



# Entrepreneurship, the initial labor force, and the location of new firms

Cristina Carias · Steven Klepper · Rui Baptista 

Accepted: 1 March 2022 / Published online: 13 April 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

**Abstract** We propose that new firm founders locate their firms close to their home region in order to hire workers they know about through their prior employment, since it is easier to find high productivity employees among talent pools for which you have significant personal experience. We test our proposition using a matched employer–employee dataset for Portugal. Consistent with our predictions, new firms in the same industry as their founder’s prior employer (i.e., spinoffs) are 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. Results suggest that the agglomeration of high performing spinoffs next to their parent firms should facilitate the emergence of successful industrial clusters.

**Plain English Summary** Industry clusters emerge because successful spinoff firms are more likely to locate in their founders’ home region and draw on the local talent pool to hire their employees. Theories about industry clusters state that entrepreneurs are drawn to regions that contain more activity in their industry because there are benefits to agglomeration, such as larger pools of specialized labor. However, empirical evidence suggests that clusters emerge because entrepreneurs tend to locate close to their home region rather than being attracted by agglomerations. We reconcile these views by proposing that new firm founders locate their firms close to their home region in order to hire workers they know about through their prior employment, since it is easier to

---

This research stems from Cristina Carias’ doctoral work in the Department of Social and Decision Sciences under Steven Klepper’s direction. An early version of this paper, co-authored by Cristina Carias and Steven Klepper, has circulated since 2010. The paper was in the review process in a different journal when Steven Klepper passed away. Carias – who had by then joined the private sector– approached Rui Baptista in 2019 for help in developing a new, updated theoretical background for the paper. Rui Baptista’s contribution is therefore limited to the first two sections of the paper.

---

Steven Klepper (Deceased, 27 May, 2013).

---

C. Carias  
IN+, Centre for Innovation, Technology and Policy  
Research, Instituto Superior Tecnico, Universidade de  
Lisboa, Lisbon, Portugal

R. Baptista (✉)  
CEG-IST, Instituto Superior Técnico, Universidade de  
Lisboa, Lisbon, Portugal  
e-mail: rui.baptista@tecnico.ulisboa.pt

S. Klepper  
Arthur Arton Hamerschlag Professor of Economics  
and Social Science, Department of Social and Decision  
Sciences, Carnegie Mellon University, 5000 Forbes  
Avenue, BP 208, Pittsburgh, PA, United States

find high productivity employees among talent pools that you know well.

We predict that new firms in the same industry as their founder's prior employer (usually called spinoffs, or spinouts) are 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. We test our predictions using a matched employer–employee dataset for Portugal. Results confirm our predictions, suggesting that the agglomeration of high-performing spinoffs next to their founders' previous employers is likely to facilitate the emergence of successful industrial clusters.

**Keywords** Location · Human Capital · Clusters · Spinoffs · Home Region · Agglomeration

**J.E.L. Classifications** L26 · M13

## 1 Introduction

Entrepreneurs are often described as “foot-loose” when choosing where to locate (Pflueger & Suedekum, 2008). Several theories of agglomeration and entrepreneurial clusters state that entrepreneurs are drawn to regions that contain more activity in their industry and overall, reflecting the influence of agglomeration economies (Delgado et al., 2010; Rosenthal & Strange, 2004).

However, empirical observation does not seem to corroborate this view of entrepreneurial mobility, suggesting instead that new firm founders do not stray far from where they previously worked and/or lived, which we call their home region. This finding is supported by observations of successful clusters as well as empirical analyses using longitudinal data for entire countries. For instance, 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 were themselves 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, while Costa and Baptista (2015) paint a similar picture for the Portuguese molds for plastic injection industry in Marinha Grande. Buenstorf and Klepper (2009) and Berchicci et al. (2011) 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 startup entrants in Portugal (Figueiredo et al., 2002) and Denmark (Dahl & Sorenson, 2010a) reveal a similar tendency of all types of new firms to locate close to where their founders previously worked and resided.

These findings raise several questions. Is this tendency to locate close to the entrepreneur's home region truer of spinoffs<sup>1</sup> 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 does this propensity play in the formation of successful industrial clusters?

We propose that the location of new firms is heavily influenced by knowledge that founders have about the skills, or human capital, of prospective hires based on the networks they have established during their prior work experience. A natural source of employees for new firms is their founder's prior employer. Also, we 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 & Sorenson, 2010b)—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, making their skills more easily translatable to the new work setting and, therefore, making 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 the same industry as their founder's prior employer.

The localized knowledge founders acquire through professional and social networks might be expected to transcend human capital. Founders may have

<sup>1</sup> The definition of ‘spinoffs’ used here follows the one adopted by Garvin (1983) and Klepper (2009)—i.e. independent startups founded by one or more individuals straight after leaving employment in an incumbent (parent company) in the same industry. These startups have no equity relationship with incumbents. This event is akin to the “spinout” concept described by Agarwal et al. (2004).

connections to family and friends (Dahl & Sorenson, 2010a) and to sources of capital (Michelacci & Silva, 2007) that could also induce them to locate close to their home region. Social ties to potential customers, technology and market specialists, sources of finance, and suppliers of intermediate goods and services developed while working for incumbents in the same industry are likely to enable spawning of better startups. Nevertheless, knowledge of local human capital is likely to play a distinctive role in the success of new firms.

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. However, the match between a new firm's ideas and its hires may also play a critical role in its performance. Recent literature highlights the existence of complementarities between workers' and entrepreneur's backgrounds and skills (Baptista et al., 2013). If employees coming from the founder's prior employer, or from firms in the same location and industry, match better with the requirements of the new startup, they should be more productive than other employees with similar levels of human capital (such as, for instance, formal education and labor experience).

The implications of the theory are potentially far reaching. If local hires improve matching between startups—and, in particular, spinoffs—and employees, new firms' hiring decisions would drive some of the patterns we often associate with successful industrial clusters—i.e., entrants locating in agglomerated areas (where successful incumbents already locate) and performing better in such areas. Hiring decisions would contribute to explain why industry linkages (i.e., the presence and strength of incumbent firms in the same or in related industries) are a particularly important determinant of the “supply of entrepreneurs” (Chatterji et al., 2014; Delgado et al., 2010, 2014) and a powerful predictor of future entrepreneurship for a city or region is (Buenstorf & Klepper, 2009, 2010; Klepper, 2007, 2010).

We test our theory using a linked employer–employee data for Portugal that provides information about all new firms and their employees, including their employee-founders. We attempt to isolate the role that knowledge of the human capital of potential employees has on industry agglomeration by examining the location of new firms and the types of workers they hire, the longevity of their hires, and the performance of the firms. We find that 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. 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.

The paper is organized as follows. In Section II we discuss the background to our theory and develop various implications for new firm location and performance. In Section III we describe the data and the variables we construct to test the implications developed in Section II. In Section IV we report estimates of various 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.

## 2 Theory

### 2.1 Background

The importance of human resources for firm success has long been recognized by human capital theory, transaction cost economics, and the resource-based view of the firm. Human resources are critical for new firm survival and growth (Cooper et al., 1994; Hayton, 2003). The early hiring decisions of new firms will influence their performance and should be difficult to emulate or change when a firm is older. As such, the quality of the labor force assembled when a firm is young may serve as the basis for the firm's capabilities.

