of questions about the formation and growth of industry agglomerations (Klepper [2010]). Furthermore, our findings suggest that the early survival of new firms is critically shaped by their initial hires, especially hires from their founder's prior employer. Based on the turnover of personnel versus products and market focus of new firms that receive venture capital and eventually go public, Kaplan et al. [2009] infer that it is primarily ideas and not people that distinguish new firms. Our findings suggest that the match between a new firm's ideas and its initial hires may also play a critical role in its early performance.

The paper is organized as follows. In Section II we lay out our model and derive various implications from it. In Section III we describe the data and the variables we construct to test the model. In Section IV we report estimates of various econometric models of location choice, employee hiring, employee longevity, and firm performance. In Section V we discuss the implications of our findings and offer concluding remarks.

II. Model

We consider a new firm founded by an employee of an incumbent firm. The location of the incumbent firm—i.e., where its founder previously worked—is referred to as the home region of the founder and also the new firm. The industry of the incumbent firm—i.e., the industry in which the founder previously worked—is referred to as the home industry of the founder and the new firm. We assume that employees receive ideas for new firms in part based on their employment experience. If the ideas are good enough, they start a new firm. The idea dictates the industry in which a new firm enters, which is referred to as the firm's chosen industry. It does not dictate where a new firm locates, which is referred to as the firm's chosen region. This choice is made by the founder of the firm.

Firms and workers are heterogeneous, limiting the number of workers that are suitable hires for any given firm. A new firm can identify suitable hires by advertising

positions and interviewing job candidates, but otherwise cannot judge the productivity of potential hires without additional information. We assume that founders learn about workers in their prior employer, which enables them to project the productivity of these workers in their new firm. We assume that in their prior employment they also interact with employees in other firms, which also enables them to project the productivity of these workers in their new firm. For simplicity, we assume that all such interactions are with firms in their home industry and region, reflecting that individuals are most likely to interact with workers in the same industry and region as their employer. Therefore, all the potential hires for which founders have distinctive information about their productivity are from their home region and industry.

We structure the model so this information is valuable only if a new firm enters in its home region, and it is most valuable if it enters in its home industry. The former assumption makes firms more likely to enter in their home region than all other regions. The latter assumption makes the home region especially attractive to firms that enter their home industry. We show how these assumptions also lead to various predictions about how the firm's location conditions the types of workers it hires, the rate of turnover of its workers, and its performance.

Setup of the Model

Let there be $j = 1, 2, \ldots, J$ regions where a potential entrant could locate. Let π_{j0} denote the expected discounted value of the potential entrant at time 0, the time of entry, if it located in region j. We assume that π_{j0} is the sum of two components: L_j, which is the expected discounted surplus the potential entrant earns from the labor it hires if it enters and locates in region j, and ε_j, which is a composite term representing all other factors affecting the potential entrant's value if it locates in region j. The ε_j, $j = 1, 2, \ldots, J$, are assumed to be independent draws on a random variable ε with support $[-d,d]$, where $d > \max_j L_j$ Potential entrants enter in the region that maximizes their expected discounted value subject to their expected discounted value being nonnegative.

Firms require a work force of H workers and thus must hire H workers when they are formed. Workers are heterogeneous in terms of their abilities and only those whose abilities match the needs of a new firm will be suitable hires. When firms are formed, all

potential hires work at a wage w equal to the value of their marginal product V. They will not work for a new firm unless they are paid a wage $w_n = w + u + rR$, where $u > 0$ is the premium required to work for a new firm with unknown prospects, $r > 0$ is a relocation premium required to get a worker to move to a new region, and R is an indicator variable equal to 1 if the new firm is located in a different region than the worker's previous employer. The value of the worker's marginal product at a new firm, V_n, equals $V - iI + \theta^*$, where $i > 0$ is the decrease in a worker's productivity if the worker moves to a different industry, I is an indicator variable equal to 1 if the firm is in a different industry than the worker's prior employer, and θ^* reflects the worker's productivity at the new firm. For simplicity, we assume that θ^* equals 0 or θ, where $\theta > u + i$. We call workers with $\theta^* = \theta$ high productivity workers (at the new firm).

Unless a worker is high productivity, a firm will not want to employ the worker as it would have to pay him a wage that exceeds the value of his marginal product. A worker might have higher productivity at a new firm than his prior employer for various reasons. The worker's prior job may not fully exploit his abilities, which can happen if the firm does not have an open position to which to promote a talented worker. Another possibility is that the worker's abilities might be better matched to the needs of the new firm than his prior employer. Suppose that a firm can identify a limited number of workers in any region whose abilities match its needs. Let δ denote the probability that $\theta^* = \theta$ for such workers. Let N_j denote the number of these workers in the firm's chosen industry in region j. To abstract from the influence of agglomeration economies, we assume that $N_j > H$ for all j, so the firm could staff all of its needs from workers in its chosen industry regardless of the region in which it locates.

We assume that some of the workers at the prior employer of the firm's founder have abilities that match the firm's needs and the founder has knowledge about them that suggests they would be more likely than δ to be high productivity workers at the founder's new firm. Let there be C such old colleagues (C stands for colleagues) with probability $\alpha > \delta$ of having productivity $\theta^* = \theta$. Suppose that through prior interactions founders have similar knowledge about some of the workers at other firms in their home industry and region. Let there be IR such workers (IR stands for industry and region) with

probability $\beta > \delta$ of having productivity $\theta^* = \theta$. We assume that founders have better knowledge about their old colleagues than workers in other firms and thus that $\beta < \alpha$.

We assume that in each period that a firm operates, there is a probability h of learning the value of θ^* for any worker whose θ^* has not already been learned. If $\theta^* = 0$ then the worker's wage exceeds the value of his marginal product and the firm replaces the worker. Firms also are affected by random developments in the external environment that affect their value in each period. We abstract from changes over time in the value of firms resulting from the replacement of workers[1] and assume that the longevity of firms is determined by the random developments in the external environment. A firm begins with an expected discounted value of π_{j0} and in each period t receives an additive shock to its expected discounted value denoted as ω_t, where for each firm the ω_t are independent draws on a random variable ω that can take on negative values. A firm exits when its expected discounted value falls below 0.

Employment Choices

The expected surplus earned by a firm from hiring a worker (with the requisite abilities) in the period in which the worker is hired is

$$E(V_n - w_n) = E[w - iI + \theta^* - (w + u + rR)] = E(\theta^*) - u - iI - rR.$$

If a firm enters in its home region and industry, it will hire first the C old colleagues, as for these workers $E(\theta^*)$ is greatest and $I = R = 0$. Next it will hire the IR workers from its home industry and region, as these workers have the next highest value of $E(\theta^*)$ and $I = R = 0$. If it needs to hire additional workers, it will hire them from its chosen industry and region (which is the same as its home industry and region) as all other workers have the same value of $E(\theta^*)$ and $I = R = 0$ for these workers. If the firm enters its home region but not its home industry, for simplicity we assume the same strategy is profitable—i.e., the expected value of the knowledge about worker productivity more than offsets the lower productivity of these workers from switching industries, which requires that $(\beta - \delta)\theta > i$.

[1] But firms factor into their expected discounted value at the time of entry and thus into their decision to enter expected increases in their labor surplus over time due to the replacement of workers.

Alternatively, the relocation premium r is assumed to be sufficiently large that $(\alpha - \delta)\theta < r - i$, which insures that if a firm did not enter in its home region then it will hire all its workers from its chosen industry and region. These assumptions insure that regardless of the industry a firm enters, it exploits its regional knowledge about labor if, and only if, it enters in its home region.

Implications

Most of the implications of the model follow straightforwardly. Consider first the expected discounted surplus of a firm from its hires. Let this surplus be denoted as S_{IR}, S_R, S_I, and S respectively for firms that enter in their home industry and home region, in their home region but not home industry, in their home industry but not home region, and in neither their home industry nor home region. For the first two cases, a firm first hires the C old colleagues and then the IR workers from its home industry and region that it knows, followed by workers in its home region and chosen industry (which in the former case is its home industry). This strategy yields a greater expected surplus in each period for firms that enter their home industry (as well as their home region) as the workers they know about do not suffer any reduction in their productivity from switching industries. A firm that does not enter in its home region hires (unknown) workers from its chosen region and industry whether it enters its home industry or not and earns the same surplus whether it enters its home industry or not. This surplus is less than if the firm located in its home region and exploited the knowledge it had about old colleagues and workers from its home industry and region. Therefore, $S_{IR} > S_R > S_I = S$.

A firm chooses the location that maximizes its expected discounted value, which means that the greater its expected discounted surplus from entering in a region then the greater the probability of entering there. Consequently, firms will be more likely to enter in their home region than elsewhere. Furthermore, if a firm enters its home industry, then the difference between its expected discounted surplus if it locates in its home region versus elsewhere is $S_{IR} - S_I$. Alternatively, if it does not enter its home industry, then the difference between its expected discounted surplus if it locates in its home region versus elsewhere is $S_R - S < S_{IR} - S_I$. This implies that the probability of a firm locating in its home region is greater for firms that enter in their home industry. Collecting results:

Proposition 1: The probability of a firm locating in its home region is greater than any other region and is greater for firms that enter in their home industry.

Consider next the hiring choices of firms. The only firms that hire the *C* old colleagues are those that enter in their home region and they hire these workers first before any other group of workers. Therefore:

Proposition 2: The probability of hiring old colleagues is initially greater for firms that enter in their home region and industry and home region but not home industry and subsequently declines for both groups of firms and eventually is equal (to 0) for all firms.

Firms that enter in their home region are also the only ones that hire workers from their home industry and region. Once they hire the *C* old colleagues and *IR* workers from their home industry and region that are known to their founder, all their subsequent hires will be from their home region and chosen industry. Since the chosen industry is the same as the home industry for firms that enter their home industry, excluding old colleagues firms that enter in their home region and industry in every period hire all their workers from their home industry and region. Alternatively, for firms that enter their home region but not their home industry, some of their hires may be from their home region and chosen, but not home, industry. If $C + IR < H$ this will be true for some of their initial hires, but otherwise it may occur for some of their later hires as they replace initial hires that prove to be low productivity. Consequently, for these firms over time the share of their hires from the home industry and region will decline. Since old colleagues are hired before workers from the home industry and region, the decline will be less severe than for old colleagues. Therefore:

Proposition 3: Excluding old colleagues, the probability of hiring workers from the firm's home industry and region is initially greater for firms that enter in their home region and industry and in their home region but not home industry, with the former probability greater than or equal to the latter. Subsequently it will remain the same for firms that enter in their home region and industry but decline (to zero) for the firms that enter in their home region but not home industry, although at a slower rate than the probability of hiring old colleagues.

