# Spin-offs: Why geography matters

Apostolos Baltzopoulos, Pontus Braunerhjelm and Ioannis Tikoudis

RESEARCH NETWORK DEBATE

SWEDISH ENTREPRENEURSHIP
FORUM

# Working Papers Series from Swedish Entrepreneurship Forum

In 2009 Swedish Entrepreneurship Forum started publishing a new series of Working Papers. These are available for download on www.entreprenorskapsforum.se, and are part of our ambition to make quality research available to a wider audience, not only within the academic world.

Scholars from different disciplines are invited to publish academic work with the common denominator that the work has  policy relevance within the field of entrepreneurship, innovation and SMEs.

The working papers published in this series have all been discussed at academic seminars at the research institution of the author.

## ABOUT SWEDISH ENTREPRENEURSHIP FORUM

Swedish Entrepreneurship Forum is the leading Swedish network organization for generating and transferring policy relevant research in the field of entrepreneurship and small enterprise development.

Swedish Entrepreneurship Forum is a network organization with the aim

- to serve as a bridge between the small business research community and all agents active in development of new and small enterprises.
- to initiate and disseminate research relevant to policy in the fields of entrepreneurship, innovation and SME.
- to offer entrepreneurship researchers a forum for idea sharing, to build national and international networks in the field and to bridge the gap between research and practical application.

Find out more on www.entreprenorskapsforum.se

# Spin-offs: Why geography matters

*Accepted for publication in Journal of Economic Geography*

Apostolos Baltzopoulos[*]

Pontus Braunerhjelm[**]

Ioannis Tikoudis[#]

November 2014

## Abstract

Based on unique data covering individuals, firms, industries, and regions for the 1999-2005 period, we contribute with new knowledge concerning the impact of regional variables on spin-offs. Implementing a large number of controls, as well as different estimation techniques and robustness tests, we show that Jacobian externalities have a positive effect on spin-offs. Moreover, using an entropy measure to disentangle unrelated and related variety, we conclude that the effect is confined to related variety. These findings are likely to be associated with strong welfare effects: a standard-deviation increase (decrease) in related (unrelated) variety increases spin-off propensity by approximately 25%. Other variables are shown to have economic effects of a similar magnitude but may have a different effect across sectors. Sensitivity analyses indicate that the impact of other determinants proposed in the literature (e.g., Marshallian externalities and scale effects) is too small to be detected.

---

[*] Department of Industrial Economics and Management, KTH Royal Institute of Technology and The Swedish Competition Authority, Stockholm, Sweden (apostolos.baltzopoulos@kkv.se).

[**] Department of Industrial Economics and Management, KTH Royal Institute of Technology, 100 44 Stockholm, Sweden. Corresponding author: Pontus Braunerhjelm (pontus.braunerhjelm@indek.kth.se).

[#] Department of Spatial Economics, VU University Amsterdam, De Boelelaan 1105, 1081HV, and Tinbergen Institute, Gustav Mahlerplein 117, 1082MS Amsterdam (i.tikoudis@vu.nl).

# 1. Introduction

Spin-offs, i.e. firms formed by employees leaving an incumbent, have been shown to perform considerably better than other start-ups: they have a higher survival rate, grow faster, generate more radical innovations, and seem to be critically important elements of cluster growth and dynamics.[1] Nevertheless, despite these insights and the societal value shown to be associated with the performance of spin-offs, knowledge of the underlying factors that trigger this type of entrepreneurial entry is still evolving.

In this paper we address how regional variables, such as market size, knowledge endowment, industrial composition (specialisation, diversification, and entropy), entrepreneurial environment and culture, influence an employee's decision to spin off a new venture. We have access to a unique, matched employer/employee data set for the entire Swedish private sector economy, and we link these data to regional-, individual-, and firm-level data, which allows us to control for the characteristics of not only the founder of the new firm but also the incumbent parent firm. In addition, we examine all spin-offs across all industries of the economy rather than spin-offs for a selected number of industries.

Our approach also allows us to compare the decision to become an entrepreneur with other occupational alternatives. Prior studies that focus on the incidence of spin-offs alone fail to consider individuals with similar backgrounds and characteristics who choose to stay in their current job, exit, or switch to a different employer rather than start a new venture. We believe that our findings should be of interest not only to the academic community but also to policy makers involved in designing policies to promote regional entrepreneurial entry.

From the perspective of the individual, a spin-off is commonly assumed to be either an exogenous process or modelled as a random utility (profit) maximisation process. Klepper and Thompson's (2007; 2010) disagreement model is however a notable exception. Theoretically, we build on the Klepper and Thompson framework, but we seek to augment this framework with a broader set of variables that are likely to influence an employee's choice between remaining an employee and spinning off a new firm. We side with the views previously set forth in the 1980s arguing that the decision to start a firm may be influenced by a combination of multilevel factors (Gartner 1985; Low and MacMillan, 1988). Hence, we take a more comprehensive approach than most other studies by focussing on the regional dimension while simultaneously taking industrial, firm, and individual factors into account.

---

[1]See Walsh et al. (1996), Klepper (2001; 2002; 2005; 2007), Gompers et al. (2005), Klepper and Sleeper (2005), Hellman (2006), Elfenbein et al. (2010), and Andersson and Klepper (2013).

Hyytinen and Maliranta (2008) represents one of the first attempts to conduct a multi-level analysis of spin-offs. Using Finnish data, that study examines a series of individual-, firm-, and industry-specific variables but disregards regional variables. The absence of regional variables is a conspicuous feature of the first wave of spin-off analyses, in which the primary concern was to understand the evolution of new industries and the role played by spin-offs (Klepper 1996; Eriksson and Kuhn 2006; Dahl and Reichstein, 2013). These studies provide strong evidence that spin-offs from successful incumbents are the highest performing entrants and can play instrumental roles in the emergence of new industries.