Despite the recognition of entrepreneurship as a collective activity (Kamm et al., 1990), earlier studies on team dynamics within the domain of entrepreneurship focus mostly on entrepreneurial or founding teams—understood as comprising those people who have a certain degree of ownership and/or control in the new venture (Ucbasaran et al., 2003)—rather than on the workforce of startups. However, employees recruited during the early stages of the organizational lifecycle can play a vital role in developing the startup and shaping its future. The initial hires a firm makes can constrain them later and affect their performance (Cardon & Stevens, 2004; Timmermans, 2009; Coad et al., 2014, 2017).

We propose a conceptual model predicting how firms choose their first employees. The model is based

on theories of human capital. These theories predict that, among other types of human capital investments, worker productivity is enhanced experience that can be firm-specific (Lazear, 2009) and industry-specific (Neal, 1995). Greater firm-specific and industry-specific experience means greater knowledge about a wide range of factors that influence productivity, such as organizational processes and routines, customer demand, products, technologies, suppliers and competitors (Helfat & Lieberman, 2002).

Former colleagues are likely to possess firm-specific knowledge that should be especially useful to the founder (Rocha et al., 2018). Dahl and Klepper (2015) argue that better firms (i.e., better founders) will hire better workers, so there will be a match between the human capital levels of founders and employees. Examining small entrepreneurial firms, Baptista et al. (2013) find that workers' wages reflect the value of the match with entrepreneur's skills. Dahl and Klepper (2015) find that entrepreneurs with better pre-entry experience start better (i.e., larger) firms and these firms hire more talented employees. Thus, workers are allocated to new firms according to the match between their abilities and the human capital of the founders, giving rise to enduring firm capabilities.

The concept of "skill relatedness" developed by Neffke and Henning (2013) provides an explanation for the human capital linkages between incumbents and spinoffs. These authors argue that, given the pivotal role of human capital in a firm's strategic asset stocks, a firm will likely focus its diversification efforts in areas that require skills already possessed by its current workforce. Thus, the intensity of labor mobility flows across industries is likely to predict the direction of firm diversification. Adopting a regional perspective, Neffke et al. (2018) measure labor mobility across firms and industries and build on the notion of skill relatedness, finding that local entrepreneurs are more likely to found spinoffs or firms in industries that are closely related to their industry of origin, while non-local entrepreneurs are more likely to start firms in industries less related to the regional industry mix. For new firm founders, the simplest way to assess the human capital of prospective employees is by observation and interaction in the same workplace or, at least, in the same industry/marketplace. Startup founders will seek to recruit employees who match their ideas regarding the market, technology, and organization of the new firm. For new firm founders,

the simplest way to assess the human capital of prospective employees is by observation and interaction in the same workplace or, at least, in the same industry/marketplace. Startup founders will seek to recruit employees who match their ideas regarding the market, technology, and organization of the new firm. The better the match, the greater the productivity of the new firm, and the greater its probability of success.

Recent studies find that labor mobility between incumbents and spinoffs plays a significant role in the development of geographical clusters. Examining the historical evolution of four prominent industry clusters—automobiles in Detroit, Michigan, tires in Akron, Ohio, semiconductors in Silicon Valley, California, and cotton garments in Dhaka, Bangladesh—Klepper (2011) finds that the main agglomeration mechanism at work in the four clusters involves employees leaving established firms to found their own firms. Examining the mobility of semiconductor inventors in Silicon Valley, Cheyre et al. (2015) find that mobility of semiconductor inventors in Silicon Valley was due primarily to spinoffs, proposing that benefits from greater inventor mobility were mainly experienced by entrants and not incumbent producers. Using linked employer–employee data, Buenstorf and Costa (2018) show that early within-industry hires contribute to the performance of spinoffs located in clusters.

## 2.2 Implications for Human Capital, Location, and Performance

A significant amount of literature in entrepreneurship (e.g., Agarwal et al., 2004; Klepper, 2009) proposes that employees with greater industry experience are more likely to recognize a new venture opportunity—i.e., come up with a good business idea—hence they are more likely to start a new firm (i.e., a spinoff, as defined in Section I above). It is reasonable to expect that, for a new firm founded by an employee of an incumbent firm, the nature of the opportunity/business idea dictates the industry chosen by the new firm, while the firm founder chooses the region of location.<sup>2</sup> Thus, in our conceptual model

<sup>2</sup> Throughout the remainder of the text we will refer to the location of the incumbent firm—i.e., where its founder previously worked—as the home region of the founder and also the new firm; we will refer to the industry of the incumbent firm—i.e., the industry in which the founder previously worked—as the home industry of the founder and the new firm.

entrepreneurs choose to enter in the region that maximizes the expected productivity of their ventures.

Startups require a workforce when they are formed. In our conceptual model, we assume that workers are heterogeneous in their abilities and only those whose abilities match the needs of a new firm will be suitable hires.<sup>3</sup> 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.<sup>4</sup> However, founders learn about workers in their prior employer (i.e., former colleagues), which enables them to project the productivity of these workers in their new firm more accurately than any alternative hire.

While working for their prior employer, founders also interact with employees in other incumbent firms, thereby enabling them to project the productivity of these workers in their new firm. Such interactions are likely to be more intense with workers in firms located in the founders' home industry and region (although not as intense as with former colleagues). Therefore, most of the potential hires for which founders have distinctive information about their productivity are from their home region and industry. Thus, as founders search for high productivity employees, they are likely to find them among talent pools for which they have significant personal experience—i.e., in their parent firm, home region and industry. This implies that new founders will prefer to locate in their home region and industry in order to access the talent pool that is better known to them.

Moving across regions and across sectors leads to a depreciation in the value of region- and industry-specific knowledge. Almeida and Kogut (1999) show that flows of knowledge associated with inter-firm mobility of engineers are embedded in regional labor networks. Neffke and Henning find that individuals changing jobs are more likely to remain in the same or a related industry because their specific skills are more valuable in those industries. Since mobility costs are likely to hinder productivity, new founders should be

more likely to locate in their home industry and look for prospective workers in the local talent pool.

The greater a firm's expected productivity from locating in a region, the greater the probability of entering there. If founders can better project the expected productivity of prospective workers originating from their previous employer, or from their home region, firms will be more likely to enter in their home region than elsewhere. Moreover, given that founders entering their home industries can better project the expected productivity (in the new firm) of workers originating from the home industry, the likelihood of locating in the home region will be greater for founders entering in their home industry than for founders entering in a different chosen industry, as the workers that founders know about should be less likely to suffer any reduction in their productivity from switching industries. Hence, we can propose the following:

**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.*

Next, our conceptual model considers the hiring choices of firms. A worker might have higher productivity at a new firm than at 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. A founder that has interacted with such workers, either in the incumbent firm that was his previous employer or in the home region and industry, is likely to have specific knowledge about them that suggests they would be more likely to be high productivity workers at the founder's new firm. If the new firm enters the home region and industry of the founder, this knowledge is especially valuable. The founder should thus be more likely to hire these workers than other workers with comparable characteristics in terms of formal education and experience. If the firm enters its home region but not its home industry, knowledge about labor will still be relevant, but likely not as much as if the founder had entered in its home industry, given that the founder is less able to project the productivity of labor in a different chosen industry.

<sup>3</sup> We assume that any region where the firm might locate has enough suitable workers matching its needs, so agglomeration costs or scarcity of labor do not play a determinant role in the selection or exclusion of a region.

<sup>4</sup> While we acknowledge that this assumption is somewhat unrealistic, we submit that the everyday contact with a co-worker should provide significantly greater knowledge about performance than second-hand information acquired in industry and professional networks.