Next consider the hazard of exit of workers from their employers. For any group of workers, after *t* periods the probability of their productivity not being learned is $(1 - h)^t$

and the probability of them being confirmed as high productivity (and thus still with the firm) is $p[1 - (1 - h)^t]$, where $p \equiv \text{prob}(\theta^* = \theta)$. Therefore, their hazard of exit at "age" $t + 1$ (i.e., in period $t + 1$ at the firm) is $(1 - h)^t h(1 - p)/\{p[1 - (1 - h)^t] + (1 - h)^t\}$, which is a decreasing function of p. All old colleagues have $p = \alpha$, the initial (*IR*) workers from the firm's home region and industry (hired by firms that locate in their home region) have $p = \beta < \alpha$ (subsequent hires from the home industry and region have $p = \delta < \beta$), and all other workers have $p = \delta$. Therefore, it follows that:

Proposition 4: For all workers the hazard of exit at each age (at the firm) is lowest for old colleagues and next lowest for the initial workers hired from the firm's home region and industry.

Last, consider the performance of firms that enter in their home region versus those that enter elsewhere. For each region j, the distribution of profits for firms entering there has support $[0, L_j + d]$. Therefore, the maximum profit firms can attain is greater for firms that enter in their home region, particularly those that also enter in their home industry. We cannot test this prediction directly given the limited data available, but we can observe a measure of firm performance, longevity, that would be expected to reflect the advantages of locating in a firm's home region. Under simplifying assumptions concerning the shocks to firm profits in each period, it can be shown that the hazard of firm exit in each period will indeed be lower for firms that enter in their home region, which we refer to as region h.

Specifically, suppose that the shocks to the firm's profits ω_t can take on only two possible values: $-g < 0$ and 0, with probabilities r and $(1 - r)$ respectively, where $L_h + d > g > \max_{j \neq h} L_j + d$. Then in each period t, firms that do not enter in their home region and that have not yet exited will have a probability of exit of r. In contrast, among firms that enter in their home region and have not yet exited, some fraction f will have profits greater than g and will not be at risk of exit, so their probability of exit will be less than r. Consequently, in every period the hazard of exit will be lower for firms that enter in their home region, and especially so for firms that also enter in their home industry (the fraction f of these firms not at risk of exit in each period will be greater than for firms that enter in their home region and not their home industry).

To analyze more general cases concerning the distributions of ω and ε, we resorted to numerical methods. We set $J = 3$ and allowed ω and ε to be uniformly distributed over $[-g, g]$ and $[-d,d]$ respectively. Alternatively, we allowed ω and ε to be normally distributed with mean 0 and standard deviation g and d respectively. We varied g, d, and also the additional surplus from entering in the firm's home region, which we denote as s. Our simulations, which involved 500,000 firms, confirmed that in each period the fraction of survivors was consistently greater and the hazard of exit consistently lower for firms that located in their home region. Furthermore, these patterns were more pronounced the greater s relative to d, as would be expected. This suggests:

Proposition 5: The hazard of firm exit will be lowest for firms that enter in their home region and industry and next lowest for firms that enter in their home region but not home industry.

III. Data and Variables

All data are drawn from the "Quadros de Pessoal," which is a matched employer-employee data set for Portugal. Submission of data is mandatory for every Portuguese firm in the economy. Each year firms report their sales, total number of employees, establishments, year of constitution, main industry (5 digit industrial code), initial capital, and share of initial capital that is foreign owned. For each establishment, they report the number of employees, location (the concelho or county where the establishment is located), and main industry (5 digit industrial code). For each worker, annual data are collected on their establishment, age, gender, education (primary, secondary, high school, or college), occupation (5 digit code), hierarchical level (nine categories that were aggregated into three groups: managers, specialized workers, and laborers), year hired, earnings, and hours worked.

A major change occurred in the industrial classification system in 1994 and the industry codes cannot all be matched before and after the change. Consequently, we restrict our focus to entrants in 1996 and later, where entrants are defined as firms whose

first appearance in the dataset matches their declared year of constitution.[2] The theory is couched in terms of the firm's early hires, which we define as its hires in its first three years. The last year for which we have data is 2006, so we consider entrants through 2004 in order to have three years of data on their hires. Information was not available for employees for 2001. Accordingly, we excluded entrants in 2000 and 2001 because we had no data on their first two years of hires.[3]

We excluded foreign owned firms and non-profit organizations (such as associations and cooperatives). We also excluded entrants involved in agriculture, energy distribution, public administration, and schools and social services whose location choices were constrained by the nature of their operations. These exclusions reduced the sample by 18%. We excluded the entrants with more than one establishment because of their multiple locations, which reduced the sample by 7%. We included only entrants with at least one employee in their first year (that was not an owner), which reduced the sample by 33%. The founders of the firm were defined as employees that were also listed as owners in the firm's first year or the second year if there were no owner-employees listed in the first year. Accordingly, we restricted the sample to entrants with at least one owner-employee in the firm's first or second year, which reduced the sample by 36%. We also restricted the sample of entrants to firms with at least one owner-employee with a known background, which reduced the sample by 47%. In total, we ended up with 10,236 entrants[4] that hired 27,282 workers in their first year, 8,851 workers in their second year, and 6,235 workers in their third year.[5]

---

[2] Tracing the history of founders and employees required us to search earlier years. The industry codes in 1995 and 1996 are the same but this was not true for some industry codes before 1995, which required the use of an algorithm (based on how the majority of firms changed industry codes from 1994 to 1995) to match industry codes. To minimize errors while preserving observations we included only entrants from 1996 onwards.

[3] We included entrants in 1999 even though we had no information on their third-year hires because the number of hires in the third year is small relative to the prior two years and we wanted to keep the sample of entrants as large as possible.

[4] The total number of new entrants in our sample period was 88,981, so our criteria resulted in retaining about 11% of the firms. We compared our sample with the full sample. The distribution of entrants by location and industry was similar for the two samples. The main difference was that average number of initial members of the firm, including owner-employees, was greater (by one) than the full sample, reflecting the requirement that firms in our sample had at least one employee in their first year.

[5] This is less than the sample of all hires (for instance, in the first year, firms hired 32,968 workers whereas our sample in the first year is composed of 27,282 workers). Some workers had to be removed from the sample because the firm failed to input the workers' correct social security number.

We determined the work history of every founder in the four years before establishing his new firm. The firm's home region is defined as the county (in Portuguese, concelho) of the establishment where its founder most recently worked. Figure 1 is a map of Portugal that distinguishes its counties, which are roughly about a quarter of the size of U.S. counties.[6] The firm's home industry is defined as the 4 digit industry of the firm where its founder most recently worked. If a firm had more than one founder then all the different counties of the founders' prior establishments were defined as a home region of the firm and all the different 4 digit industries of the founders' prior firms were defined as a home industry of the firm. The chosen region of the firm is defined as the county of the firm in its first year and its chosen industry is defined as its 4 digit industry in its first year. The founder's tenure at his prior employer was defined as the number of years the founder worked at the employer and the founder's tenure in his home region was defined as the total number of years the founder previously worked at establishments in the same county as his previous employer.[7] In the case of firms with more than one founder, these variables are computed as averages for all the founders. We also identified the founder's prior position (manager, specialized worker, laborer) according to the last position he held at his prior employer.

We traced the background of every employee in the four years before joining a new firm. We focused on four (non-exhaustive) categories of workers: old colleagues, workers from the firm's home region and industry, workers from the firm's chosen industry and region, and workers with an unknown background. Old colleagues are employees whose most recent job was at the founder's prior employer. Workers from the firm's home industry and region are employees whose most recent job was at an establishment in the same county as the founder's prior establishment and in the same 4 digit industry as the founder's prior employer. Workers from the firm's chosen region and industry are employees whose most recent job was at an establishment in the firm's chosen county and 4 digit industry; if the firm's chosen industry and region were the

[6] Currently, Portugal (excluding islands) is subdivided into 278 counties. Three counties were introduced in 1998 and correspond to subdivisions of previously existing counties. Throughout the analysis we merged the new counties with the older ones and so considered only the 275 counties that existed as of 1996.

[7] For every founder, we traced his work history from 1987 (the first year of the dataset) onwards to determine tenure in the home region.

same as its home industry and region, these employees are classified both as workers from the firm's home region and industry and from the firm's chosen industry and region. Employees were classified as unknown if they did not show up in the data set in the four years before being hired by the firm, which includes individuals that previously were unemployed, students, or worked in the public sector.

Table 1 provides various descriptive statistics about the 10,237 firms in the sample. In terms of location, 66% of the firms located in their home county, 8% located within 10 kilometers (km) of their home county (but not in it), 10% located between 10 and 20km of their home county, 7% located between 20 and 40km of their home county, and the remaining 8% located elsewhere.[8] Half of the firms located in just three districts: Porto (20%), Lisboa (19%), and Braga (12%). In terms of the relationship between the firm's home and chosen industry, 44% of the firms located in their home (4 digit) industry, 4% entered the same 3 but not 4 digit industry as the firm's home industry, 11% entered the same 2 but not 3 or 4 industry as the firm's home industry, and the remaining 40% entered other industries. Regarding the entrants, 35% had more than one founder, 36% had founders that were managers in their prior establishment, 50% had founders that worked four or more years in their prior establishment, 50% had founders that worked six or more years in their home county, and 50% had two or more employees in their first year.