Subsequently, more attention has been directed towards industrial dynamics, life-cycle patterns, and feedback among agglomeration economies, entry, and social factors (Boschma and Wenting 2007; Frenken et al., 2011). More precisely, the empirical literature has thoroughly analysed how regional agglomeration economies (localisation and urbanisation) affect entrepreneurial entry and spin-offs compared with *home field advantages* (Sorenson and Audia 2000; Figueiredo et al. 2002a; Buenstorf and Klepper 2007, 2010a).[2] The advantages of remaining in a (home) region – where the entrepreneur can exploit superior information and previously established networks – are claimed to outweigh the benefits of agglomeration economies. In addition, the importance of social ties and proximity to family and friends has been stressed in this context (Dahl and Sorenson, 2009). Hence, higher entry rates in more agglomerated industries may simply reflect the presence of a larger number of firms (i.e., a region's birth potential) and stronger entrepreneurial opportunities rather than agglomeration economies. Alternatively, such economies are more valid for firms that have decided to move across regions or nations, e.g., foreign or horizontally organised firms (Figueiredo et al. 2002b; Buenstorf and Klepper, 2010b).

Our results suggest that the regional context has a major impact on the propensity of an individual to spin-off a new firm. Implementing an entropy measure (Frenken et al. 2007), we conclude that local *related industrial variety* (RV) exerts a positive impact on spin-offs; no such effect has previously been detected for diversity. This finding corroborates the submarket hypothesis (Buenstorf and Klepper 2010). However, *unrelated variety* (UV) is shown to have a negative impact on spin-offs. We also examine the economic significance of the findings mentioned above. Rough calculations suggest that the spin-off rate can increase by 0.10% with a one-standard-deviation change in either of these indices (increase in RV or decrease in UV). This effect is considerable (approximately 25%), given that the average

---

[2]Examining the fashion industry, Wenting (2008) concludes that knowledge spill-overs are limited to those spin-offs that move between fashion agglomerations.

spin-off rate was 0.37% during the period considered in the present study. We find that the average firm size in the region has a similar economic effect, except in knowledge-intensive sectors for which no statistical significance was attained. Surprisingly, the corresponding effects of localisation economies (defined as industrial specialisation) and market size are too small to be detected (they vanish with alternative specifications and estimation techniques).

An additional contribution made by the econometric part of this study is that we explicitly address the assumption of *Independence of Irrelevant Alternatives* (IIA). Previous studies have used logit models to explain occupational choice, assuming that the *IIA* assumption holds. However, this assumption can be questioned because omitted factors (e.g., psychological factors) are likely to be strongly correlated across alternatives (i.e., stay in the same job, switch to a new job, or retire).[3] We assess the *IIA* assumption with multiple tests, which lean towards rejecting the logit model. We show that alternative estimations with Nested Logit models yield results that differ both quantitatively and qualitatively: the marginal effects for specific regional explanatory variables are approximately 50% larger than the corresponding effects from the *Multinomial Logit* model, and the causal effects of some specific determinants proposed in the literature disappear.

To the best of our knowledge, this is the first multi-level study that focuses on the regional determinants of spin-offs while controlling for a series of other important factors at the individual, firm, and industry levels. Furthermore, we add to the literature by using a straightforward measure of entrepreneurship. For instance, Hyytinen and Maliranta (2008) define entrepreneurs indirectly on the basis of their participation in an insurance scheme for self-employed individuals. In the current study, firm ownership and entrepreneurial ventures – as opposed to employment – are observed directly.[4]

The remainder of this paper is organised as follows. Section 2 outlines the theoretical framework, whereas the data, definitions and econometric method are discussed in section 3. Section 4 introduces the explanatory variables, and section 5 presents the results. Section 6 summarises and concludes.

## 2. Entry through spin-offs – A conceptual framework

Standard economic geography arguments for the location and entry of new ventures typically refer to the size of the market (agglomeration economies, economies of scale, specialisation,

---

[3]For instance, these factors might include individual idiosyncrasies related to the perception of risk. See Figueiredo et al. (2002b) for a discussion of Carlton's (1985) empirical approach.
[4]Only Danish data allow for a similar analysis; see for instance, Dahl and Sorenson (2009).

and pecuniary linkages), trade costs (transport, tariffs, and non-tariff barriers), and non-pecuniary linkages (density, knowledge spill-overs, and externalities).[5] These factors are also relevant for the decision to spin off a new firm; however, there are other theoretical considerations driving the underlying mechanism that impels an individual to leave her employment and engage in a risky and uncertain entrepreneurial endeavour. The occupational choice model analyses the choice between employment and entrepreneurship in terms of profit-maximising agents making rational evaluations of future pay-offs associated with the respective occupation (Kanbur 1979; Wit, 1993). Despite the intuitive logic of this model, a number of other factors influence the choice of entrepreneurship, such as psychological traits and risk attitudes.[6] Furthermore, the geographical dimension is neglected in the occupational choice model, which is remedied – to some extent – in the random profit models that include most of these dimensions but disregard individual-level factors. A third set of models are based on information asymmetries between employees and employers and how new ideas and innovations may disrupt existing production facilities.[7]

According to Klepper and Thompson (2007), most attempts to model spin-offs fail to comply with empirical observations. Based on that insight, they proposed a theoretical model that includes heterogeneous individuals, decision making under uncertainty and divergent views over the optimal firm strategy.[8] Their *disagreement model* is explicitly designed to capture how employees' different attitudes regarding a firm's future strategy may lead to spin-offs. Nevertheless, even in this case, the determinants behind spin-offs are limited to individual abilities and generalised costs.

Disagreement alone is obviously not sufficient for a spin-off to occur. An employee must also make a subjective evaluation of entry costs, her entrepreneurial capacity, future earning potentials, and the overall prevailing conditions to start a new venture. The regional component is neglected, in particular in the Klepper-Thompson model. A more elaborate model should explicitly incorporate the following categories of variables: i) a set of regional conditions ($X$), including market size, industrial specialisation, and knowledge endowments; ii) past and current performance of the incumbent firm ($Y$); and, finally, iii) individual characteristics ($H$), such as education level and tenure (see section 4).

---

[5] See, e.g., Fujita et al. (1999) and Braunerhjelm et al. (2000).

[6] See Parker (2009) for a survey.

[7] The random profit maximisation model by Carlton (1985) is assumed to weigh all regional characteristics that may influence profits and firms' locational decisions. In addition, theoretical explanations of market entry through spin-offs relate to contract theory, asymmetric information and adverse selection (Anton and Yao 1995; Klepper and Sleeper 2005; Chatterjee and Rossi-Hansberg, 2007). Franko and Filson (2006) examine the effects of employee learning and imitation on spin-off activities.

[8] See Klepper and Thompson (2007; 2010).