For new firms that locate away from the founder's home region, costs ensuing from relocation make it less likely that former colleagues and other workers originating from the home region will be hired. For firms entering an industry other than the home industry and locating outside the home region, costs from relocation and switching industries will make hiring workers from the home region the most unlikely.

In the years following entry, it is likely that founders will gradually deplete their knowledge of specific hires from the incumbent firm, home region, and industry that are willing to move to their firm. At the same time, as firms acquire more knowledge about workers in their chosen region and industry, they will be more likely to hire workers originating from both. Thus, while founders that enter in their home region and industry may still find value in hiring workers from their home industry and region (because they coincide with the chosen region and chosen industry), the same will not be true for founders entering their home region but not home industry, or other firms. We can therefore propose the following:

**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.*

**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 likely 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 for other firms.*

We now consider the hazard of exit of workers from their employers. Our conceptual model assumes that, in each period that the firm operates, it acquires additional knowledge about productivity for any of its employees, as well as for workers in its chosen industry and region. As the firm adjusts its expectations of workers' productivity levels, it will replace those employees whose productivity is lower than expected at the time of hire. The longer a worker stays with the firm, the more likely it is that (s)he was confirmed as a high productivity employee. Those initial employees

whose expected productivity was better known to the founder at startup (i.e., former colleagues and those originating from home region and industry) are more likely to remain with the firm longer, as the likelihood of the founder overestimating their productivity is lower. 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. According to our theory, firms enter their home region in order to be able to exploit regional knowledge about labor; they thus employ workers with higher expected productivity, which maximizes the firms' expected productivity. Since the likelihood of employing higher productivity workers is greater for those firms that entered their home industry and located in their home region, firms that enter in their home region and home industry are more likely to be productive and profitable. Firms that enter in their home region (but not home industry) can also exploit regional knowledge about labor, but this knowledge will be less valuable, since the founder is less able to predict the performance of the old colleagues and workers from the home region and industry in the (different) chosen industry. We cannot test these predictions directly given the limitations of the 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. We therefore propose the following:

**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.*

### 3 Data and Variables

Data are drawn from the "Quadros de Pessoal," which is a matched employer–employee database for Portugal. Submission of data is mandatory for every



Portuguese firm with at least one employee. 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 creation.<sup>5</sup> 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.<sup>6</sup>

We excluded from the sample 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. We

excluded entrants with more than one establishment because of their multiple locations.

We identify founders/entrepreneurs as employees that are 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). We included only entrants with at least one employee in their first year (that was not an owner). We consider only firms where there is information allowing us to trace the owners' and employees' backgrounds (i.e., previous employer, home firm, and home region). In total, we ended up with 10,236 entrants<sup>7</sup> that hired 27,282 workers in their first year, 8,851 workers in their second year, and 6,235 workers in their third year.<sup>8</sup>

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. Portuguese counties are roughly about a quarter of the size of US counties.<sup>9</sup> 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

<sup>5</sup> 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.

<sup>6</sup> 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.

<sup>7</sup> 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 not significantly different 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.

<sup>8</sup> 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.

<sup>9</sup> 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.

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.<sup>10</sup> 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, and laborer) according to the last position (s)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 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 dataset 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 10,236 firms in the sample. In terms of location, 66% of the firms located in their home county, 8% located within 10 km (km) of their home county (but not in it), 10% located between 10 and 20 km of their home county, 7% located between 20

and 40 km of their home county, and the remaining 8% located elsewhere.<sup>11</sup> 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 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.

<sup>10</sup> We traced the work history of every founder from 1987 (the first year of the dataset) onwards to determine tenure in the home region.

<sup>11</sup> In Portugal, the time to commute increases discontinuously with the number of county borders one has 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 15 km travel from a county nearby Lisboa to the county of Lisboa may take 15 min 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.



**Table 1** – Descriptive statistics of the firms in the sample

Location of Entrants relative to their home county	
Home County (Home)	66%
Within 10 km from the home county	7%
Within 10-20 km from the home county	10%
Within 20-30 km from the home county	6%
Within 30-40 km from the home county	2%
Elsewhere	8%
Industry of Entrants relative to their home industry	
Home 4-digit industry (H4I)	44%
Home 3 but not 4-digit industry (H3I)	4%
Home 2 but not 3 or 4-digit industry (H2I)	11%
Not home 2-, 3-, or 4-digit industry (HOI)	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 including 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 in years (Tenure/FT)	Range: 1–61; Mean: 6; Median:4
Tenure in previous county in years (Regional Tenure/RT)	Range: 1–61; Mean: 8; Median:6
More than one founder (Multiplefounders/MF)	35%
Founder was previously manager (Highlevel/HL)	36%

Variable names in parentheses

### 4 Empirical Analysis

We first test our concerning the location choices of firms, followed by the propositions regarding the hiring decisions of firms, the hazard of exit of their workers, and the hazard of exit of the firms. Table 3 defines the variables used in our estimations.

#### 4.1 Location Choices

We test the location predictions 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 a variable in  $\mathbf{x}_{ij}$ ,  $Home$ , which equals one if county  $j$  is the firm’s home county and 0 otherwise.<sup>12</sup> Proposition 1 implies that the coefficient of these variables, denoted as  $\beta_{Home}$  ( $\beta_{Dist}$  in the alternative specification) 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  $Home$  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 that  $\beta_{Distx4I} < 0$ ).

The estimates for this specification, which is denoted as Model 1, are presented in Table 4. Consistent with Proposition 1,  $\hat{\beta}_{Home}$  (the hat denotes an estimate) and  $\hat{\beta}_{Homex4I}$  are positive and significant.

<sup>12</sup> We estimated additional models using a variable  $Dist$ , which is the number of kilometers between the center of county  $j$  and the center of the firm’s home county. These models are available upon request as an appendix to the paper.

**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%) <i>1,407 (5%)</i>	789 (9%) <i>376 (4%)</i>	552 (9%) <i>237 (4%)</i>	3,782 2020
Unknown	8,602 (32%)	3,822 (43%)	2,675 (43%)	15,099
Female	40%	36%	36%	38%
College	2%	3%	3%	2%
Managers (Highoccupation)	3%	4%	4%	4%
Specialized workers (Middleoccupation)	61%	57%	54%	59%
Age	32	32	32	32
Total	27,282	8,851	6,235	42,368

Thus, all else equal firms are more likely to locate in their home county.

We probe the importance of the industry entered by adding to Model 1 *Home* 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  $\beta_{Homex4I} > \beta_{Homex3I} > \beta_{Homex2I} > 0$ . All the coefficient estimates of this specification, denoted as Model 2 in Table 4, are significant and have the predicted signs and ordering.<sup>13</sup>

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 (e.g., Baptista et al., 2014). 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., 2011), then we also need to control for factors related to the performance of firms to reliably test 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 employees, which numerous studies have found is a proxy for the longevity of the firm.<sup>14</sup> 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.

<sup>13</sup> Specifically, the estimated coefficients for variables  $Homex4I$ ,  $Homex3I$ , and  $Homex2I$  are significantly different.

<sup>14</sup> 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.