Table 2 provides descriptive statistics about the 42,368 employees in the sample. For employees hired in the firm's first, second, and third year it reports the fraction that were old colleagues, workers from the firm's home industry and region, workers from the firm's chosen industry and region, and workers of unknown background. In year 1, 34% of the workers hired were old colleagues, 5% were from the firm's home industry and region, 9% were from the firm's chosen industry and region (if those also from the firm's

[8] In Portugal, the time to commute increases discontinuously with the number of county borders one needs to cross. This is true both for commuters using public transportation and for commuters that use their own vehicle. Public transportation systems are usually county bounded and integration mechanisms are time consuming. On the other hand, time consuming traffic lines are usual in the most densely populated metropolitan areas. For instance, a 15km travel from a county nearby Lisboa to the county of Lisboa may take 15 minutes on a Sunday morning and one hour or more in rush hour, while inside Lisboa public transportation systems such as the subway ensure those discrepancies do not exist. This explains the preference of firms to stay so close to their home county.

home industry and region are excluded, this drops to 4%), and 32% had an unknown background. In year 2 the percentage of old colleagues drops sharply from 34% to 8% and then drops further to 4% in year 3. In contrast, the percentage of workers from the home industry and region rises from 5% to 6% in year 2 and 7% in year 3 while the percentage of workers from the chosen industry and region stays steady at 9% in all three years and the percentage of workers with an unknown background rises from 32% to 43% in years 2 and 3.

IV. Econometric Specifications and Estimates

We first test the predictions of the model concerning the location of firms, followed by the predictions concerning the hires of firms, the hazard of exit of their workers, and the hazard of exit of the firms.

A. Location Choices

We test the location predictions of the model using a conditional logit model:

$$p_{ij} = \frac{exp(\mathbf{x}_{ij}\boldsymbol{\beta})}{\sum_{m=0}^{J} exp(\mathbf{x}_{im}\boldsymbol{\beta})}$$

where p_{ij} is the probability of firm i locating in county j, $\mathbf{x}_{ij}$ is a vector of explanatory variables pertaining to firm i and county j, and $\boldsymbol{\beta}$ is a vector of coefficients to be estimated. Proposition 1 predicts that the probability of locating in a region is greater if the region is the firm's home region. To test this prediction, we include two variables in $\mathbf{x}_{ij}$, *Home*, which equals one if county j is the firm's home county and 0 otherwise, and *Dist*, which is the number of kilometers between the center of county j and the center of the firm's home county. Proposition 1 implies that the coefficient of these variables, denoted as β_{Home} and β_{Dist}, should be positive and negative respectively. Proposition 1 also predicts that the probability of locating in the home region will be greater for firms that enter their home industry. We test this by interacting *Home* and *Dist* with a variable *4I*, which equals 1 if the firm entered its home (4 digit) industry and 0 otherwise. Proposition 1 implies that $\beta_{Homex4I} > 0$ and $\beta_{Distx4I} < 0$.

The estimates for this specification, which is denoted as Model 1, are presented in Table 3. Consistent with Proposition 1, $\hat{\beta}_{Home}$ (the hat denotes an estimate) and $\hat{\beta}_{Homex4I}$ are positive and significant and $\hat{\beta}_{Dist}$ and $\hat{\beta}_{Distx4I}$ are negative and significant. Thus, all

else equal firms are more likely to locate in their home county and closer to their home county if not in their home county, with this being especially so for firms that enter their home (4 digit) industry.

We probe the importance of the industry entered by adding to Model 1 *Home* and *Dist* each interacted with two other dummies, denoted as *3I* and *2I*, which are equal to 1 for firms that did not enter their home 4 digit industry but entered in the same 3 digit category and 2 digit category respectively as their home 4 digit industry. Based on an ordering of "closeness" of industries using the 4 digit industry code system, Proposition 1 implies that $\beta_{Homex4I} > \beta_{Homex3I} > \beta_{Homex2I} > 0$ and $\beta_{Distx4I} < \beta_{Distx3I} < \beta_{Distx2I} < 0$. The coefficient estimates of this specification, denoted as Model 2 in Table 3, are all significant and have the predicted signs and ordering.[9]

We next consider other factors that may bear on the location of firms, especially factors related to the closeness between the firm's chosen and home industry. Numerous studies of new industries have found that among new entrants, those founded by individuals that previously worked for firms in the same industry outperformed other startup entrants (Klepper and Thompson [2010]). If better firms are also more likely to locate close to their home region, which could hold for a number of reasons (cf. Berchicci et al. [2010]), then we also need to control for factors related to the performance of firms to test reliably Proposition 1.

We construct three variables that might be expected to affect the performance of firms. The first, *Firmtenure*, is the log of the average number of years the firm's founders worked at their prior employers. It is a measure of the extent of the knowledge founders acquired at their prior employers and is measured in log form to allow for diminishing returns. The second, *Highlevel*, is a 1-0 dummy variable equal to 1 for firms with one or more founders that worked as managers in their prior employer. The third, *Multiplefounders*, is a 1-0 dummy variable equal to 1 for firms with more than one founder. Each of these variables is positively correlated with the firm's initial number of

[9] In terms of significant differences, $\hat{\beta}_{Homex4I}$ and $\hat{\beta}_{Homex2I}$ are significantly different and $\hat{\beta}_{Distx4I}$ is significantly different than $\hat{\beta}_{Distx3I}$ and $\hat{\beta}_{Distx2I}$.

employees, which numerous studies have found is a proxy for the longevity of the firm.[10] We also construct a fourth variable, denoted as *Regionaltenure*, which is measured as the log of the average number of years the firm's founders worked in their home county. It is both a measure of knowledge pertaining to the region and thus a potential cause of firm performance and also a proxy for local ties to family, friends, and others that might induce firms to locate close to their founder's home region.

We interact each of these four variables with *Home* and *Distance* and add these interactions to Model 2, which defines Model 3 in Table 3. Regarding the added interactions, the only interaction with *Home* that is significant is for *Highlevel*, which has a positive coefficient estimate, implying that firms with high-level founders were more likely to locate in their home county. In terms of the interactions with *Distance*, the coefficient estimates of *Highlevel*, *Multiplefounders*, and *Regionaltenure* are negative and significant, implying that firms with high-level founders, multiple founders, and founders with longer tenure in their home county were more likely to locate closer to their home county if they did not enter in their home county. All these estimates are consistent with better firms being more likely to locate closer to their home county. The only exception to this pattern is the coefficient estimate of *Firmtenure* interacted with *Distance*, which is positive and significant, which implies that firms with founders with more tenure at their prior employer were more likely to locate further away from their home county if they did not locate in their home county. Most important, the addition of these variables has little effect on the other coefficient estimates, which continue to support Proposition 1.

As noted in the introduction, modern geography theories posit that firms are attracted to regions with more workers in their industry and overall based on the idea that locating in such regions improves a firm's performance. Recent studies have found, however, that the attractive force of such regions is stronger when the region is not the

[10] Consistent with other studies, in our sample firm longevity was positively and significantly related to the initial number of employees of the firm. However, when *Firmtenure*, *Highlevel*, and *Multiplefounders* were also allowed to affect firm longevity, the effect of the initial number of employees on the firm hazard was no longer significant while the effects of *Firmtenure*, *Highlevel*, and *Multiplefounders* were. Accordingly, we did not include the initial number of employees of the firm in our analyses, although its inclusion has little effect on our estimates.

founder's home region (Figueiredo at al. [2002], Buenstorf and Klepper [2009b]). To explore the role of regional characteristics and to test the robustness of the findings to such factors, we include in Model 4 the density of workers in county *j* in the firm's chosen industry, denoted as *Workerdensity*, and the density of all workers in county *j*, denoted as *Populationdensity*. Following Figueiredo at al. [2002], we interact each variable with *Home* and (1 – *Home*) in Model 4 to allow these variables to have different effects for the firm's home county and elsewhere.

The coefficient estimates of *Workerdensity x (1 – Home)* and *Populationdensity x (1 – Home)* are both positive and significant whereas the coefficient estimate of *Workerdensity x Home* is small and insignificant and the coefficient estimate of *Populationdensity x Home* is actually negative and significant. Consistent with Figueiredo at al. [2002], firms are more likely to enter in counties outside of their home county that have a greater density of workers in their industry and overall, but they are not more likely to enter in their home county if it has a greater density of workers in their industry and less likely to enter in their home county if it is more densely populated. The other coefficient estimates are largely unaffected by the inclusion of these density variables and thus continue to support the theory.

We estimate one last specification for the conditional logit, Model 5, in which we include dummy variables for each of the 275 counties. This allows for unobservables, such as local amenities and proximity to a port or other kind of transportation, to affect the location of firms. The inclusion of the fixed effects has little effect on the main coefficient estimates related to Proposition 1. In terms of the four variables representing firm performance, the coefficient estimates of both *Multiplefounders* and *Regionaltenure* interacted with *Home* are now positive and significant, implying that both increase the probability of firms locating in their home county (as well as closer to their home county if not in their home county, as before). The other main change is the coefficient estimate of *Populationdensity* for the home county, which is about half as large and no longer significant. Thus, firms are no more or less likely to enter in their home region based on characteristics of the region. Rather, consistent with Proposition 1 the decision to locate in the home region appears to be driven mainly by whether a firm enters its home industry or an industry close to its home industry. In contrast, if the firm does not locate

in its home region, then it is attracted to regions with more workers in its chosen industry and overall.

To put the conditional logit estimates in perspective, Table 4 reports the fraction of firms that entered in their home county and 1-10, 10-20, 20-30, 30-40, and greater than 40 kilometers from their home county according to the industry the firm entered. Overall, 66% of firms entered in their home county, another 8% entered within 10 kilometers of their home county, and 8% entered over 40 kilometers away from their home county. The differences in these percentages according to the industry the firm entered are striking, particularly the percentage locating in their home county. For firms that entered their home (4 digit) industry, 77.4% entered their home county versus 69.3%, 66.5%, and 53.6% for firms that entered the same 3 digit, 2 digit and other industries relative to their home industry. Similarly, among the firms that did not enter their home county, generally they entered closer to their home county the closer the industry they entered relative to their home industry.