Spin-offs (*s*) might, therefore, be expressed as a function of these three vectors of explanatory variables:

$$s = f(H, Y, X). \qquad (1)$$

Thus, our empirical analysis uses the Klepper and Thompson model as a point of departure, from which the explanatory variables have been classified in different categories and levels of aggregation (see Appendix A for a modified version of the model).

## 3. Data and empirical design

*Data and definitions*

We use an unbalanced panel database compiled by Statistics Sweden, which is based on census data, tax declarations, firm registries, and other complementary sources. This data set is unique insofar that it covers all individuals and firms in the Swedish private sector. It contains information on all firms and establishments (with organisational and financial data) and is linked to individuals' education, occupation, place of origin, and residence.

By isolating the information on a single individual at years *t* and *t+1*, we can identify whether she has stayed in the same job, taken up a position in a different firm, founded her own firm, or stopped working. Thus, we can trace all the incumbents who spawned spin-offs in the economy. Similarly, we can compare the founders of spin-offs with individuals who chose to stay with their current employer or switch to another employer. Because the location of all firms is known, we can include regional variables at different levels of aggregation.

For the level of industrial aggregation, we use the 2-digit SNI codes.[9] Regarding the regional dimension, we consider 81 functional labour market regions in Sweden. These represent groupings of municipalities that are characterised by a high degree of self-contained commuting flows. Therefore, regional data on industrial structure and market size are combined with data at the firm and individual levels. The analysis is carried out for the 1999-2005 period. We primarily focus on individuals who at time *t* were employed in a firm that remained extant at time *t+1* and the change (or lack thereof) in their occupation between the two time periods.

Individuals who were not employed at a certain time period are not included in the database, so individuals exiting from the panel from one time period to the next are identified as

---

[9]The SNI codes (i.e, Swedish Standard Industrial Classification used by Statistics Sweden), correspond almost perfectly to the NACE codes (i.e., the Statistical Classification of Economic Activities in the European Community). The private economy (on which this paper focuses) is broken down into 43 2-digit classes. In constructing the entropy measures (see section 4), we utilise a finer level by considering the 5-digit classes that fall under the 2-digit classes.

choosing to *Exit*. We do not consider an individual to have left her previous employment unless the identification codes pointing to her employing firm and establishment have both changed. As long as no clear change can be identified, individuals are classified as choosing to *Stay*.[10] *Switching* refers to the transition from being employed in a certain establishment of a certain firm to being employed in an entirely different firm.

*Spin-off* is the key concept of this paper that, along with entrepreneurship, lacks a stringent and generally accepted definition.[11] Klepper defines a spin-off as a minimum number of employees breaking off from an incumbent to found a new firm in the same industry. Here, we impose no such restrictions. We follow Hyytinen and Maliranta's definition, implying that firms founded by a former employee are counted as spin-offs, as long as additional information (ensuring that the entrepreneurial entry is not the product of a split, downsizing or shut down of the mother firm, can be retrieved from our data sources. The interested reader is referred to Appendix B for a detailed description of how we identify a spin-off given our data sources. The same appendix provides information on the robustness of the key empirical findings (see below) across alternative definitions of spin-offs.

Access to individual- and firm- level data allows us to observe the transition from employment to entrepreneurship directly without being dependent on proxies or secondary sources. Using this information, we define *Spin-offs* as firms started at time *t+1* by individuals who were employees in a firm at time *t*. We require that the latter firm remain extant at time *t+1*. Cases in which employees became owners of firms due to splits[12] are excluded from the category of *Spin-offs*. Therefore, in contrast to previous studies, start-ups are here linked to the founders (who can be compared to those who chose not to spin-off).

Table 1 portrays the employment choices of all working individuals during each time period. We assume that at the end of each year, the individual decides her course of action for the approaching year. The choice of employment is then identified by comparing her employment status between year *t* and year *t+1*. Therefore, the column labelled "1999" in Table 1 describes the choice of employment for the year 2000 for the 1,720,364 people who were employed in a firm in 1999; 79.3% of these individuals remained in the same job position,

---

[10]Individuals who stay in the same working position are surprisingly difficult to identify, although it is the most common occupational choice and should be relatively simple to recognise. Splits, mergers, buy-outs, and internal firm reorganisations introduce considerable turbulence in the identification codes for firms and establishments. Cases in which establishments where an individual was working were acquired by a different firm are identified as *Stay* since not much has changed from the viewpoint of the individual.

[11]See, for instance, Audretsch (2002) or Andersson and Klepper (2013) for commonly used definitions.

[12]Sweden Statistic's definition of a firm split is based on the number of employees who leave to work in the same firm in the following year. Different thresholds are used, but if more than 50 percent of the employees in a newly registered firm all originated from the same firm, then it is defined as a split and not a genuine spin-off.

14.1% switched employers, 6.4% stopped working, and 0.3% (i.e., slightly more than 5000) founded their own venture.

As illustrated in Table 1, no particular trend is discernible. The incidence of spin-offs shows a clear increase over the previous two years, although it remains a relatively rare occurrence, with an average percentage of 0.4%. The percentage of people choosing to *Switch* to different forms of employment declines rather steadily up to 2003 but rises again slightly in 2004. The percentage of people deciding to *Stay* is higher after 2001, although the respective percentage of people deciding to *Exit* seems to fluctuate approximately 7%.

Table 2 summarises the choices of the pooled sample according to industry category and originating firm size. Although the actual values for the percentages vary across industry categories, certain tendencies can be identified. The percentage of employees choosing to *Spin-off* consistently declines in all sectors as the size of the firm increases, as does the *Exit* alternative, although the decline is not as pronounced in specific sectors (*financial and real-estate services* and *knowledge-intensive business services*). The propensity to *Stay* in the same firm increases with the size of the firm in all industry categories except for *knowledge-intensive business service* (KIBS) firms.[13] The *Switch* alternative seems to show the greatest variance among the different industry sectors; it drops in some industry categories as the size of the firm increases but shows the exact opposite trend in other industries. It is notable that the trends for large firms (of more than 501 employees) deviate from most of the trends of smaller firms. Entrepreneurial spin-offs are much more likely to originate from small rather than large firms in all four industry sectors considered. This result corroborates the findings of Glaeser and Kerr (2009) and Hyytinen and Maliranta (2008).