**Table 3** Description of variables used in the estimations

Variable Name	Description
Home	Dummy variable equal to one if the county is the firm’s home county and 0 otherwise
Dist	Number of kilometers between the center of a county and the center of the firm’s home county
4I/3I/ 2I	Dummy variable equal to one if the firm enter its home (4/ 3 but not 4/ 2 but not 4 or 3) digit industry
Firmtenure (FT)	Log of the average number of years the firm’s founders worked at their prior employers
Highlevel (HL)	Dummy variable equal to 1 for firms with one or more founders that worked as managers in their prior employer
Multiplefounders (MF)	Dummy variable equal to 1 for firms with more than one founder
Regionaltenure (RT)	Log of the average number of years the firm’s founders worked in their home county
WD/ Workerdensity	Density of workers in the county in the firm’s chosen industry
PD/ Populationdensity	Density of all workers for each county
H4IR/ H3IR/ H2IR/ HOIR	Dummy variable equal to 1 for firms that entered in their home county and home (4-digit/ 3-digit but not 4/ 2-digit but not 3 or 4/ other) industry
H4I, H3I, H2I	Dummy variable equal to 1 for firms that did not enter in their home county but entered their home (4-digit/ 3-digit but not 4/ 2-digit but not 3 or 4/) industry
HR	Dummy variable equal to 1 for firms that entered in their home county
C1/C2/C3	Dummy variable equal to one for old colleagues of the founders hired on year 1/2/3
WHIR1/ WHIR2/ WHIR3	Dummy variables equal to one for workers from the firm’s home 4-digit industry and home county hired in years 1/ 2/ 3
Age	Age of the individual when hired
Female	Dummy variable equal to 1 for females
College	Dummy variable equal to 1 for college educated workers
Highoccupation	Dummy variable equal to 1 for managers
Middleoccupation	Dummy variable equal to 1 for specialized technical workers
UK_Age	Dummy variable equal to 1 if age is unknown
UK	Dummy variable equal to 1 if worker’s background is unknown
%OldColleagues	% of the firm’s hires in year 1 that were old colleagues
%Workersfromhomeindustry&region	% of the firm’s hires in year 1 that were workers from the home county and home 4-digit industry

We interact each of these four variables with *Home* and add these interactions to Model 2, leading to Model 3 in Table 4. Regarding the added interactions, all interactions with *Home* are significant; with *Multiplefounders*, *Regionaltenure*, and *Highlevel* having a positive coefficient estimate, implying that firms with multiple flounders, higher regional tenure, and high-level founders were more likely to locate 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 *Home*, which is negative and significant, which implies that firms with founders with more tenure at their prior employer were more likely to locate outside 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 founder’s home region (Buenstorf & Klepper, 2010; Figueiredo et al., 2002). 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 et al. [2002],

**Table 4** – Conditional Logit estimates

	Model 1	Model 2	Model 3	Model 4
Home	5.217*** (0.031)	5.044*** (0.035)	4.137*** (0.057)	4.557*** (0.058)
Home x 4I	1.070*** (0.048)	1.243*** (0.051)	1.127*** (0.052)	1.058*** (0.053)
Home x 3I		0.681*** (0.112)	0.579*** (0.115)	0.570*** (0.118)
Home x 2I		0.610*** (0.073)	0.546*** (0.076)	0.538*** (0.077)
Home x FT			-0.248*** (0.046)	-0.305*** (0.047)
Home x HL			0.513*** (0.050)	0.548*** (0.051)
Home x MF			0.358*** (0.050)	0.335*** (0.051)
Home x RT			0.624*** (0.046)	0.701*** (0.047)
WD x (1-H)				0.251*** (0.084)
PD x (1-H)				0.051*** (0.002)
WD x H				-0.036 (0.075)
PD x H				-0.040 (0.026)
Observations	10,236 × 275 = 2,814,900			
Log likelihood	-23,990.715	-23,942.129	-23,687.694	-23,066.233

\*\*\* significance at the 0.01 level; \*\* significance at the 0.05 level; \* significance at the 0.1 level. Standard errors are in parentheses

Variables: Home—dummy variable equal to 1 if the location corresponds to the home location of at least one founder; 4I—dummy variable equal to one if firm is in the same 4-digit industry that any one of the previous employers of the founders; 3I—dummy variable equal to one if firm is in the same 3-digit industry that any one of the previous employers of the founders, but not in the same 4-digit industry; 2I—dummy variable equal to one if firm is in the same 3-digit industry that any one of the previous employers of the founders, but not in the same 4- or 3-digit industry. FT, HL, MF, and RT, stand for Firmtenure, Highlevel, Multiplefounders, and Regionaltenure, respectively. Multiplication signs correspond to interactions

An Appendix shows results for alternative specifications of the model (using Distance: distance in km from home region to the firm's chosen region, rather than Home; and alternative computations to control for the fact that, due to computational limitations, we were unable to add the level variables for Firmtenure, Highlevel, Multiplefounders, and Regionaltenure in Model 4).

we interact each variable with *Home* and  $(1 - \text{Home})$  in Model 4 to allow these variables to have different effects for the firm's home county and elsewhere.

The coefficient estimates of *Workerdensity*  $\times$   $(1 - \text{Home})$  and *Populationdensity*  $\times$   $(1 - \text{Home})$  are both positive and significant whereas the coefficient estimates of *Workerdensity*  $\times$  *Home* and *Populationdensity*  $\times$  *Home* are both small and insignificant. Consistent with Figueiredo et 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.

To put the conditional logit estimates into perspective, Table 5 reports the fraction of firms that entered in their home county as well as in counties distanced

1–10, 10–20, 20–30, 30–40, and more than 40 km 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 km of their home county, and 8% entered over 40 km away from their home county. The differences in these percentages are striking, particularly regarding 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*

**Table 5** – Characteristics of the firms in the sample, by geographical localization (columns) and according to relationship with home 4-digit industry (rows)  
In every cell (apart for total column/row), shares correspond to percentages of row total

Geographical Localization Relation to home industry	Distance to Home County (km)						Total
	Home County	1–10	10.1–20	20.1–30	30.1–40	> 40	
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 industry, not 3/ 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%

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 *High-level*, *Multiplefounders*, and *Regionaltenure* are all positive, and the coefficient estimate of *Firmtenure* is negative.

#### 4.2 Hires

Next we test the predictions 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, and this holds independent of the industry entered. 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 6.

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.

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.<sup>15</sup> 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 6 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.<sup>16</sup>

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 6. 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,  $\beta_{H4IR}$  and  $\beta_{HR}$  should decline in year 2 and then decline further in year 3. These predictions are strongly supported by the results. Compared with year 1,  $\hat{\beta}_{H4IR}$  and  $\hat{\beta}_{HR}$  decline sharply in years 2 and 3, especially  $\hat{\beta}_{H4IR}$ , which is not much larger and is no longer significantly different than  $\hat{\beta}_{H4I}$  by year 3.

<sup>15</sup> 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.

<sup>16</sup> 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 possibly the hiring of 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 7 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 (different) 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. These percentages help put the probit estimates into 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. Our 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  $\beta_{HR}$  and not  $\beta_{H4IR}$  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  $\beta_{HR}$  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 8.