The main effect of the industry entered appears to be on whether firms located in their home county. Accordingly, we also estimated a simple probit model for the probability of a firm locating in its home county. The explanatory variables were *4I*, *3I*, and *2I* as well as the four firm controls, *Workerdensity* in the firm's home county, and dummies for each of the 275 counties. Consistent with the conditional logit estimates, all the coefficient estimates except for *Workerdensity* are significant. The coefficient estimates of *4I*, *3I*, and *2I* are all positive and descending in magnitude, as expected, the coefficient estimates of *Highlevel*, *Multiplefounders*, and *Regionaltenure* are all positive, and the coefficient estimate of *Firmtenure* is negative.

B. Hires

Next we test the predictions of the theory regarding the workers firms hired by estimating a series of probit models. We first estimate a probit model for the probability of the firm's hires in year 1 being old colleagues. Standard errors for this and all subsequent probits have been computed by clustering observations for all the hires of each firm.

Proposition 2 predicts that firms that locate in their home region are more likely to hire old colleagues. In our theoretical model we assumed this holds independent of the industry entered, which is also reflected in Proposition 2. To test these predictions, we include as explanatory variables three 1-0 dummies, denoted as *H4IR*, *HR*, and *H4I*, which equal 1 respectively for firms that entered in their home county and home (4 digit) industry, their home county but not their home (4 digit) industry, and in their home (4 digit) industry but not their home county, with the omitted group firms that did not enter their home (4 digit) industry or home county. If firms that locate in their home county are more likely to hire old colleagues then $\beta_{H4IR} > \beta_{H4I}$ and $\beta_{HR} > 0$. If this holds independent of the industry entered, then $\beta_{H4IR} = \beta_{HR}$ and $\beta_{H4I} = 0$.

The coefficient estimates and standard errors for these variables are reported under Model 1 in Table 5. They are all positive and significant. Consistent with Proposition 2, $\hat{\beta}_{H4IR}$ is significantly greater than $\hat{\beta}_{H4I}$ and $\hat{\beta}_{HR}$ is positive and significant, indicating that firms that entered in their home county were more likely to hire old colleagues whether they entered their home (4 digit) industry or not. At the same time, $\hat{\beta}_{H4IR}$ is significantly greater than $\hat{\beta}_{HR}$ and $\hat{\beta}_{H4I}$ is positive and significant, which indicates that industry is also relevant, in contrast to our assumption in the model.

We probe these regional and industry effects further by distinguishing in Model 2 firms that entered the same 3 (but not 4) and same 2 (but not 3 or 4) digit industry as their home (4 digit) industry. This defines eight dummy variables, which are denoted as *H4IR*, *H3IR*, *H2IR*, *HOIR*, *H4I*, *H3I*, and *H2I*, where the abbreviations reflect the overlap between the industry entered and the home industry (the same 4, 3, 2 digit or Other industries) and region (*R* if the firm entered its home county). All the coefficient estimates are positive and significant. More important, for each type of industry (4, 3, 2 digit or other industries), the coefficient estimate is always significantly larger for firms entering in their home county than elsewhere, consistent with firms locating in their home region being more likely to hire old colleagues. Industry continues to be important, as reflected in the decreasing magnitudes of $\hat{\beta}_{H4IR}$, $\hat{\beta}_{H3IR}$, $\hat{\beta}_{H2IR}$, and $\hat{\beta}_{HOIR}$ and the (near) decreasing magnitudes of $\hat{\beta}_{H4I}$, $\hat{\beta}_{H3I}$, and $\hat{\beta}_{H2I}$. However, the estimates are such that firms that entered in their home county and at least the same 2 digit industry as their home (4 digit) industry were significantly more likely to hire old colleagues than all other firms

that did not enter their home county, including those that entered in their home (4 digit) industry. Thus, while industry is relevant, firm location is still paramount in terms of hiring old colleagues.

We next estimate a model with the same variables as in Model 1 plus the control variables for firm performance.[11] Three of these variables, *Firmtenure*, *Highlevel*, and *Multiplefounders*, might also bear on the amount of knowledge founders possessed about their prior colleagues, which would be expected to increase the likelihood of hiring old colleagues. This would not be true of the fourth variable, *Regionaltenure*, which would bear more on knowledge of other workers in the region once *Firmtenure* is controlled. The coefficient estimates of this specification, which are presented in Table 5 under Model 3, are consistent with these conjectures. The coefficient estimates of *Firmtenure*, *Highlevel*, and *Multiplefounders* are positive and significant and the coefficient estimate of *Regionaltenure* is small and insignificant. Most important, the addition of the control variables has little effect on $\hat{\beta}_{H4IR}$, $\hat{\beta}_{HR}$, and $\hat{\beta}_{H4I}$ and thus their support for Proposition 2.[12]

We next add to Model 3 fixed effects for 2 digit industries (36 in total) and 18 regions ("distritos"). This has little effect on the coefficient estimates, which are reported under Model 4 in Table 5. Last, we estimate Model 4 for hires in years 2 and 3, which are denoted as Models 4-2 and 4-3. According to Proposition 2, over time the rate of hiring of old colleagues by firms that entered their home county should decline. Consequently, β_{H4IR} and β_{HR} should decline in year 2 and then decline further in year 3. These predictions are strongly supported. From year 1 to years 2 and 3 $\hat{\beta}_{H4IR}$ and $\hat{\beta}_{HR}$ decline sharply, especially $\hat{\beta}_{H4IR}$, which is not much larger and is no longer significantly different than $\hat{\beta}_{H4I}$ by year 3.

---

[11] The patterns are similar if we maintain the industry distinctions in Model 2, but we report the estimates for the simpler model to facilitate comparisons with the patterns for hires in years 2 and 3 reported below.

[12] We also included a variable equal to the number of employees of the parent firm in the year prior to the new firm's entry to test how the size of the pool of old colleagues influenced the hiring of old colleagues. Curiously, the coefficient estimate of this variable was negative and significant. We added a quadratic term for the number of employees of the parent firm, which had a positive and significant coefficient estimate, indicating that eventually hiring old colleagues increased with the pool to choose from. One interpretation of these findings is that small firms facilitate the kind of interactions among employees that motivate the hiring of old colleagues. The inclusion of the number of employees of the parent firm did not alter the patterns regarding the other coefficient estimates.

Table 6 reports the percentage of hires in years 1, 2, and 3 that were old colleagues, from the firm's home industry and region, from the firm's chosen industry and region, and had an unknown background according to whether the firm entered in its home county and/or home (4 digit) industry. The patterns regarding old colleagues help put the probit estimates for the old colleagues in perspective. Early on firms that entered in their home county were more likely to hire old colleagues—in year 1 50% and 30% of the hires of firms that entered in their home county and in their home (4 digit) or another industry respectively were old colleagues versus 29% and 12% of their counterparts that entered elsewhere. In contrast, in the next two years less than 12% of the hires of all four groups of firms were old colleagues. Firms hired nearly twice as many workers in their first year than the next two combined, so the first-year differences in the rate of hiring of old colleagues contributed to pronounced differences in the composition of the early labor force of new firms.

Consider next workers from the firm's home industry and region. The model's predictions regarding these workers, as reflected in Proposition 3, are similar to those for old colleagues in Proposition 2. The main difference is that in contrast to old colleagues, only β_{HR} and not β_{H4IR} is expected to decline over time (firms that entered in the home region and industry continue to hire workers from their home industry and region over time). Furthermore, the theory predicts that old colleagues will be hired before workers from the home industry and region and hence the predicted decline in β_{HR} will be less sharp than for old colleagues. We test these predictions by estimating probits for the probability of all hires except old colleagues being workers from the firm's home county and home (4 digit) industry. We estimate the same progression of models as for old colleagues. The estimates are presented in Table 7.

The coefficient estimates for Model 1 are similar to those for the old colleagues: $\hat{\beta}_{H4IR}$, $\hat{\beta}_{HR}$, and $\hat{\beta}_{H4I}$ are all positive and significant and $\hat{\beta}_{H4IR}$ is significantly greater than $\hat{\beta}_{H4I}$. Thus, consistent with Proposition 3, among firms that entered either their home (4 digit) industry or other industries, those that located in their home county were significantly more likely in their first year to hire workers from their home (4 digit) industry and county. In contrast to the assumption in the model that industry does not matter but similar to old colleagues, $\hat{\beta}_{H4I}$ is positive and significant and $\hat{\beta}_{H4IR}$ is

significantly greater than $\hat{\beta}_{HR}$ (although this is not necessarily inconsistent with the model).

In Model 2, distinctions are added for firms that entered the same 3 (but not 4) and 2 (but not 3 or 4) digit industry as their home (4 digit) industry. They yield similar patterns to the old colleagues. Consistent with Proposition 3, firms that entered in their home county and either the same 4, 3, 2 digit or other industries were significantly more likely than comparable firms that did not enter their home county to hire workers from their home (4 digit) industry and county in their first year. Also similar to old colleagues, industry mattered; whether firms entered their home county or not they were generally more likely to hire workers from their home (4 digit) industry and county the "closer" the industry they entered was to their home (4 digit) industry.[13]

In Model 3, which adds to Model 1 the control variables for firm performance, the coefficient estimates of all but *Multiplefounders* are significant. Firms with higher-level founders and founders with longer tenure at their prior employer were significantly less likely to hire workers from their home 4 digit industry and county, which is the opposite of the results for old colleagues. Furthermore, firms with founders with more tenure in their home county were more likely to hire workers from the home industry and region, which is consistent with longer tenure in the region providing knowledge about local workers outside of the founder's prior employer. Similar to the old colleagues, the addition of the control variables has little effect on the other coefficient estimates and thus their support for Proposition 3.

In Model 4 industry and regional fixed effects are added, which has little effect on the estimates. Models 4-2 and 4-3 pertain to the workers hired in years 2 and 3. $\hat{\beta}_{H4IR}$, $\hat{\beta}_{HR}$, and $\hat{\beta}_{H4I}$ are somewhat smaller than in year 1, but in contrast to the estimates for old colleagues their decline over time is modest and all remain significant by year 3 with $\hat{\beta}_{H4IR}$ still significantly larger than $\hat{\beta}_{H4I}$ and $\hat{\beta}_{HR}$ in year 3. Thus, consistent with Proposition 3, the decline in the hiring of workers from the home region and industry is less sharp than for old colleagues, and firms that entered in their home county and industry continued to hire these workers at greater rates than all other firms.