Finally, the spatial origin-destination composition of spin-offs displays a weak variation across the four aggregate sectors of the private economy. In other words, the vast majority (slightly below 85%) of spin-offs occur intra-regionally: 81.8% in manufacturing, 83.1% in low-end services, 85.7% in financial and real-estate services and 82.7% in KIBS.

*Econometric models*

Following Hyytinen and Maliranta (2008), at the end of each time period *t*, each employee (labelled *n*) faces the following alternatives in time period *t+1*: she can decide to *Stay* in her current employment, *Switch* to a different employer, *Exit* the workforce, or *Spin-off* and start her own firm. These four mutually exclusive alternatives comprise the choice set $C =$

---

[13]KIBS refers to NACE 72-74.

$\{stay, switch, exit, spin\ off\}$. Assuming that, by choosing alternative $i$, individual $n$ obtains utility $U_{ni}$, the choice probability of this alternative is: $P_n(i) = P(U_{ni} \geq U_{nj}, \forall j \in C, j \neq i)$. We assume that this utility can be decomposed into an observed and an unobserved part:

$$U_{ni} = V_{ni}(a_i, \boldsymbol{x}_n, \boldsymbol{\beta}_i) + \varepsilon_{ni}, \tag{2}$$

where the systematic utility, $V_{ni}$, is a function of the alternative-specific constant $\alpha_i$, a vector of *case (individual) specific variables* $(\boldsymbol{x}_n)$, and a vector of alternative-specific parameters $(\boldsymbol{\beta}_i)$ that are used to generate utility differences across alternatives. The unobserved portion of utility $(\varepsilon_{ni})$ incorporates omitted factors and measurement errors. Under a *linear-in-parameter* specification, the choice probability of $i$ becomes: $P_n(i) = P(\varepsilon_{nj} - \varepsilon_{ni} \leq a_i + \boldsymbol{x}_n{}' \boldsymbol{\beta}_i - a_j - \boldsymbol{x}_n{}' \boldsymbol{\beta}_j, \forall j \in C, j \neq i)$.

The most convenient way to compute this probability is to assume that the error terms are distributed *i.i.d. extreme value type I*, which gives rise to logistically distributed error differences, $\varepsilon_{nj} - \varepsilon_{ni}$, and a *multinomial logit* (MNL) choice probability:

$$P_n(i) = \frac{\exp(a_i + \boldsymbol{x}_n{}' \boldsymbol{\beta}_i)}{\sum_{j \in C} \exp(a_j + \boldsymbol{x}_n{}' \boldsymbol{\beta}_j)}. \tag{3}$$

The likelihood that the observed sample is realised is:

$$L_n(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \prod_n \prod_{i \in C} \{\frac{\exp(a_i + \boldsymbol{x}_n{}' \boldsymbol{\beta}_i)}{\sum_{j \in C} \exp(a_j + \boldsymbol{x}_n{}' \boldsymbol{\beta}_j)}\}^{y_{ni}}, \tag{4}$$

where $y_{ni} = 1$ if individual $n$ has chosen alternative $i$ and zero otherwise.[14]

The MNL model is associated with certain drawbacks (Train, 2002). Among others, the above model postulates a very strict substitution pattern between alternatives, which is summarised by the property of *Independence from Irrelevant Alternatives (IIA)*. IIA dictates that the choice probability ratio between any pair of arbitrary alternatives $(i, j) \in C$:

---

[14]Because the model contains only *alternative specific constants* $(\boldsymbol{\alpha} = \alpha_{stay}, \alpha_{switch}, \alpha_{exit}, \alpha_{spin})$ and alternative specific parameters $(\boldsymbol{\beta} = \boldsymbol{\beta}_{stay}, \boldsymbol{\beta}_{switch}, \boldsymbol{\beta}_{exit}, \boldsymbol{\beta}_{spin})$, it suffices to set $\alpha_{stay} = 0$, $\boldsymbol{\beta}_{stay} = \boldsymbol{0}$, and to normalise the scale of utility (i.e., set $Var(\varepsilon_{ni}) = \frac{\pi^2}{6}$) in order for the MNL model to be identified.

$$\frac{P_n(i)}{P_n(j)} = \frac{\exp\left(x_n'\boldsymbol{\beta}_i\right)}{\exp\left(x_n'\boldsymbol{\beta}_j\right)} \tag{5}$$

is independent of any alternatives other than $i$ and $j$; consequently, when an alternative $k$ is added in $C$, the decrease in choice probabilities of $i$ and $j$ is such that the ratio $P_n(i)/P_n(j)$ remains intact. This substitution pattern might appear to be unrealistic in the context of this study. For instance, one would expect the probability ratio between *Stay* and *Spin* might be disturbed if the option *Switch* is present in the worker's choice set. This issue is non-trivial and constitutes a recurring problem in discrete choice estimations; accordingly, we devote considerable space to addressing it.

*Testing the IIA assumption*

We conduct a series of IIA tests to examine how realistic the IIA assumption is in the context of this study. The *Hausman-McFadden* (1984) test is based on repeating the MNL estimation for each choice subset $C_R = \{\alpha_1, \alpha_2, \alpha_3\}$, with alternatives $\alpha_1 \neq \alpha_2 \neq \alpha_3$ chosen from $C$. The restricted MNL estimation in each of these subsets yields a vector of ML estimators, $\breve{\boldsymbol{\beta}}_R = (\boldsymbol{\alpha}_R, \boldsymbol{\beta}_R)$, and a covariance matrix $\boldsymbol{\Sigma}_R$. The test statistic (MFH) is the quadratic form,

$$MFH = [\breve{\boldsymbol{\beta}}_U - \breve{\boldsymbol{\beta}}_R]'[\boldsymbol{\Sigma}_U - \boldsymbol{\Sigma}_R]^{-1}[\breve{\boldsymbol{\beta}}_U - \breve{\boldsymbol{\beta}}_R], \tag{6}$$

where $\breve{\boldsymbol{\beta}}_U = (\boldsymbol{\alpha}, \boldsymbol{\beta}_U)$, $\boldsymbol{\beta}_U$ refers to the ML estimator sub-vector that can be obtained by dropping from the full model the estimators that correspond to the restricted model's excluded parameters, and $\boldsymbol{\Sigma}_U$ is derived from the covariance matrix of the full model, $\boldsymbol{\Sigma}$, by dropping the rows and columns that correspond to excluded parameters in the restricted model. It can be shown that, asymptotically, the MFH test statistic follows the $\chi^2$ distribution with $r(\boldsymbol{\Sigma}) - r(\boldsymbol{\Sigma}_R)$ degrees of freedom, where $r$ stands for the rank of the covariance matrix.