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 our assumption that industry does not matter,  $\hat{\beta}_{H4I}$  is positive and significant and  $\hat{\beta}_{H4IR}$  is significantly greater

**Table 6** – 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	27,282	27,282	27,282	27,277 [1]	8,770 [1]	6,057 [1]
Log Pseudolikelihood	-16,222.195	-16,037.695	-15,559.331	-15,268.338	-2,304.4512	-1,037.7086

\*\*\* significance at the 0.01 level; \*\* significance at the 0.05 level; \* significance at the 0.1 level. Standard errors, robust to intra-firm correlation, are in parenthesis

[1] Observations were dropped due to collinearity

Variables: H4IR/ H3IR/ H2IR/ HOIR—dummy variable equal to 1 for firms that entered in their home county and home (4-digit/ 3-digit but not 4/ 2-digit but not 3 or 4/ other) industry; H4I, H3I, H2I—dummy variable equal to 1 for firms that did not enter in their home county but entered their home (4-digit/ 3-digit but not 4/ 2-digit but not 3 or 4/) industry; HR—dummy variable equal to 1 for firms that entered in their home county; Firmtenure—log of the average number of years the firm’s founders worked at their prior employers; Highlevel—dummy variable equal to 1 for firms with one or more founders that worked as managers in their prior employer; Multiplefounders—dummy variable equal to 1 for firms with more than one founder; Regionaltenure—log of the average number of years the firm’s founders worked in their home county

than  $\hat{\beta}_{HR}$  (although this is not necessarily inconsistent with our predictions).

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.<sup>17</sup>

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,

<sup>17</sup> It should be noted that *H4IR* is significantly different from *H3IR* and *H2IR*.

**Table 7** – 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: Unknown background

Shares correspond to the fraction of workers with a given relationship with their founder, hired by a certain group of founders in a given year (i.e., in their first year, 50% of the workers hired by founders that entered their home region and industry were old colleagues). Shares in italic correspond to shares of workers excluding old colleagues

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.

The patterns concerning the workers from the home industry and region in Table 7 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.

#### 4.3 Worker Hazard of Exit

We next test our predictions 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

**Table 8** – 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	17,925	17,925	17,925	17,764 [1]	7,991 [1]	5,895 [1]
Log Pseudolikelihood	-4,211.3007	-4,185.28	-4,145.8754	-3,914.1788	-1,631.211	-1,247.0042

\*\*\* significance at the 0.01 level; \*\* significance at the 0.05 level; \* significance at the 0.1 level. Standard errors, robust to intra-firm correlation, are in parenthesis

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

Variables: H4IR/ H3IR/ H2IR/ HOIR—dummy variable equal to 1 for firms that entered in their home county and home (4-digit/ 3-digit but not 4/ 2-digit but not 3 or 4/ other) industry; H4I, H3I, H2I—dummy variable equal to 1 for firms that did not enter in their home county but entered their home (4-digit/ 3-digit but not 4/ 2-digit but not 3 or 4/) industry; HR—dummy variable equal to 1 for firms that entered in their home county; Firmtenure—log of the average number of years the firm’s founders worked at their prior employers; Highlevel—dummy variable equal to 1 for firms with one or more founders that worked as managers in their prior employer; Multiplefounders—dummy variable equal to 1 for firms with more than one founder; Regionaltenure—log of the average number of years the firm’s founders worked in their home county

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  $\beta_{C1}$ ,  $\beta_{C2}$ , and  $\beta_{C3}$  will all be negative and would be expected to be smaller than  $\beta_{WHIR1}$ ,  $\beta_{WHIR2}$ , and  $\beta_{WHIR3}$ , respectively. We expect that, over time, an increasing fraction of the workers hired from the home industry and region will not be known to the founder, implying 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 9 under Model 1. First,  $\hat{\beta}_{C1}$ ,  $\hat{\beta}_{C2}$ , and  $\hat{\beta}_{C3}$  all are 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 being 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

**Table 9** – Cox proportional hazard model of the annual hazard of worker exit

	Model 1	Model 2	Model 3
C1	-0.617 <sup>***</sup> (0.025)	-0.587 <sup>***</sup> (0.026)	-0.552 <sup>***</sup> (0.025)
C2	-0.291 <sup>***</sup> (0.064)	-0.271 <sup>***</sup> (0.064)	-0.247 <sup>***</sup> (0.064)
C3	-0.358 <sup>***</sup> (0.117)	-0.341 <sup>***</sup> (0.116)	-0.359 <sup>***</sup> (0.114)
WHIR1	-0.319 <sup>***</sup> (0.050)	-0.305 <sup>***</sup> (0.041)	-0.268 <sup>***</sup> (0.049)
WHIR2	-0.176 <sup>***</sup> (0.063)	-0.152 <sup>**</sup> (0.061)	-0.125 <sup>**</sup> (0.061)
WHIR3	-0.094(0.072)	-0.071(0.072)	-0.021(0.069)
Age	-0.002 <sup>***</sup> (0.001)	-0.002 <sup>***</sup> (0.001)	-0.003 <sup>***</sup> (0.001)
Female	-0.123 <sup>***</sup> (0.019)	-0.118 <sup>***</sup> (0.019)	-0.077 <sup>***</sup> (0.018)
College	0.036(0.050)	0.023(0.046)	0.027(0.048)
Highoccupation	-0.162 <sup>***</sup> (0.044)	-0.175 <sup>***</sup> (0.045)	-0.186 <sup>***</sup> (0.043)
Middleoccupation	-0.049 <sup>***</sup> (0.016)	-0.052 <sup>***</sup> (0.016)	-0.053 <sup>***</sup> (0.016)
UK_Age	-0.078(0.065)	-0.073(0.066)	-0.121 <sup>*</sup> (0.068)
UK	0.156 <sup>***</sup> (0.016)	0.157 <sup>***</sup> (0.016)	0.145 <sup>***</sup> (0.015)
Firmtenure		-0.052 <sup>***</sup> (0.019)	-0.046 <sup>**</sup> (0.013)
Highlevel		0.035(0.021)	0.014(0.021)
Multiplefounders		-0.012(0.020)	-0.009(0.020)
Regionaltenure		-0.040 <sup>*</sup> (0.019)	-0.023(0.018)
Year Dummies	Included	Included	Included
2-Digit Industry Dummies			Included
District Dummies			Included
Subjects	40,993 [1]		
Failures	21,914		
Observations	97,420		
Log likelihood	-22,0960.87	-22,0897.22	-22,0676.25

[1] Out of the 42,368 initial sample on Table 2, 1375 subjects were hired on 2006, the last year of analyses and were thus excluded from the analyses

<sup>\*\*\*</sup>significance at the 0.01 level; <sup>\*\*</sup>significance at the 0.05 level; <sup>\*</sup>significance at the 0.1 level. Standard errors, robust to intra-firm correlation, in parenthesis

Variables: C1/C2/C3—dummy variable equal to one for old colleagues of the founders hired on year 1/2/3; WHIR1/ WHIR2/ WHIR3—dummy variables equal to one for workers from the firm's home 4-digit industry and home county hired in years 1/ 2/ 3; Age—Age of the individual when hired; Female—dummy variable equal to 1 for female; College—dummy variable equal to 1 for college educated workers; Highoccupation—dummy variable equal to 1 for managers; Middleoccupation—dummy variable equal to 1 for specialized technical workers; UK\_Age—dummy variable equal to 1 if worker's background is unknown; Firmtenure—log of the average number of years the firm's founders worked at their prior employers; Highlevel—dummy variable equal to 1 for firms with one or more founders that worked as managers in their prior employer; Multiplefounders—dummy variable equal to 1 for firms with more than one founder; Regionaltenure—log of the average number of years the firm's founders worked in their home county; %OldColleagues—% of the firm's hires in year 1 that were old colleagues; %Workersfromhomeindustry&region—% of the firm's hires in year 1 that were workers from the home county and home 4-digit industry

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 the four controls for firm performance to Model 1. 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.