---

[13] In terms of significant differences, $\hat{\beta}_{H4IR}$ is significantly different from $\hat{\beta}_{H3IR}$ and $\hat{\beta}_{H2IR}$.

The patterns concerning the workers from the home industry and region in Table 6 helps put the probit estimates in perspective. Excluding old colleagues, in all three years around 18% of the hires of firms that entered in their home county and home (4 digit) industry came from their home (4 digit) industry and county versus 1%-5% for the other three groups of firms. Furthermore, among firms that did not enter in their home (4 digit) industry, around 3% of their hires were workers from their home industry and county if they entered in their home county versus 1% for those that entered elsewhere.

C. Worker Hazard of Exit

We next test the predictions of the model concerning the hazard of workers exiting their employer by estimating a Cox proportional hazard model of the annual hazard of worker exit, which obviates having to specify how time with the employer affects the worker hazard of exit. Workers are assumed to exit their firm if they leave before the firm exits (through 2006); otherwise they are treated as censored. All standard errors have been computed by clustering observations for all the workers of each firm.

Proposition 4 predicts that the hazard should be lowest for old colleagues and next lowest for workers initially hired from the firm's home region and industry. We test these predictions by including as explanatory variables six 1-0 dummies, denoted as *C1*, *C2*, *C3*, *WHIR1*, *WHIR2*, and *WHIR3*, which equal 1 respectively for old colleagues hired in years 1, 2, and 3 and workers from the firm's home (4 digit) industry and home county hired in years 1, 2, and 3. If some of the old colleagues hired after the first year are known to the founder, then β_{C1}, β_{C2}, and β_{C3} will all be negative and would be expected to be smaller than β_{WHIR1}, β_{WHIR2}, and β_{WHIR3} respectively. The theory predicts that over time an increasing fraction of the workers hired from the home industry and region will not be known to the founder, which implies that $\beta_{WHIR1} < \beta_{WHIR2} < \beta_{WHIR3} \leq 0$.

We also include as explanatory variables individual characteristics that have been found to be related to worker turnover, including *Age*, *Female*, *College*, *Highoccupation*, and *Middleoccupation*, where *Age* is the age of the individual when hired, *UK_Age* is a 1-0 dummy variable equal to 1 for individuals whose age is not known), and the other variables are 1-0 dummies equal to 1 respectively for females, college-educated, managers, and specialized workers. We also include a 1-0 dummy, denoted as *UK*, which equals 1 for hires whose backgrounds could not be determined, which includes

individuals that were previously public employees, students, and the unemployed. Last, we include dummies for each year of a worker's employment to allow for economy-wide factors that could affect the worker hazard.

The coefficient estimates are presented in Table 8 under Model 1. First, $\hat{\beta}_{C1}$, $\hat{\beta}_{C2}$, and $\hat{\beta}_{C3}$ are all negative and significant and as predicted they are smaller than the respective coefficient estimates of $\hat{\beta}_{WHIR1}$, $\hat{\beta}_{WHIR2}$, and $\hat{\beta}_{WHIR3}$, with the differences in the coefficients significant in years 1 and 3. Also as predicted, $\hat{\beta}_{WHIR1}$, $\hat{\beta}_{WHIR2}$, and $\hat{\beta}_{WHIR3}$ increase in value, with only the first two significant and negative. The estimates imply that old colleagues hired in each year and the workers from the home industry and region hired in years 1 and 2 had markedly lower annual hazards than other workers with known backgrounds—46%, 25%, and 30% lower for the old colleagues hired respectively in years 1, 2, and 3 and 27% and 16% lower for the workers from the home industry and region hired respectively in years 1 and 2. In terms of the individual characteristics, workers whose background is unknown had a higher hazard, as might be expected. Furthermore, those that were older, female, and held higher positions had a significantly lower hazard than other workers.

Next we estimate Model 2, which adds to Model 1 the four controls for firm performance. It might be expected that better firms would have less turnover in their employees. Consistent with this expectation, workers in firms with founders that had longer tenure at their prior employer and in their home county had significantly lower hazards, although workers in firms with higher-level founders had significantly higher hazards (the effect of multiple founders was not significant). The inclusion of these variables had little effect on the other coefficient estimates, which is also true in Model 3, which adds fixed effects for regions and industries. Thus, the worker hazard estimates continue to support Proposition 4.

D. Firm Hazard of Exit

Last, we test the predictions of the theory concerning the hazard of firm exit by estimating a Cox proportional hazards model of firm exit, which obviates having to specify how firm age affects the hazard. Firms that did not exit by 2006 are treated as censored. Proposition 5 predicts that at every age the hazard of firm exit will be lower for firms that entered in their home region, particularly those that entered as well in their

home industry. To test this, we estimate an initial hazard model in which the explanatory variables include *H4IR*, *HR*, and *H4I*, where it is expected that $\beta_{H4IR} < \beta_{HR} < \beta_{H4I} = 0$. We also include year dummies to allow the hazard to vary according to economy-wide conditions.

The coefficient estimates are presented in Table 9 under Model 1. Consistent with Proposition 5, $\hat{\beta}_{H4IR}$ and $\hat{\beta}_{HR}$ are both negative and significant, with $\hat{\beta}_{H4IR}$ significantly smaller than $\hat{\beta}_{HR}$ and $\hat{\beta}_{HR}$ significantly smaller than $\hat{\beta}_{H4I}$. However, $\hat{\beta}_{H4I}$ is also negative and significant (at the .10 level), which is not predicted by the theory but is consistent with earlier findings that industry matters, even among firms that did not enter in their home county. Relative to the omitted group of firms that did not enter in their home industry or home county, the coefficient estimates imply a 37%, 23%, and 9% lower annual hazard for firms that entered in their home county and home industry, their home county but not home industry, and their home industry but not home county respectively. The former two effects are certainly sizable, consistent with Proposition 5.

In Model 2 we break down the firms that did not enter their home (4 digit) industry into those that entered the same 3 (but not 4) and 2 (but not 3 or) digit industry as their home (4 digit) industry, which yields eight variables, as in the probits for hires: *H4IR*, *H3IR*, *H2IR*, *HOIR*, *H4I*, *H3I*, and *H2I*. Consistent with Proposition 5, $\hat{\beta}_{H4IR}$, $\hat{\beta}_{H3IR}$, $\hat{\beta}_{H2IR}$, and $\hat{\beta}_{HOIR}$ are all negative and significant while $\hat{\beta}_{H4I}$, $\hat{\beta}_{H3I}$, and $\hat{\beta}_{H2I}$ are all insignificant, which implies that firms that entered in their home region had lower hazards regardless of the industry they entered. With the omitted group now changed to firms that did not enter their home county or even the same 2 digit industry as their home (4 digit) industry, industry no longer matters among firms that did not enter in their home region, consistent with Proposition 5. Furthermore, $\hat{\beta}_{H4IR}$, $\hat{\beta}_{H3IR}$, $\hat{\beta}_{H2IR}$, and $\hat{\beta}_{HOIR}$ are ordered as expected based on Proposition 5 with the exception of the first two, which are very close in magnitude.[14] Thus, among firms that entered in their home county, those that entered closer to their home industry tended to have a lower hazard of exit, as would be expected based on Proposition 5.

[14] In terms of significant differences, $\hat{\beta}_{H4IR}$, $\hat{\beta}_{H3IR}$, $\hat{\beta}_{H2IR}$ are significantly different than $\hat{\beta}_{HOIR}$.

Next we add to Model 1 the four control variables for firm performance, which defines Model 3. The coefficient estimates of *Firmtenure*, *Highlevel*, and *Multiplefounders* are negative and significant, which is consistent with better firms having lower hazards, while the coefficient estimates of *Regionaltenure* is small and insignificant. $\hat{\beta}_{H4IR}$ and $\hat{\beta}_{HR}$ continue to be negative and significant with $\hat{\beta}_{H4IR}$ significantly smaller than $\hat{\beta}_{HR}$ and $\hat{\beta}_{H4I}$ is still negative but no longer significant. In Model 4 we add regional and industry controls, which has little effect on the coefficient estimates except that $\hat{\beta}_{H4I}$ is once again significant.

We perform one last test of the theory regarding the firm hazard of exit. We add to the model two variables, denoted as *%OldColleagues* and *%Workersfromhomeindustry&region*, which equal the percentage of the firm's hires in year 1 that were old colleagues and workers from the home county and home (4 digit) industry firms respectively. While these are crude measures of the benefits of hiring old colleagues and workers from the firm's home industry and region, they would still be expected to be related to the performance of the new firms. This clearly comes through for old colleagues. The coefficient estimate of *%OldColleagues* is negative and significant; it implies that if half of the firm's initial hires are old colleagues then its annual hazard would be 16% lower. It does not come through for workers from the firm's home county and 4-digit industry, as the coefficient estimate of *%Workersfromhomeindustry&region* is small and insignificant. Perhaps this is to be expected given that the initial percentage of workers hired from the firm's home industry and region was more limited than the percentage of old colleagues hired and the hazard of these workers was not as low as the old colleagues.

We did one last analysis. The theory implies that firms that entered in their home region have a lower hazard because they make greater use of old colleagues and workers from the home industry and region that stay longer with the firm. This should translate into their workers having lower hazards than the workers hired by firms that do not enter in their home region, especially for the firms that also entered in their home industry. We tested this by adding to Model 3 of the worker hazard analysis the variables *H4IR*, *HR*, and *H4I* and then estimated this model without and with the variables *C1*, *C2*, *C3*, *WHIR1*, *WHIR2*, and *WHIR3*. Without these variables, $\hat{\beta}_{H4IR}$, $\hat{\beta}_{HR}$. and $\hat{\beta}_{H4I}$ respectively

equal -0.141*** (0.030), -0.063** (0.029), and -0.018 (0.039), with their standard errors in parentheses. Consistent with the theory, $\hat{\beta}_{H4IR}$ and $\hat{\beta}_{HR}$ are negative and significant and $\hat{\beta}_{H4I}$ is insignificant, with $\hat{\beta}_{H4IR}$ significantly smaller than $\hat{\beta}_{HR}$.[15] Furthermore, when *C1*, *C2*, *C3*, *WHIR1*, *WHIR2*, and *WHIR3* are added to the specification, the coefficient estimates are close to 0 and insignificant.[16] Thus, consistent with the theory, the lower hazard of workers in firms that entered in their home region appears to be attributable to their greater hiring of old colleagues and workers from the home industry and region.