There is no guarantee that the finite sample distribution of $MFH$ is $\chi^2$. In fact, $MFH$ can yield negative values in finite samples. Although negative values support the IIA assumption, the test results must be accompanied by an additional test, the *suest-based Hausman* (SBH), which is a modified version of the Hausman-McFadden test and allows for safer conclusions (Long and Freese, 2006).

We also perform the IIA test proposed by Small and Hsiao (1985). The test randomly divides the data set into two subsamples of roughly equal size; applying maximum likelihood estimation under the MNL assumptions in the two full models yields two estimator vectors of

equal size: $\widehat{\boldsymbol{\beta}}_{1F} = (\boldsymbol{\alpha}_{1F}, \boldsymbol{\beta}_{1F})$ and $\widehat{\boldsymbol{\beta}}_{2F} = (\boldsymbol{\alpha}_{2F}, \boldsymbol{\beta}_{2F})$. The estimation can be repeated with the second sample for each choice subset $C_R = \{\alpha_1, \alpha_2, \alpha_3\}$, with alternatives $\alpha_1 \neq \alpha_2 \neq \alpha_3$ chosen from $C$. For each restricted choice set, we obtain $\widehat{\boldsymbol{\beta}}_{2R} = (\boldsymbol{\alpha}_{2R}, \boldsymbol{\beta}_{2R})$ and construct the Small-Hsiao (SH) test statistic:

$$SH = -2\left[L_R\left(\widehat{\boldsymbol{\beta}}_{12U}\right) - L_R\left(\widehat{\boldsymbol{\beta}}_{2R}\right)\right]. \tag{7}$$

The parameter vector $\widehat{\boldsymbol{\beta}}_{12U}$ is obtained from $\widehat{\boldsymbol{\beta}}_{12F} = \frac{1}{\sqrt{2}}\widehat{\boldsymbol{\beta}}_{1F} + [1 - \frac{1}{\sqrt{2}}]\widehat{\boldsymbol{\beta}}_{2F}$ by dropping the estimators that correspond to the restricted model's excluded estimators, and $L_R$ is the likelihood function of the restricted model.

Such formal tests based on restricted choice sets have been criticised (Cheng and Long, 2007) because they frequently yield contradictory results, which is also the case for our specification (the IIA tests provide no universal acceptance or rejection; see section 5). Our Nested Logit (NL) models allow for some correlation of the unobserved factors in the choice subset $C_A = \{stay, switch, exit\}$.[15] The rationale behind using an NL model is that the error terms of these alternatives might contain a common characteristic, e.g., *the lack of entrepreneurial risk*. Conversely, the alternative in the degenerate nest, i.e., spin-off, involves financial and perhaps legal risk that is not present in the other three alternatives. The NL model assumes a more flexible Generalised Extreme Value (GEV) distribution for the error terms, which gives rise to the choice probability:

$$P_n(i \in C_A) \tag{8}$$
$$= \frac{\exp(V_{ni}/\lambda_A)\left[\exp(V_{n,stay}/\lambda_A) + \exp(V_{n,switch}/\lambda_A) + \exp(V_{n,exit}/\lambda_A)\right]^{\lambda_A - 1}}{\left[\exp\left(\frac{V_{n,stay}}{\lambda_A}\right) + \exp\left(\frac{V_{n,switch}}{\lambda_A}\right) + \exp\left(\frac{V_{n,exit}}{\lambda_A}\right)\right]^{\lambda_A} + \exp(V_{n,spin\,off})},$$

for any $i \in C_A$ and:

$$P_n(spin\,off) \tag{9}$$
$$= \frac{\exp(V_{n,spin\,off})}{\left[\exp\left(\frac{V_{n,stay}}{\lambda_A}\right) + \exp\left(\frac{V_{n,switch}}{\lambda_A}\right) + \exp\left(\frac{V_{n,exit}}{\lambda_A}\right)\right]^{\lambda_A} + \exp(V_{n,spin\,off})},$$

---

[15]All NL estimations that are not discussed in the robustness section are available upon request.

for the spin-off alternative.[16] Train (2002) suggests that $(1 - \lambda_A)$ is a measure of the strength of the correlation of errors in the $C_A$ nest. In other words, if $\lambda_A = 1$, the correlation becomes zero, and the above NL choice probabilities collapse to a logit form. Because MNL is a special case of NL, testing the MNL model[17] is encompassed in the estimation of an NL model.[18]

## 4. Explanatory variables and hypotheses

*Regional level variables: knowledge and culture*

Both non-pecuniary and pecuniary effects are associated with regions. Regarding the pecuniary effects, more sizeable markets imply a larger pool of potential customers, exploitation of economies of scale, and a distribution of sunk costs over larger volumes of production. In addition, the probability of identifying unexplored market niches that target a particular segment of the market increases.[19]

Regarding non-pecuniary effects, an impressive number of studies support the hypothesis that knowledge spill-overs are geographically bounded (stressing the importance of spatial proximity) and that knowledge is embedded in regionally immobile agents (Jaffe 1989; Audretsch and Feldman 1996; Keller 2002). Knowledge is thus partly tied to certain geographic entities, and the size of the market has a direct effect on both the opportunity space and the feasibility of exploiting opportunities. The measure that we use to control for the size of the region is the logarithm of the number of active workers in the region.

Simultaneously, the size of the region encapsulates local knowledge endowments. Two different measures of regional knowledge resources are considered: first, the region's

---

[16]Note that it is possible to rewrite the choice probabilities of nest *A* as products of a marginal probability of choosing nest *A*, i.e., the probability of choosing a non-risky behaviour times the conditional probability of choosing a specific alternative, given this non-risky behaviour.