#### 4.4 Firm Hazard of Exit

Last, we test the predictions 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 10 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 0.10 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 (empirical work on new firm survival has consistently shown that firms that enter the same industry as the founders' parent firm have a greater likelihood of survival, regardless of location—cf. Bosma et al., 2004; Baptista et al., 2014). 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 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.<sup>18</sup> 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.

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

<sup>18</sup> In terms of significant differences,  $H4IR$ ,  $H3IR$ ,  $H2IR$  are significantly different than  $HOIR$ .

**Table 10** – Cox proportional hazard model of the annual hazard of firm exit

	Model 1	Model 2	Model 3	Model 4	Model 5
<i>%OldColleagues</i>					-0.004*** (0.001)
<i>%Workersfromhomeindustry&amp;region</i>					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	10,236				
Failures	4177				
Observations	42,904				
Log likelihood	-36,278.932	-36,269.663	-36,133.491	-36,007.127	-35,983.737

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

Variables: *%OldColleagues*—% of the firm’s hires in year 1 that were old colleagues; *%Workersfromhomeindustry&region*—% of the firm’s hires in year 1 that were workers from the home county and home 4-digit industry; H4IR/ H3IR/ H2IR/ HOIR—dummy variable equal to 1 for firms that entered in their home county and home (4-digit/ 3-digit but not 4/ 2-digit but not 3 or 4/ other) industry; H4I, H3I, H2I—dummy variable equal to 1 for firms that did not enter in their home county but entered their home (4-digit/ 3-digit but not 4/ 2-digit but not 3 or 4/) industry; HR—dummy variable equal to 1 for firms that entered in their home county; Firmtenure—log of the average number of years the firm’s founders worked at their prior employers; Highlevel—dummy variable equal to 1 for firms with one or more founders that worked as managers in their prior employer; Multiplefounders—dummy variable equal to 1 for firms with

industry and region, they would still be expected to correlate with 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 of exit would be 16% lower. However, the same 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.

Finally, we performed one last analysis. Our theoretical remarks imply that firms that entered in their home region have a lower hazard of exit 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 (worker hazard of exit) the variables *H4IR*, *HR*, and *H4I*, and then estimated this model without and with the variables *C1*, *C2*, *C3*, *WHIR1*, *WHIR2*, and *WHIR3*. In the absence of these variables,  $\hat{\beta}_{H4IR}$ ,  $\hat{\beta}_{HR}$  and  $\hat{\beta}_{H4I}$  are, respectively, equal to -0.141\*\*\* (0.030), -0.063\*\* (0.029),

and  $-0.018$  (0.039).<sup>19</sup> Consistent with our theoretical remarks,  $\hat{\beta}_{H4IR}$  and  $\hat{\beta}_{HR}$  are negative and significant, and  $\hat{\beta}_{H4I}$  is insignificant, with  $\hat{\beta}_{H4IR}$  significantly smaller than  $\hat{\beta}_{HR}$ .<sup>20</sup> Furthermore, when  $C1$ ,  $C2$ ,  $C3$ ,  $WHIR1$ ,  $WHIR2$ , and  $WHIR3$  are added to the specification, the coefficient estimates are close to 0 and insignificant.<sup>21</sup> Thus, consistent with the out theoretical remarks, 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.

## 5 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.

<sup>19</sup> Standard errors are in parentheses.

<sup>20</sup> Given that the percentages of hires that were old colleagues and workers from the home industry and region are similar for firms that entered in their home county but not their home industry, and home industry but not home county (in contrast to what we predicted), it might have been expected that  $HR$  and  $H4I$  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 compared with the other three groups of firms. This result suggests that firms that do not enter in their home region, whether they entered their home industry or not, do not benefit by hiring old colleagues and workers from the home industry and region as much as firms that enter in their home region.

<sup>21</sup> When just  $C1$ ,  $C2$ ,  $C3$  are added,  $H4IR$ ,  $HR$  and  $H4I$  equal  $-0.044$  (0.029),  $-0.022$  (0.029),  $0.023$  (0.024), and when  $WHIR1$ ,  $WHIR2$ , and  $WHIR3$  are added as well then  $H4IR$ ,  $HR$  and  $H4I$  equal  $-0.023$  (0.029),  $-0.018$  (0.029),  $0.029$  (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. As mentioned in the previous footnote, these estimates are available from the authors upon request.

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 our predictions. 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.

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 with 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.

It can therefore be suggested that an important (though by no means exclusive) force driving the buildup of industry clusters is the tendency of new firms—in particular spinoffs—to locate near their

parent firms in order to access human capital that matches better with the new firms ideas regarding markets, technology, and organization. As pointed out earlier, a significant body of literature reveals a similar tendency of all types of new firms to locate close to where their founders previously worked and resided (Berchicci et al., 2011; Buenstorf & Klepper, 2009; Dahl & Sorenson, 2010a; Figueiredo et al., 2002; Klepper, 2007, 2010). 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. When those geographic roots comprise a set of already successful incumbents/parent firms, the local buildup of successful new firms will contribute to cluster-type agglomerations. 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.

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, regions with different levels and mix of economic activities, for firms with a single employee-owner, which did not alter our conclusions. Since our data ends in 2006 (before the emergence of the Great Recession, which hit Portugal particularly hard), we examined data for the only year available to us following the economic downturn (2016). Again, our findings did not alter our conclusions. The results are also robust to restricting the sample to firms for which the parent firm of the founder was still in operation when the new firm was started. Our findings further resonate with studies of firm location and survival in other countries. Dahl and Sorenson (2010a) 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. (2012) find that new Brazilian firms are more likely to locate close to their parent firm if they entered the same industry as their parent. Buenstorf and Klepper (2010) 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 & Sorenson, 2009), potential sources of capital (Michelacci & Silva, 2007), and location-specific public policy incentives (Leitão and Baptista, 2009) 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 the question of why all firms—or, at least, all spinoffs—don't enter close to home. This is a central question for our analysis. A key assumption for our analysis is that 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.

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? 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. Furthermore, 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 in the home region of founders 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.

An additional factor that might affect worker mobility—in particular between incumbents and spinoffs—is the existence of non-compete clauses in workers' contracts. Marx et al. (2009) find that the enforcement of non-compete agreements significantly attenuates labor mobility, and the effect is more severe for those workers with firm-specific skills. Using data for US states, Samila and Sorensen (2011) find that the enforcement of non-compete clauses significantly impedes entrepreneurship, innovation, and employment growth. Similar to most European countries, Portuguese law does allow for non-compete agreements to be enforced after a worker exits a firm (Sousa, 2020). However, since non-compete clauses are most likely to hinder labor mobility, their impact should counteract the effects being tested, so the validity of our results is not affected.

Certainly knowledge possessed by new firms at the time of entry is not limited to knowledge about prospective hires. Surely founders also acquire

knowledge about how to organize and operate a firm in their prior industry, as well as knowledge about access to financing sources, local suppliers, and the institutional and public policy environment. 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.

**Acknowledgements** We thank Serguey Braguinsky, David Greenstreet, Brian Kovak, Lee Branstetter, and participants in several seminars and conferences for helpful comments.

Cristina Carias acknowledges support from a Carnegie Mellon-Portugal Program doctoral scholarship. Support for this research was provided by Fundação para a Ciência e a Tecnologia through the Carnegie Mellon Portugal Program (Project Ref. CMUPT/Etech/0036/2008) and through Project Ref. PTDC/EGE-OGE/31152/2017.

We thank the Portuguese Ministry of Labor, Solidarity and Social Security for granting us access to the data.