V. Discussion

Our findings indicate that the location of new Portuguese firms was heavily influenced by the industry that they entered. Firms that entered in their home county as well as their home industry were in turn more likely initially to hire old colleagues and workers from the home industry and region, and both types of workers had markedly lower annual hazards of exit than other workers, particularly the old colleagues. Firms that located in their home county, and especially those that entered in their home industry as well, also had markedly lower annual hazards of exit. All of these patterns are consistent with the theory. The main departures from the theory pertain to the simplifying assumption that the benefits of hiring old colleagues and workers from the home industry and region are restricted to firms that enter in their home region.

---

[15] Actually, given that the percentage of hires that were old colleagues and workers from the home industry and region were similar for firms that entered in their home county but not home industry and home industry but not home county (in contrast to what was predicted by the theory), it might have been expected that $\hat{\beta}_{HR}$ and $\hat{\beta}_{H4I}$ would both be negative and significant and similar in magnitude. We probed this by allowing *C1*, *C2*, *C3*, *WHIR1*, *WHIR2*, and *WHIR3* in Model 3 of the worker hazard to have different effects for the four groups of firms (corresponding to whether they entered their home county and/or their home industry). These estimates revealed that except for the old colleagues hired in year 1, the hazards of the old colleagues and workers from the home industry and region were not nearly as low for the firms that entered their home industry and not home county than the other three groups of firms, including those that entered in their home county and not home industry. This brings the estimates more in line with the assumption in the theory that firms that do not enter in their home region, whether they entered their home industry or not, cannot benefit by hiring old colleagues and workers from the home industry and region as much as firms that enter in their home region.

[16] When just *C1*, *C2*, *C3* are added, $\hat{\beta}_{H4IR}$, $\hat{\beta}_{HR}$ and $\hat{\beta}_{H4I}$ equal -0.044 (0.029), -0.022 (0.029), 0.023 (0.024), and when *WHIR1*, *WHIR2*, and *WHIR3* are added as well then $\hat{\beta}_{H4IR}$, $\hat{\beta}_{HR}$. and $\hat{\beta}_{H4I}$ equal -0.005 (0.029), -0.013 (0.030), 0.034 (0.039). This suggests that hiring of old colleagues is the most important factor differentiating the worker hazards of firms but hiring workers from the home industry and region is also important.

We also examined the influence of agglomeration economies related to regional activity in the firm's chosen industry and overall on its location. Similar to a prior study of Portuguese startups (Figueiredo et al. [2002]), we found that these regional characteristics affected only the attractiveness of locating in regions other than the home region. One interpretation of these patterns advanced by Figueiredo et al. [2002] is that firms possess some kind of localized knowledge about their home region that substitutes for or offsets the benefits associated greater local activity in the firm's industry and overall. Our findings suggest a leading candidate for this localized knowledge is information about prospective hires that founders of new firms acquire through their prior work experience.

The hiring patterns related to firm location that are most pronounced concern old colleagues. The first year of hires seems especially critical, as firms hire nearly twice as many workers in their first year as the next two combined. The prior employers of the firm's founders were a key source of hires in this year, especially for firms that entered in their home region. Given the markedly lower annual hazard of exit of the old colleagues hired in the first year, this appears to have provided a sizable advantage to firms that located in their home region. Indeed, in the firm hazard analysis, firms with a greater fraction of initial hires that were old colleagues had a lower hazard of exit even controlling for the region and industry entered, suggesting that the hiring of old colleagues conferred considerable benefits, as featured in the theory. While successful young firms experience considerable turnover in their management (Kaplan et al. [2009]), our findings suggest that the composition of the initial labor force plays a key role in early labor turnover and in the performance of the firm. In Kaplan et al.'s [2009] terminology, it is not just the horse but also the jockeys that influence the prospects of young firms, and firms can hire the most suitable jockeys by locating close to their geographic roots.

Can we expect findings for Portugal to apply broadly to other countries? To probe how our findings are sensitive to the composition of activity in Portugal, we conducted our analyses for different sectors and for firms with a single employee-owner, which did not alter our conclusions. Our findings also resonate with studies of firm location and survival in other countries. Dahl and Sorenson [2008] find that new Danish

firms that located in their home region survived longer, and Timmermans [2009] finds that among new Danish firms those that hired a greater fraction of old colleagues survived longer. Muendler et al. [2009] replicate our finding for Brazil that new firms were more likely to locate close to their parent firm if they entered the same industry as their parent. Buenstorf and Klepper [2009b] find that regional characteristics related to agglomeration economies influenced the attractiveness of locations outside of a firm's home county but not the home county itself. We don't doubt that other factors besides knowledge about prospective hires, such as connections to family and friends (Dahl and Sorenson [2009]) and potential sources of capital (Michelacci and Silva [2007]), may influence where firms locate. But our findings and those of related studies suggest that knowledge about potential hires, particularly old colleagues, provides a strong incentive to locate close to home, especially for firms that enter in their home industry.

Our findings naturally raise questions about why firms do not always or even regularly enter in their home industry (40% of our new firms did not enter the same two digit industry as their founder's prior employer) and why firms do not always locate in their home region whether they enter their home industry or not. The former question is certainly beyond the scope of our investigation. However, it is an important question for our analysis if the industry in which a firm enters is endogenously determined with its location choice. We abstracted from this possibility by assuming that the industry entered was based on information acquired at the prior employer before any consideration of location choice. While this seems consistent with recent industry studies of new entrants (Buensfort and Klepper [2009a], Klepper [2010]), surely there is much to be learned about the motivations for entry by different types of participants.

That brings us back to the question of why all firms don't enter close to home. This is a central question for our analysis. A key assumption for our analysis is the other, unobservable factors that influence location choice are unrelated to the central variables of our analysis. Implicit in our approach is the assumption that the founders that choose to locate in or close to their home region are not inherently different in terms of their background and quality of ventures they start from other founders. We tried to guard against this possibility by controlling for various aspects of new firms related to their founders that were related to their longevity. We are heartened by the fact that these

controls had little effect on our main findings, but acknowledge that the crudeness of the controls certainly tempers any claims that can be advanced about this key issue.

Indeed, evidence amassed in Berchicci et al. [2010] concerning the location of disk drive spinoffs that resonates with more fragmentary evidence on historical tire entrants by Buenstorf and Klepper [2009a] suggests that new firms that enter in their home industry that choose to locate in their home region may in fact be different from the outset in ways that induce them to locate close to their geographic roots. But in a number of ways, their findings actually reinforce our theory. Both studies suggest that spinoffs that at birth are closest to the technological frontier of their industry are more likely to want to hire old colleagues, which induces them to locate close to home, and provide supportive evidence for these claims. While this could explain in part why such firms survive longer, it also suggests that hiring old colleagues is a key element of their success, and this is facilitated by locating close to home. Our findings suggest that it is not just old colleagues but other workers from the firm's home region as well that might motivate them to locate close to home.

Might it be possible, though, that all of our results are driven by firms that choose to locate in their home region, especially ones that also enter their home industry, simply being superior firms? Exactly why that would be the case would have to be explained. But it could be that firms that locate in their home region are more likely to hire old colleagues and workers from the home industry and region merely because they are close by and do not have to pay a premium to hire them. Furthermore, it might be these workers stay longer with their new employers because of past connections to the firm's founders but are no more productive than other workers. Yet even if they were not more productive, their lower turnover would itself seem beneficial given the costs of hiring and training workers. Even if that were not true, though, the negative coefficient estimate of the percentage of old colleagues initially hired in the firm hazard analysis suggests that it is not just locating in the home region but also hiring old colleagues that improves the performance of firms. It may well be that firms that locate at home are innately superior performers, but their hiring practices also seem to set them apart and thus provide an incentive for them to locate in their home region.

Certainly knowledge possessed by new firms at the time of entry is not limited to knowledge about prospective hires. Just as entrepreneurs acquire knowledge about old colleagues based on their prior work experience, surely they also acquire knowledge about how to organize and operate a firm in their prior industry. This might help explain why among the firms that entered in their home region, those that entered in their home industry survived longer, and why it has been repeatedly found that new firms that enter the same industry as their parents outperform other startups. However, our findings indicate that among the firms that entered their home industry, those that entered in their home region survived longer, suggesting that the knowledge firms possess transcends industry specific knowledge. Local knowledge related to prospective hires as well as industry specific knowledge may both play a role in influencing the performance of firms and thus in turn where they locate.

To the extent that incumbent firms are the source of new firms and new firms do not locate far from their geographic roots, then a region's entry rate of new firms will be critically shaped by its incumbent producers. Klepper [2007, 2010] and Buenstorf and Klepper [2009a] argue that the extreme clustering of the historical automobile and tire industries and the modern semiconductor industry in the U.S. was attributable to such forces. Berchicci et al. [2010] argue that the most fertile spinoffs (in terms of spawning further spinoffs) are even more likely to locate close to their geographic roots, which reinforces the influence of incumbent producers on regional birth rates. To the extent that clustering is fundamentally related to the tendency of spinoffs to locate close their geographic roots, such forces need to be built into theories of industry agglomeration (Klepper [2010]).

Not surprisingly, a firm's parent appears to be a key source for its initial employees. Even if a spinoff does not directly compete with its parent firm, siphoning off employees from the parent will surely create tensions between spinoff and parent and potentially harm the parent. To understand the welfare implications of spinoffs, we will need to gain a deeper understanding of both the importance of old colleagues to new firms and how their departure from the spinoff's parent affects its performance. Recently various papers have begun to examine the latter issue (Phillips [2002], McKendrick et al. [2009], Bermiss and Murmann [2009], Campbell et al. [2009], Ioannou [2009]). Our

findings suggest that further investigation of the initial hires of new firms will provide deeper insights into the performance of the firms, their location, and the welfare implications of the ubiquitous spawning of new firms by existing producers.