[17]The remainder of the limitations of MNL models discussed in Train (2002) include the inability to i) capture random taste variation and ii) to handle dynamics in the unobserved factors. The first limitation is highly relevant in choice situations that involve *attributes* of the alternatives (alternative-specific variables), which are not available in this study; although random taste variation can be addressed with a more advanced (e.g., mixed logit) model, its presence is assumed away. Access to panel data allows for the use of dynamic modelling approaches that explicitly incorporate the time dimension. However, Rhody (1998) and Dunn and Holtz-Eakin (2000) indicate that the gains from such methods are limited. Thus, we pool the data from all years and follow the two traditional approaches discussed above: i) a straightforward MNL estimation with the inclusion of time dummies and ii) an NL estimation consistent with Random Utility Maximisation to account for the effects of entrepreneurial risk as an unobserved factor.

[18]The dissimilarity parameter $(1 - \lambda_A)$ is strongly significant (p values below 0.01) in all NL estimations reported throughout the paper: parameter $\lambda_A$ ranges between 0.277 and 0.419 across the NL estimations.

[19]Urbanisation/agglomeration is one of two processes found to have a consistently positive effects on regional start-up rates; the presence of small firms and economic specialisation represents the other (Henderson and Thisse, 2004). Moreover, larger markets tend to protect new ventures from the hold-up costs of customisations (McLaren, 2000).

*knowledge intensity*, which is measured as the ratio of employees with a tertiary education over the total number of employees in the region, and, second, *R&D investments* by private firms and universities. Both measures of knowledge resources are found to be highly correlated with the size of the region (ρ=0.87 and ρ=0.94, respectively), and their addition failed to improve the fit of the empirical model.[20] As a result, *knowledge intensity* hereinafter refers to firms (see below), not regions.

A more specific non-pecuniary regional effect is related to the importance of a *local entrepreneurial culture*. Glaeser and Kerr (2009) conclude that environments dominated by smaller and independent firms are more conducive to entrepreneurship than environments hosting large monopolists. Rosenthal and Strange (2010) find a strong correlation between local average establishment size and subsequent employment growth through start-ups. Building on these findings, we use an average firm size index (ASIZE) to illustrate attitudes towards small businesses,

$$ASIZE_{ij} = \frac{S_{ij}/E_{ij}}{\sum_j S_{ij}/\sum_j E_{ij}} \tag{10}$$

where *S* refers to the numbers of firms, *E* is the number of employees, and the subscripts refer to industry (*i*) and region (*j*). The ASIZE index, i.e., the inverse of the average firm size, might also be argued to control for scale effects, i.e., technology, in the corresponding industry. High values of the ASIZE index indicate a local industry populated by many small firms, whereas low values indicate fewer but larger firms that enjoy more market power.

*Regional industry structure: specialisation versus diversity*

Whether the specialisation or diversity of industries in a region promotes entry is the subject of on-going debate.[21] As indicated early on by Marshall (1890), and as further emphasised by Arrow (1962) and Romer (1986), knowledge spill-overs are suggested to be industry specific and may only be appropriated in regions with high industry-specific concentrations (*MAR externalities*). Jacobs (1969) proposes the opposite view, advocating that knowledge exchange

---

[20]Note that R&D data are only available for 2001, and temporal extrapolations are seldom trustworthy. Considering relative knowledge endowments (R&D per inhabitant) mitigates but does not extinguish the problem because the correlation remains high (0.69).
[21]See Paci and Usai 1999; Duranton and Puga 2000; Klepper 2002; Rosenthal and Strange 2003; van Ort and Atzema 2004; van der Panne and van Beers, 2006. Efforts to disentangle the effect of the different externalities have focused on either overall employment growth or on innovation output and have overlooked the effect of spatial externalities on entrepreneurial activity. See Henderson and Thisse (2004) for a survey.

among a diverse spectrum of industries facilitates experimentation, innovation, and market entry. Although the empirical evidence is mixed, slightly more weight is given to specialised regions in explaining entrepreneurship, and the opposite seems to prevail for innovative activities.[22]

We adopt a location quotient as our measure of specialisation. The production specialisation index (PS) captures the relative specialisation of each industry at the 2-digit NACE level in the corresponding region,

$$PS_{ij} = \frac{E_{ij}/\sum_j E_{ij}}{\sum_j E_{ij}/\sum_i \sum_j E_{ij}} \qquad (11)$$

where $E$ denotes employment and subscripts $i$ and $j$ denote industries and regions (functional labour markets), respectively. A common normalisation is to calculate the ratio $\frac{PS-1}{PS+1}$, which is balanced and restricted between -1 and 1, unlike (11).[23] Values of this corrected PS index larger than zero indicate a higher degree of industrial specialisation compared with national industrial composition, whereas values smaller than zero indicate the opposite.

Regarding diversity, it is important to distinguish between related and unrelated variety when considering the significance of Jacobian externalities.[24] Following Frenken et al. (2007), we apply an entropy measure to capture related versus unrelated variety. The latter (UV) is expressed as the entropy of the two-digit industry-level distribution of regional employment. Intuitively, UV captures the degree to which local employment is evenly distributed across sectors. *Related variety* (RV) is the weighted sum of the entropy at the five-digit level within each two-digit class. It shows the degree to which a sector's (two-digit) local employment is evenly distributed across its (five-digit) sub-sectors in a region. Indexing the two-digit SNI present in a region with $i$, *Unrelated Variety* (UV) (suppressing the regional index $j$) can be expressed as:

$$UV = \sum_{i=1}^{43} P_i \ln\left(\frac{1}{P_i}\right), \qquad (12)$$

---

[22]See Desrochers and Sautet (2008) and Beaudry and Schiffauerova (2009).
[23]We apply the same standardisation for the ASIZE index.
[24]Technological breakthroughs in the motorbike industry are more likely to find their way into the automobile industry than, for instance, into the pulp or pharmaceutical industries. Some industries are simply more complementary than others, and Jacobian externalities are more likely to appear among such industries.

where $P_i$ is the two-digit local employment share. Similarly, indexing all five-digit SNI sectors that fall exclusively under the two-digit industry $i$ with $g$ (note that $g$ may vary across $i$), *Related Variety* (RV) is given by:

$$RV = \sum_{i=1}^{43} P_i H_i,$$ (13)

and

$$H_i = \sum_{g \in i} \frac{P_g}{P_i} \ln \left( \frac{1}{P_g \big/ P_i} \right),$$ (14)

where $P_g$ refers to the five-digit local employment shares.[25]

In summary, the specialisation index captures the employment share of the industry (that a prospective entrepreneur works in) but provides no information regarding how the various industries in a region relate to one another. Thus, we need a second measure that addresses industrial variety, but when analysing the variety of other industries, industries that are related and industries that are completely unrelated may be distinguished from one another.