## Appendix

In this appendix, we present alternative specifications for the conditional logit model in Table 4 of our manuscript.

We use the variable *Dist* (distance from the home country to the chosen county in km) instead of *Home*. Proposition 1 implies that  $\beta_{Distx4I} < 0$ . As seen below, consistent with Proposition 1,  $\beta_{Dist}$  and  $\beta_{Distx4I}$  are negative and significant. Proposition 1 also implies that  $\beta_{Homex4I} > \beta_{Homex3I} > \beta_{Homex2I} > 0$  and  $\beta_{Distx4I} < \beta_{Distx3I} < \beta_{Distx2I} < 0$ . Again, as seen from the table in the main manuscript and the results below, results are as predicted by the theory (in model 2 all coefficients are statistically different at the 0.1 significance level; in model 3, the estimated coefficient of  $\beta_{Distx4I}$  is statistically different than  $\beta_{Distx2I}$ , but  $\beta_{Distx4I}$  is not statistically different than  $\beta_{Distx3I}$ ).

## Tables 11 and 12

The estimations with the conditional logit, and again with the probit, we find that: entrepreneurs that enter in the same industry as their former employers are more likely to choose the home region; entrepreneurs that enter the same 3-digit and 2-digit industry of their employers are more likely to enter the home region as entrepreneurs that enter a different industry; entrepreneurs that enter the same 3-digit and 2-digit industry are less likely to enter the home region than entrepreneurs that enter the same 4-digit industry (the differences between the estimated coefficients for 4I, and 3I are statistically significant; but the difference for the estimated coefficients for 3I and 2I are not statistically significant, across Models 2–4); higher employer tenure, number of founders, and the hierarchical level of the employee also increase the likelihood of choosing the location of the previous employer. These results are in line with our predictions.



**Table 11** – Conditional Logit estimates

	Model 1	Model 2	Model 3
Dist	-0.062 <sup>***</sup> (0.001)	-0.056 <sup>***</sup> (0.001)	-0.022 <sup>***</sup> (0.001)
Dist x 4I	-0.056 <sup>***</sup> (0.002)	-0.063 <sup>***</sup> (0.002)	-0.049 <sup>***</sup> (0.002)
Dist x 3I		-0.030 <sup>***</sup> (0.003)	-0.020 <sup>***</sup> (0.003)
Dist x 2I		-0.023 <sup>***</sup> (0.002)	-0.015 <sup>***</sup> (0.002)
Dist x FT			0.023 <sup>***</sup> (0.002)
Dist x HL			-0.016 <sup>***</sup> (0.001)
Dist x MF			-0.022 <sup>***</sup> (0.001)
Dist x RT			-0.041 <sup>***</sup> (0.002)
Observations	10,236 × 275 = 2,814,900		
Log likelihood	-25,052.643	-24,928.842	-23,813.088

<sup>\*\*\*</sup> significance at the 0.01 level; <sup>\*\*</sup> significance at the 0.05 level; <sup>\*</sup> significance at the 0.1 level. Standard errors are in parentheses. Variables: 4I—dummy variable equal to one if firm is in the same 4-digit industry that any one of the previous employers of the founders; 3I—dummy variable equal to one if firm is in the same 3-digit industry that any one of the previous employers of the founders, but not in the same 4-digit industry; 2I—dummy variable equal to one if firm is in the same 3-digit industry that any one of the previous employers of the founders, but not in the same 4- or 3-digit industry; Distance: distance (in km) from home region to the firms chosen region. FT, HL, MF, and RT, stand for Firmtenure, Highlevel, Multiplefounders, and Regionaltenure, respectively. Multiplication signs correspond to interactions. In the conditional logit estimations presented in our manuscript we were unable to include the level variables Firmtenure, Highlevel, Multiplefounders, and Regionaltenure due to computational reasons. As a result, we also include here an alternative econometric model. Instead of estimating a conditional logit model in which each possible location is a choice for the entrepreneur, we estimated the factors influencing the likelihood that the entrepreneur would choose the home region, given the entrepreneur's characteristics and choice of new industry, via a probit model. [Table 12](#) shows our estimation results

**Table 12** – Probit estimates

	Model 1	Model 2	Model 3	Model 4
4I	0.563 <sup>***</sup> (0.027)	0.660 <sup>***</sup> (0.028)	0.602 <sup>***</sup> (0.030)	0.622 <sup>***</sup> (0.031)
3I		0.411 <sup>***</sup> (0.065)	-0.354 <sup>***</sup> (0.066)	0.345 <sup>***</sup> (0.069)
2I		0.334 <sup>***</sup> (0.043)	0.309 <sup>***</sup> (0.043)	0.321 <sup>***</sup> (0.045)
FT			0.139 <sup>**</sup> (0.015)	0.147 <sup>***</sup> (0.015)
HL			0.273 <sup>**</sup> (0.028)	0.311 <sup>***</sup> (0.030)
MF			0.248 <sup>***</sup> (0.028)	0.269 <sup>***</sup> (0.029)
Constant	0.189 <sup>***</sup> (0.017)	0.091 <sup>***</sup> (0.020)	-0.244 <sup>***</sup> (0.029)	0.119(0.170)
Observations				
Log likelihood	-6313.8342	-6268.7749	-6119.4741	-5645.3047

<sup>\*\*\*</sup> significance at the 0.01 level; <sup>\*\*</sup> significance at the 0.05 level; <sup>\*</sup> significance at the 0.1 level

Standard errors in parentheses. FT, HL, MF, and RT, stand for Firmtenure, Highlevel, Multiplefounders, and Regionaltenure, respectively

## References

- Agarwal, R., Echambadi, R., Franco, A. M., & Sarkar, M. B. (2004). Knowledge transfer through inheritance: Spin-out generation, development, and survival. *Academy of Management Journal*, 47(4), 501–522.
- Almeida, P., & Kogut, B. (1999). Localization of knowledge and the mobility of engineers in regional networks. *Management Science*, 45(7), 905–917.
- Baptista, R., Karaöz, M., & Mendonça, J. (2014). The impact of human capital on the early success of necessity versus opportunity-based entrepreneurs. *Small Business Economics*, 42(4), 831–847.