Table 1 – Descriptive statistics of the firms in the sample

Location of Entrants relative to their home county	
Home County	66%
Within 10km from the home county	7%
Within 10-20km from the home county	10%
Within 20-30km from the home county	6%
Within 30-40km from the home county	2%
Elsewhere	8%

Industry of Entrants relative to their home industry	
Home 4 digit industry	44%
Home 3 but not 4 digit industry	4%
Home 2 but not 3 or 4 digit industry	11%
Not home 2, 3, or 4 digit industry	40%

Geographic location of entrants (district)	
District of Porto	20%
District of Lisboa	19%
District of Braga	12%
Elsewhere	48%

Entrants	
Nr. of workers in entry year (inc. owners)	Range: 2-86; Mean: 5; Median:3
Number of employees in entry year	Range: 1-83; Mean: 3; Median:2
Tenure in previous employer (years)	Range: 1-61; Mean: 6; Median:4
Tenure in previous county (years)	Range: 1-61; Mean: 8; Median:6
More than one founder	35%
Founder was previously manager	36%

Table 2 – Characteristics of the hires in the sample

	Year 1	Year 2	Year 3	Total
Old Colleagues	9,357 (34%)	716 (8%)	264 (4%)	10,337
Workers from the Home Region and Industry (WHIR)	1,396 (5%)	567 (6%)	434 (7%)	2,397
Workers from the Chosen Region and Industry (excluding WHIR in italics)	2,441 (9%) *1,407 (5%)*	789 (9%) *376 (4%)*	552 (9%) *237 (4%)*	3,782 *2020*
Unknown	8,602 (32%)	3,822 (43%)	2,675 (43%)	15,099
Total	27,282	8,851	6,235	42,368

Table 3 – Conditional Logit estimates

	Model 1	Model 2	Model 3	Model 4	Model 5
Home	3.209*** (0.039)	3.113***(0.044)	2.736***(0.067)	3.176***(0.069)	2.773***(0.070)
Dist	-0.028***(0.001)	-0.027***(0.001)	-0.015***(0.001)	-0.014***(0.001)	-0.014***(0.001)
Home x 4I	0.454***(0.069)	0.549***(0.072)	0.594***(0.071)	0.560***(0.070)	0.658***(0.069)
Home x 3I		0.393** (0.164)	0.401** (0.157)	0.408***(0.155)	0.502***(0.152)
Home x 2I		0.224** (0.104)	0.246** (0.099)	0.263***(0.098)	0.302***(0.097)
Dist x 4I	-0.012***(0.001)	-0.013***(0.001)	-0.010***(0.001)	-0.009***(0.001)	-0.008***(0.001)
Dist x 3I		-0.007** (0.003)	-0.005* (0.003)	-0.004*(0.002)	-0.002 (0.002)
Dist x 2I		-0.006***(0.002)	-0.005** (0.002)	-0.004**(0.001)	-0.004***(0.001)
Home x FT			0.097 (0.065)	0.006 (0.064)	0.010 (0.064)
Dist x FT			0.011***(0.002)	0.009***(0.001)	0.008*** (0.001)
Home x HL			0.347*** (0.067)	0.422***(0.066)	0.471***(0.065)
Dist x HL			-0.002* (0.001)	-0.001 (0.001)	-0.002*(0.001)
Home x MF			0.048 (0.068)	0.059 (0.067)	0.111* (0.066)
Dist x MF			-0.018***(0.002)	-0.007***(0.001)	-0.006***(0.001)
Home x RT			-0.055 (0.064)	0.090 (0.063)	0.142**(0.063)
Dist x RT			-0.018***(0.002)	-0.016***(0.001)	-0.014***(0.001)
WD x (1-H)				0.147**(0.078)	0.147* (0.089)
PD x (1-H)				0.051*** (0.002)	0.059** (0.028)
WD x H				0.009 (0.072)	0.071 (0.073)
PD x H				-0.032***(0.002)	-0.016 (0.028)
CD					Included

Observations	10236 x 275=2814900				
Log likelihood	-21285.600	-21228.879	-20738.405	-20238.869	-19186.766

***significance at the 0.01 level; **significance at the 0.05 level; *significance at the 0.1 level. FT, HL, MF, RT, stand for Firmtenure, Highlevel, Multiplefounders, Regionaltenure, respectively.

Table 4 – Characteristics of the firms in the sample, by geographical localization (columns) and according to relationship with home 4 digit industry (rows)

Geo. localization / Rel. w. home ind.	Distance to Home County (km)						
	Home County	1-10	10.1-20	20.1-30	30.1-40	>40	Total
Same 4 digit industry	77%	5%	7%	4%	1%	4%	44%
Same 3 digit industry, not 4	69%	9%	9%	6%	2%	6%	4%
Same 2 digit ind., not 3 or 4	66%	8%	10%	7%	2%	7%	11%
Not 2 digit industry	54%	10%	12%	9%	3%	13%	40%
Total	66%	7%	10%	6%	2%	8%	100%

In every cell (apart for total column/row), shares correspond to percentages of row total.

Table 5 – Probit regressions of the probability that a new hire is from the founders' previous employer, for years 1, 2, and 3

	Year 1				Year 2	Year 3
	Model 1	Model 2	Model 3	Model 4	Model 4	Model 4
H4IR	1.177*** (0.046)	1.360*** (0.053)	1.049*** (0.047)	1.067*** (0.049)	0.511*** (0.080)	0.412*** (0.115)
H3IR		1.285*** (0.091)				
H2IR		1.010*** (0.065)				
HOIR		0.701*** (0.059)				
H4I	0.619*** (0.073)	0.802*** (0.077)	0.606*** (0.071)	0.661*** (0.068)	0.433*** (0.101)	0.262** (0.134)
H3I		0.811*** (0.132)				
H2I		0.568*** (0.095)				
HR	0.667*** (0.048)		0.608*** (0.049)	0.596*** (0.049)	0.219*** (0.078)	0.247** (0.123)
Firmtenure			0.238*** (0.032)	0.223*** (0.033)	0.236*** (0.055)	0.124* (0.070)
Highlevel			0.323*** (0.031)	0.333*** (0.031)	0.208** (0.063)	0.108 (0.076)
Multiplefounders			0.115*** (0.030)	0.093*** (0.030)	0.011 (0.060)	-0.045 (0.077)
Regionaltenure			0.021 (0.033)	0.009 (0.034)	-0.117** (0.056)	-0.096 (0.075)
2 Digit Industry Dummies				Included	Included	Included
District Dummies				Included	Included	Included
Constant	-1.185*** (0.041)	-1.369*** (0.048)	-1.697*** (0.050)	-1.621*** (0.111)	-1.729*** (0.239)	-1.861*** (0.308)
Observations	27282	27282	27282	27277 [1]	8770 [1]	6057 [1]
Log Pseudolikelihood	-16222.195	-16037.695	-15559.331	-15268.338	-2304.4512	-1037.7086

***significance at the 0.01 level; **significance at the 0.05 level; *significance at the 0.1 level. Robust standard errors in parenthesis.
[1] Observations were dropped due to collinearity.

Table 6 – Characteristics about the hires in the sample, per group of entrepreneurs, excluding and including old colleague

	Year 1	Year 2	Year 3
Entrepreneur enters home region and industry	C: 50% WHIR: 9% *(19%)* WCIR: 9% *(19%)* UK: 26%	C: 12% WHIR: 15% *(17%)* WCIR: 15% *(17%)* UK: 41%	C: 5% WHIR: 16% *(17%)* WCIR: 15% *(16%)* UK: 43%
Entrepreneur enters home region, not industry	C: 30% WHIR: 3% *(4%)* WCIR: 8% *(12%)* UK: 32%	C: 7% WHIR: 3% *(3%)* WCIR: 7% *(7%)* UK: 45%	C:4% WHIR: 3% *(3%)* WCIR: 6% *(6%)* UK: 43%
Entrepreneur enters home industry, not region	C: 29% WHIR: 4% *(5%)* WCIR: 10% *(13%)* UK: 33%	C: 9% WHIR: 4% *(4%)* WCIR: 8% *(9%)* UK: 41%	C: 4% WHIR: 5% *(5%)* WCIR: 8% *(8%)* UK: 43%
Entrepreneur does not enter home region, nor home industry	C: 12% WHIR: 1% *(1%)* WCIR: 9% *(10%)* UK: 40%	C: 4% WHIR: 1% *(1%)* WCIR: 5% *(5%)* UK: 45%	C: 3% WHIR: 1% *(1%)* WCIR: 5% *(5%)* UK: 42%

C: old Colleagues; WHIR: Workers from the Home Region and Industry; WCIR: Workers from the Chosen Industry and Region; UK: Uknown background.
Shares correspond to the fraction of workers with a given relationship hired by a certain group of entrepreneurs in a given year (ie, in their first year, 50% of the workers hired by entrepreneurs that entered their home region and industry were old colleagues). Shares in italic correspond to shares of workers excluding old colleagues.

Table 7 – Probit regressions of the probability that a new hire is from the founders' home region and industry, for years 1, 2, and 3

	Year 1				Year 2	Year 3
	Model 1	Model 2	Model 3	Model 4	Model 4	Model 4
H4IR	1.539*** (0.085)	1.624*** (0.107)	1.504*** (0.086)	1.353*** (0.088)	1.191*** (0.116)	1.188*** (0.126)
H3IR		1.168*** (0.248)				
H2IR		0.867*** (0.124)				
HOIR		0.587*** (0.118)				
H4I	0.808*** (0.104)	0.893*** (0.123)	0.800*** (0.106)	0.719*** (0.111)	0.573*** (0.137)	0.658*** (0.154)
H3I		0.217 (0.233)				
H2I		0.319* (0.168)				
HR	0.642*** (0.095)		0.590*** (0.100)	0.555*** (0.100)	0.430*** (0.129)	0.436*** (0.138)
Firmtenure			-0.196*** (0.056)	-0.155*** (0.041)	0.040 (0.044)	-0.126*** (0.051)
Highlevel			-0.138** (0.055)	-0.044 (0.050)	-0.023 (0.069)	-0.052 (0.070)
Multiplefounders			-0.008 (0.053)	-0.016 (0.050)	0.148** (0.064)	0.040 (0.066)
Regionaltenure			0.262*** (0.055)	0.177*** (0.041)	0.054 (0.048)	0.200*** (0.055)
2 Digit Industry Dummies				Included	Included	Included
District Dummies				Included	Included	Included
Constant	-2.427*** (0.0760)	-2.512*** (0.100)	-2.538*** (0.093)	-1.950*** (0.281)	-2.689*** (0.265)	-3.177*** (0.344)
Observations	17925	17925	17925	17764 [1]	7991 [1]	5895 [1]
Log Pseudolikelihood	-4211.3007	-4185.28	-4145.8754	-3914.1788	-1631.211	-1247.0042

***significance at the 0.01 level; **significance at the 0.05 level; *significance at the 0.1 level. Robust standard errors in parenthesis.