*Control variables: firm and industry level*

We construct dummy variables to classify firms into six size categories. The underlying assumption is that large firms that invest more in R&D, should also spin-off more firms because new knowledge creates opportunities that are not always suitable for integration into firms' on-going operations.[26] In addition, we use the ratio of employees with a tertiary education over total firm employment as a proxy for firms' innovativeness. Unfortunately, detailed data on innovative output are extremely difficult to obtain at the disaggregated level pursued in this analysis.

Moreover, controls for the age of the firm[27], the logarithm of productivity (defined as value added per employee), and four different industry dummies (according to the groups in Table

---

[25]Intuitively, *UV* takes a value of zero when all local employment falls under the same 2-digit sector. Its value increases as employment spreads to more 2-digit sectors. *RV* takes a value of zero if for all 2-digit sectors are present in a region and employment is concentrated under only a single 5-digit sector within each 2-digit sector. As employment within each 2-digit sector spreads in more 5-digit sectors, the *RV* index increases in value.

[26] Previous studies are ambiguous regarding how firm size affects spin-offs. See, for instance, Gompers et al. (2005) and Kim and Marschke (2005).

[27]The age of the firm, as with individuals' tenure, is left truncated because the point of reference is the birth of the database in 1985, i.e., when observations began to be collected rather than when firms were created. The longest period attainable is therefore 18 years, but adding a dummy variable to distinguish truncated firms does not improve the fit of the model.

2) are also used in the analysis.[28] We also include a *declining employment* dummy, which equals one if there has been a drop in the firm's employment between time *t* and *t-2*, and a *declining sales* dummy, which equals one if there has been a drop in the sales of the firm between *t* and *t-2*. These two dummies are controls for push-out effects caused by a decline in the business of the incumbent firm. In particular, the choice to *Exit* is often involuntary, and these dummies allow us to control for whether the firm is downsizing.

*Control variables: individual level*

Finally, we control for the individual determinants that are most commonly referred to in the literature on entrepreneurship: tenure, age, age squared, gender (which equals one for males and zero otherwise), wage, and education (Berglann et al., 2009). The individual's education level is captured by a set of seven dummy variables ranging from primary education to PhD, as categorised by Statistics Sweden.[29]

The independent variables included in the econometric analysis are summarised in Table 3. Approximately 15% of the population (as described in Tables 1 and 2) were excluded because of missing data in certain variables.[30] Note that the summary statistics were calculated after pooling the data for the 1999-2004 period. Approximately 70% of the individuals included in the study are males; approximately one-third of the individuals are employed in a firm exhibiting declining sales/employment; approximately one-third of the individuals work in firms of fewer than 50 employees, another one-third work in firms employing between 51 and 500 individuals, and the remaining third work in firms of 501 or more employees; the majority of the individuals (83.5%) work either in manufacturing or in low-end services.

## 5. Estimation results

We have performed both MNL and NL estimations (the base outcome was considered *Stay*) to evaluate the effects of the above explanatory variables. The parameter estimates can be used to compute the effects of marginal changes in the explanatory variables on the choice

---

[28]The 43 industry branches are grouped into these four categories in order to limit the number of controls.
[29]These categories are as follows: 1. primary and lower secondary education, shorter than 9 years; 2. primary and lower secondary education, 9 (10) years; 3. upper secondary education, 2 years or shorter; 4. upper secondary education, longer than 2 years but at most 3 years; 5. post-secondary education, shorter than 3 years; 6. post-secondary education, 3 years or longer (excluding PhD); 7. PhD.
[30]Missing values were mainly due to the imperfect reporting of the financial variables of some firms. After comparing kernel densities, no pattern emerged that distinguishes these firms in any way, leading us to the conclusion that there is no selection bias.

probabilities, i.e., the marginal effect on the propensity to choose each of the four alternatives considered.

The results are reported in Tables 4-12. We first report the results for the entire Swedish private sector (Table 4) and then for knowledge-intensive manufacturing and knowledge-intensive services only (Table 6). Thereafter, we undertake a number of robustness tests and estimations with alternative econometric specifications (Tables 5 and 7-12).

Tables 4 and 6 report the marginal effects of the MNL estimations and their standard errors. Overall, the different regional variables are shown to have a substantial influence on spin-offs. These primary results suggest that regional size has a positive and significant impact on the propensity of individuals to *Spin-off* (although the effect is much weaker for knowledge-intensive sectors; see Table 6). The results corroborate findings from previous studies but fail to prove robust under more elaborate specifications and after an explicit treatment of the IIA assumption (see below).

Regarding the effects of the regional industrial composition, it appears that *MAR externalities* have a positive effect on the propensity of the individual to *Spin-off,* but this effect is confined to high-tech manufacturing and KIBS (Table 6). However, this effect vanishes in the sensitivity analysis, i.e., when the MNL model is replaced by an NL model (Table 12).

As for the effect of *Jacobian externalities*, the findings are significant across all industry sectors and survive the sensitivity analysis. Unrelated variety, which may capture the negative aspects of dense agglomerations, such as congestion and high rents, has a negative effect on the propensity to *Spin-off*. Related variety, which captures knowledge externalities between diverse but related industries, is shown to have a positive and significant effect on the propensity to *Spin-off*.[31] Note also that the regional prevalence of small firms (ASIZE) positively affects the probability of spin-offs in the analysis for all industries, but the result is not robust for the knowledge-intensive sectors (Table 12).

With respect to the firm-level control variables, the results deviate from those of Hyytinen and Maliranta (2008), particularly when differences between spin-offs in general and in high-tech sectors are taken into account. As in the Finnish study, smaller firms are found to be more likely to spawn new firms. In addition, the Finnish analysis revealed that less productive and less innovative firms seem to be consistently more likely to spawn entrepreneurs. However, the latter result is not entirely the case in our study. When considering the entire private sector

---

[31]These results are partly consistent with those of Baptista and Swann (1999) and Swann and Prevezer (1996), who find that specialisation has a positive effect on particular high-tech industries. See also van Oort and Atzema (2004).

(Table 4), higher productivity has a negative effect on the choice to *Spin-off*, but that is not the case for KIBS (Table 6). This negative effect may arise from increasing alternative costs (risks) associated with spin-offs in industries not characterised by the same opportunities as those in more knowledge-intensive industries.