- Baptista, R., Lima, F., & Torres Preto, M. (2013). Entrepreneurial skills and workers' wages in small firms. *Small Business Economics*, 40(2), 309–323.
- Berchicci, L., King, A., & Tucci, C. L. (2011). Does the apple always fall close to the tree? The stay-leave choice of spinouts. *Strategic Entrepreneurship Journal*, 5(2), 120–136.
- Buenstorf, G., & Klepper, S. (2009). Heritage and Agglomeration: The Akron Tire Cluster Revisited. *Economic Journal*, 119, 705–733.
- Buenstorf, G., & Klepper, S. (2010). Why does entry cluster geographically? Evidence from the US tire industry. *Journal of Urban Economics*, 68(2), 103–114.
- Buenstorf, G., & Costa, C. (2018). Drivers of spin-off performance in industry clusters: Embodied knowledge or embedded firms?. *Research Policy*, 47(3), 663–673.
- Cabral, L., & Mata, J. (2003). On the evolution of the firm size distribution: Facts and theory. *American Economic Review*, 93(4), 1075–1090.
- Cardon, M. S., & Stevens, C. E. (2004). Managing human resources in small organizations: What do we know? *Human Resource Management Review*, 14(3), 295–323.
- Chatterji, A., Glaeser, E., & Kerr, W. (2014). Clusters of entrepreneurship and innovation. *Innovation Policy and the Economy*, 14(1), 129–166.
- Cheyre, C., Klepper, S., & Veloso, F. (2015). Spinoffs and the mobility of US merchant semiconductor inventors. *Management Science*, 61(3), 487–506.
- Coad, A., Daunfeldt, S.-O., Johansson, D., & Wennberg, K. (2014). Whom do high-growth firms hire? *Industrial and Corporate Change*, 23(1), 293–327.
- Coad, A., Nielsen, K., & Timmermans, B. (2017). My first employee: An empirical investigation. *Small Business Economics*, 48(1), 25–45.
- Cooper, A. C., Javier Gimeno-Gascon, F., & Woo, C. Y. (1994). Initial human and financial capital as predictors of new venture performance. *Journal of Business Venturing*, 9(5), 371–395.
- Costa, C. & Baptista, R. (2015) Organizational heritage and entrepreneurship: Steven Klepper's theories reflected in the emergence and growth of the plastic molds industry in Portugal. In: Baptista R., & Leitão J. (Eds.), *Entrepreneurship, human capital, and regional development*. International Studies in Entrepreneurship, vol 31. Springer, Cham. [https://doi.org/10.1007/978-3-319-12871-9\\_14](https://doi.org/10.1007/978-3-319-12871-9_14)
- Dahl, M. S., & Klepper, S. (2015). Whom do new firms hire? *Industrial and Corporate Change*, 24(4), 819–836.
- Dahl, M., & Sorenson, O. (2009). The embedded entrepreneur. *European Management Review*, 6, 172–181.
- Dahl, M. S., & Sorenson, O. (2010a). The social attachment to place. *Social Forces*, 89(2), 633–658.
- Dahl, M., & Sorenson, O. (2010b). The migration of technical workers. *Journal of Urban Economics*, 67, 33–45.
- Delgado, M., Porter, M. E., & Stern, S. (2010). Clusters and entrepreneurship. *Journal of Economic Geography*, 10(4), 495–518.
- Delgado, M., Porter, M. E., & Stern, S. (2014). Clusters, convergence, and economic performance. *Research Policy*, 43(10), 1785–1799.
- Figueiredo, O., Guimarães, P., & Woodward, D. (2002). Home-field advantage: Location decisions of Portuguese entrepreneurs. *Journal of Urban Economics*, 52, 341–361.
- Garvin, D. A. (1983). Spin-offs and the new firm formation process. *California Management Review*, 25, 3–20.
- Geroski, P. A., Mata, J., & Portugal, P. (2010). Founding conditions and the survival of new firms. *Strategic Management Journal*, 31(5), 510–529.
- Hayton, J. C. (2003). Strategic human capital management in SMEs: An empirical study of entrepreneurial performance. *Human Resource Management*, 42(4), 375–391.
- Helfat, C. E., & Lieberman, M. B. (2002). The birth of capabilities: Market entry and the importance of pre-history. *Industrial and Corporate Change*, 11(4), 725–760.
- Kamm, J., Shuman, J. C., Seeger, J. A., & Nurick, A. J. (1990). Entrepreneurial teams in new venture creation: A research agenda. *Entrepreneurship Theory & Practice*, 14(4), 7–17.
- Kaplan, S. N., Sensoy, B. A., & Strömberg, P. (2009). 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, 75–115.
- Klepper, S. (2007). Disagreements, spinoffs, and the evolution of Detroit as the capital of the U.S. automobile industry. *Management Science*, 53, 616–631.
- Klepper, S. (2009). Spinoffs: A review and synthesis. *European Management Review*, 6(3), 159–171.
- Klepper, S. (2010). The origin and growth of industry clusters: The making of silicon valley and Detroit. *Journal of Urban Economics*, 67, 15–32.
- Klepper, S. (2011). Nano-economics, spinoffs, and the wealth of regions. *Small Business Economics*, 37(2), 141–154.
- Lazear, E. P. (2009). Firm-specific human capital: A skill-weights approach. *Journal of Political Economy*, 117(5), 914–940.
- Leitão, João, & Rui, Baptista, (Eds). (2009). *Public policies for fostering entrepreneurship: A European perspective* (Vol. 22). Springer Science & Business Media.
- Marx, M., Strumsky, D., & Fleming, L. (2009). Mobility, skills, and the Michigan non-compete experiment. *Management science*, 55(6), 875–889.
- Michelacci, C., & Silva, O. (2007). Why so many local entrepreneurs? *Review of Economics and Statistics*, 89, 415–433.
- Muendler, M. A., Rauch, J. E., & Tocoian, O. (2012). Employee spinoffs and other entrants: Stylized facts from Brazil. *International Journal of Industrial Organization*, 30(5), 447–458.
- Muendler, M.-A., & Rauch, J. E. (2009). *Employee spinoffs and industrial clustering: Evidence from Brazil*. Unpublished manuscript, University of California, San Diego .
- Neal, D. (1995). Industry-specific human capital: Evidence from displaced workers. *Journal of Labor Economics*, 13(4), 653–677.
- Neffke, F., & Henning, M. (2013). Skill relatedness and firm diversification. *Strategic Management Journal*, 34(3), 297–316.
- Neffke, F., Hartog, M., Boschma, R., & Henning, M. (2018). Agents of structural change: The role of firms and entrepreneurs in regional diversification. *Economic Geography*, 94(1), 23–48.
- Niels, B., Mirjam van Praag, A., Thurik, R., & de Wit, G. (2004). The value of human and social capital investments for the business performance of startups. *Small Business Economics*, 23, 227–236.
- Pflueger, M., & Suedekum, J. (2008). A synthesis of footloose-entrepreneur new economic geography models: When is agglomeration smooth and easily reversible? *Journal of Economic Geography*, 8, 39–54.

- Rocha, V., Carneiro, A., & Varum, C. (2018). Leaving employment to entrepreneurship: The value of co-worker mobility in pushed and pulled-driven start-ups. *Journal of Management Studies*, 55(1), 60–85.
- Rogerson, R., Shimer, R., & Wright, R. (2005). Search-theoretic models of the labor market: A survey. *Journal of Economic Literature*, 43(4), 959–988.
- Rosenthal, S. S., & Strange, W. C. (2004). Evidence on the nature and sources of agglomeration economies. In: J. Vernon Henderson, & Jacques-François Thisse (Eds.), *Handbook of regional and urban economics*. Elsevier, vol 4, Chap 49, pp. 2119–2171. [https://doi.org/10.1016/S1574-0080\(04\)80006-3](https://doi.org/10.1016/S1574-0080(04)80006-3)
- Samila, S., & Sorenson, O. (2011). Noncompete covenants: Incentives to innovate or impediments to growth. *Management Science*, 57(3), 425–438.
- Sousa, D. A. (2020). Impacts of social networks in the employment environment—from the traditional subjects to the particular case of employment non-compete clauses. In: Carby-Hall, J. & Méndez. L.M. (Eds.) *Labour Law and the Gig Economy: Challenges Posed by the Digitalisation of Labour Processes* (1st ed.), pp. 125–138. Routledge. <https://doi.org/10.4324/9781003033721>
- Timmermans, B. (2009). The effect of previous co-worker experience on firm survival. In: *DRUID Summer Conference 2009*.
- Ucbasaran, D., Lockett, A., Wright, M., & Westhead, P. (2003). Entrepreneurial founder teams: Factors associated with member entry and exit. *Entrepreneurship Theory & Practice*, 28, 107–128.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.