[1] Observations were dropped due to collinearity due to the large number of variables.

Table 8 – Cox proportional hazard model of the annual hazard of worker exit

	Model 1	Model 2	Model 3
C1	-0.617*** (0.025)	-0.587*** (0.026)	-0.552*** (0.025)
C2	-0.291*** (0.064)	-0.271*** (0.064)	-0.247*** (0.064)
C3	-0.358*** (0.117)	-0.341*** (0.116)	-0.359*** (0.114)
WHIR1	-0.319*** (0.050)	-0.305*** (0.041)	-0.268*** (0.049)
WHIR2	-0.176*** (0.063)	-0.152** (0.061)	-0.125** (0.061)
WHIR3	-0.094 (0.072)	-0.071 (0.072)	-0.021 (0.069)
Age	-0.002*** (0.001)	-0.002*** (0.001)	-0.003*** (0.001)
Female	-0.123*** (0.019)	-0.118*** (0.019)	-0.077*** (0.018)
College	0.036 (0.050)	0.023 (0.046)	0.027 (0.048)
Highoccupation	-0.162*** (0.044)	-0.175*** (0.045)	-0.186*** (0.043)
Middleoccupation	-0.049*** (0.016)	-0.052*** (0.016)	-0.053*** (0.016)
UK_Age	-0.078 (0.065)	-0.073 (0.066)	-0.121* (0.068)
UK	0.156*** (0.016)	0.157*** (0.016)	0.145*** (0.015)
Firmtenure		-0.052*** (0.019)	-0.046** (0.013)
Highlevel		0.035 (0.021)	0.014 (0.021)
Multiplefounders		-0.012 (0.020)	-0.009 (0.020)
Regionaltenure		-0.040* (0.019)	-0.023 (0.018)
Year Dummies	Included	Included	Included
2 Digit Industry Dummies			Included
District Dummies			Included

Subjects	40993		
Failures	21914		
Observations	97420		
Log likelihood	-220960.87	-220897.22	-220676.25

***significance at the 0.01 level; **significance at the 0.05 level; *significance at the 0.1 level. Robust standard errors in parenthesis.

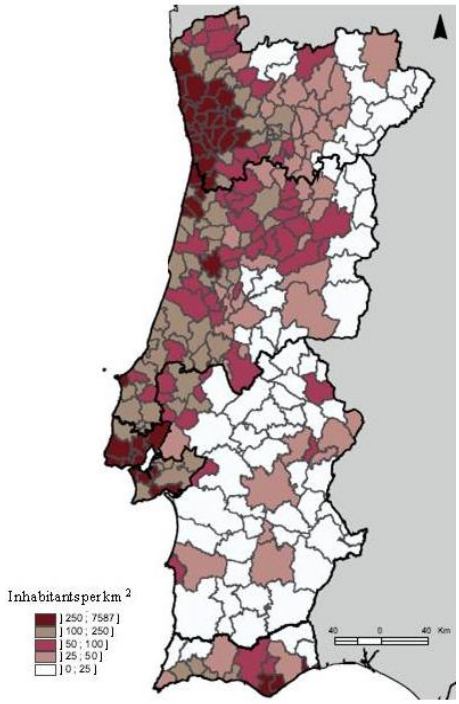
Table 9 – Cox proportional hazard model of the annual hazard of firm exit

	Model 1	Model 2	Model 3	Model 4	Model 5
%OldColleagues					-0.004*** (0.001)
%Workersfromhomeindustry&region					0.000 (0.001)
H4IR	-0.455*** (0.041)	-0.445*** (0.044)	-0.314*** (0.043)	-0.351*** (0.044)	-0.266*** (0.047)
H3IR		-0.472*** (0.104)			
H2IR		-0.401*** (0.068)			
HOIR		-0.180*** (0.046)			
H4I	-0.097* (0.055)	-0.086 (0.057)	-0.074 (0.055)	-0.111** (0.056)	-0.079 (0.056)
H3I		-0.180 (0.133)			
H2I		0.097 (0.077)			
HR	-0.266*** (0.040)		-0.187*** (0.040)	-0.166*** (0.041)	-0.130*** (0.041)
Firmtenure			-0.226*** (0.027)	-0.209*** (0.027)	-0.194*** (0.028)
Highlevel			-0.181*** (0.034)	-0.160*** (0.035)	-0.124*** (0.035)
Multiplefounders			-0.309*** (0.034)	-0.313*** (0.035)	-0.304*** (0.035)
Regionaltenure			0.015 (0.027)	0.009 (0.027)	0.011 (0.027)
Year Dummies	Included	Included	Included	Included	Included
2 Digit Industry Dummies				Included	Included
District Dummies				Included	Included

Subjects	10236				
Failures	4177				
Observations	42904				
Log likelihood	-36278.932	-36269.663	-36133.491	-36007.127	-35983.737

*** significance at the 0.01 level; ** significance at the 0.05 level; * significance at the 0.1 level. Standard errors in parenthesis.

Figure 1 – Population density in mainland Portugal, 2002 (adapted from "Retrato Territorial de Portugal, 2003", Instituto Nacional de Estatistica)



References

Berchicci, Luca, Andrew King, and Christopher L. Tucci, "Does the Apple Always Fall Close to the Tree? The Stay-Leave Choice of Spinouts," 2010, mimeo.

Bermiss, Y. Sekou and Johann P. Murmann. "The Individual and the Organization: Where Do Knowledge-Based Capabilities Reside in Professional Service Firms?," 2009, mimeo.

Buenstorf, Guido and Steven Klepper. "Heritage and Agglomeration: The Akron Tire Cluster Revisited," *Economic Journal* 119 (2009a), 705-733.

Buenstorf, Guido and Steven Klepper. "Why Does Entry Cluster Geographically? Evidence from the U.S. Tire Industry," 2009b, mimeo.

Campbell, Benjamin A., Martin Ganco, April M. Franco, and Rajshree Agarwal. "Who Leaves, Where to, and Why Worry? Employee Mobility, Employee Entrepreneurship, and Effects on Source Firm Performance," 2009, mimeo.

Carlton, Dennis W. "Why New Firms Locate Where They Do: An Economic Model," in *Interregional Movements and Regional Growth*, 1979, William C. Wheaton, ed., Washington D.C.: The Urban Institute.

Carlton, Dennis W. "The Location and Employment Choices of New Firms: An Econometric Model with Discrete and Continuous Endogenous Variables," *The Review of Economics and Statistics* 65 (1983), 440-449.

Dahl, Michael and Olav Sorenson. "The Social Attachment to Place," 2008, mimeo.

Dahl, Michael and Olav Sorenson. "The embedded entrepreneur," *European Management Review* 6 (2009), 172-181.

Dahl, Michael and Olav Sorenson. "The migration of technical workers," *Journal of Urban Economics* 67(2010), 33-45.

Figueiredo, Octavio, Paulo Guimaraes, and Douglas Woodward. "Home-field advantage: location decisions of Portuguese entrepreneurs," *Journal of Urban Economics* 52 (2002), 341-361.

Ioannou, Ioannis. "The Impact of Different Types of Spinoffs on Firm Survival: Evidence from the US Automobile Industry 1890-1986," 2009, mimeo.

Kaplan, Steven N., Berk A. Sensoy, and Per Strömberg, "Should Investors Bet on the Jockey or the Horse? Evidence from the Evolution of Firms from Early Business Plans to Public Companies," *Journal of Finance* 64 (2009), 75-115.

Klepper, Steven. "Disagreements, Spinoffs, and the Evolution of Detroit as the Capital of the U.S. Automobile Industry," *Management Science* 53 (2007), 616-631.

Klepper, Steven. "The Origin and Growth of Industry Clusters: The Making of Silicon Valley and Detroit," *Journal of Urban Economics* 67(2010), 15-32.

Klepper, Steven and Peter Thompson, "Disagreements and Intra-industry Spinoffs," *International Journal of Industrial Organization*, forthcoming (2010).

McKendrick, David G., James B. Wade, and Jonathan Jaffee. "A Good Riddance? Spin-offs and the Technological Performance of Parent Firms," *Organization Science* 20 (2009), 979-992.

Michelacci, Claudio and Olmo Silva. "Why So Many Local Entrepreneurs?," *Review of Economics and Statistics* 89 (2007), 415-433.

Muendler, Marc-Andreas and James E. Rauch. "Employee Spinoffs and Industrial Clustering: Evidence from Brazil," 2009, mimeo.

Pflueger, Michael and Jens Suedekum. "A synthesis of footloose-entrepreneur new economic geography models: when is agglomeration smooth and easily reversible?," *Journal of Economic Geography* 8 (2008), 39-54.

Phillips, Damon J. "A Genealogical Approach to Organization Life Chances: The Parent-Progeny Transfer," *Administrative Sciences Quarterly* 47 (2002), 474-506.

Rosenthal, Stuart S. and William C. Strange. "Evidence on the Nature and Sources of Agglomeration Economies," *Handbook of Regional and Urban Economics, Vol 4: Cities and Geography*, 2004, J. Vernon Henderson and Jacques-Francois Thisse (eds.), Elsevier: Amsterdam.

Timmermans, Bram. "The Effect of Previous Co-Worker Experience on Firm Survival," 2009, mimeo.