A firm's innovativeness, approximated by its knowledge intensity, has a positive effect on the propensity to *Spin-off* across all industry sectors. These results corroborate those of Hyytinen and Maliranta (2008) and partly corroborate those of Gompers et al. (2005), who suggest that more productive high-tech firms, i.e., those exhibiting a higher degree of knowledge intensity, are more likely to spawn entrepreneurs.[32] More complex and knowledge-intensive production can be expected to foster more diversified ideas about the best strategy regarding how best to operate such firms. In general, the estimates of our firm[33] and individual[34] control variables are consistent with previous findings and are robust across all specifications. Therefore, these variables are excluded from the sensitivity analyses, which focus exclusively on the effects of regional determinants.

*Sensitivity analyses*

Table 5A presents the marginal effects on the decision to *Spin-off*, where the MNL models have been separately estimated on regional-, individual-, and firm-level controls (models I, II, and III, respectively), and compares them with the marginal effects from the complete model (model IV in Table 5A is a copy of the last column of Table 4).[35]

---

[32]In Hyytinen and Maliranta (2008), an R&D dummy was used. However, as noted by Machin and van Reenen (1998), an industry's R&D intensity may determine the demand for skilled labour, which raises the question of whether it is possible to disentangle the two effects when simultaneously controlling for the education level (skill) of the individual employee and the innovativeness of the firm.

[33]The age of the incumbent firm seems to have a weak but negative effect on the propensity to *Spin-off* but only when we consider the entire private sector. When we focus on high-tech manufacturing and KIBS firms, no such age effect is detected. Declining employment and sales have strong and negative effects on *Staying* and positive and significant effects on all the other alternatives (exit, spin-offs, and switch).

[34]More educated males with shorter tenure exhibit the highest propensity to become entrepreneurs, while the current wage of the individual, representing the opportunity cost of any change in employment status, has a positive effect on *Staying* and a negative effect on all three other choices. Although the individual's current wage and the incumbent's productivity may both be considered measures of one's opportunity cost, it is notable that they are not highly correlated (0.07). Age has a non-linear impact on all four choices. Furthermore, men are much more likely than women to *Switch*, *Exit*, or *Spin-off*. It is also notable that the likelihood of *Staying* in the same position drops as the level of education increases. By contrast, the probabilities of *Switching* and *Spinning-off* increase with the level of education.

[35]Furthermore, the following alternative specifications were also estimated: i) using only two instead of seven education dummies, splitting the population into those with several years of university education (more than 3 years) and those without; ii) using the log of the number of employees in a firm as a firm size control instead of the six size category dummies; iii) excluding the knowledge intensity of the firm as a control; and iv) adding a spatially lagged set of the regional independent variables to check for the presence of spatial autocorrelation. These modifications had minor effects on the remaining variables. All these results are available upon request. A

To further test for omitted variable bias, we estimate the (MNL) model with a step-wise removal of groups of explanatory variables. These groups correspond to the three levels of explanatory variables discussed earlier (individual, firm, and industry), although regional variables are always included. The rationale behind this test is that if estimator values for the spatial indices do not change drastically (changes in the signs or large changes relative to their initial value) when the explanatory variables from other the levels are dropped, they are less likely to do so if potentially omitted variables (from these levels) are included in the model.

The third column of Table 5B displays the marginal effects of the regional indices of interest for a specification from which all firm and industry characteristics have been removed. The corresponding estimates in the fourth column are from a model estimated with the regional indices and year dummies only. The results show that all indices, apart from the specialisation index (PS), are generally stable. Moreover, the RV and UV indices are relatively less sensitive than ASIZE and specialisation. This result is consistent with the results from the robustness checks based on alternative specifications and estimation techniques (see below). Tables 7A and 7B display the results of repeating the same steps as in Table 5 (A and B) where the analysis is restricted to the subgroup of individuals who are employed in KIBS to assess the different tendencies for this group of entrepreneurs.

Regarding our tests of the IIA assumption, the results of the formal tests described in section 3 are presented in Tables 8 to 10 (Hausman-McFadden, suest-based Hausman, and Small-Hsiao tests). The results do not point to the universal acceptance or rejection of the IIA assumption (in accordance with Cheng and Long, 2007), but they clearly lean towards rejection. As a further test, the NL specification described above is also estimated on a random sample (10%) of the population.[36]

Both the benchmark model and the NL model are estimated with standard errors clustered at the regional level because conventional estimation techniques may understate the standard errors of the regional indices. Comparisons between the benchmark MNL model and Model I and between the NL model (Model III) and Model IV in Tables 11 and 12 do not reveal any qualitative changes (the statistically significant variables remain significant) except for the log of regional size, which ceases to be significant for KIBS.

Finally, we include a dummy variable for the three largest agglomerations (Stockholm, Gothenburg, and Malmo) to control for potentially omitted variables that are generally stable

---

similar regression was also estimated for the high-tech manufacturing and KIBS sub-sectors (compare Table 6), again with a minor impact on the remaining variables.

[36]A 10% sample was chosen due to the heavy computational burden and prolonged estimation time associated with NL models.

over a short period – such as the variable considered in the study – but that differ across the three urban areas above and the rest of the country. Such variables may include, but are not limited to, unemployment rates and the value of public infrastructure in an area. The results from Models II and V (again, presented in Tables 11 and 12) show that the inclusion of these area-specific constants absorbs the effect of regional size for all industries, and the associated variable becomes insignificant when we use either the MNL or the NL specification (Table 11).

It is notable that the choice of estimation technique (MNL versus NL) results in both qualitative and quantitative changes, which is evident from the model pair comparisons (benchmark versus model III, model I versus model IV, and model II versus model V) in Tables 11 and 12. *Qualitatively*, the log of regional size in the case of KIBS and the average size of the firm cease to be significant with the NL models. Thus, the only strongly significant factors affecting spin-offs are UV and RV. *Quantitatively*, the marginal effects of these two indices are up to 50% larger in the NL models compared with their MNL counterparts.

The magnitude of these effects is conditional on the values of the explanatory variables on which the effect is measured. Figure 1 presents the predicted spin-off probability (MNL estimation) calculated at the mean observation vector for KIBS industries in the most recent year (2004). This probability varies considerably within the range of the indices of interest: it doubles within the range of RV and quadruples within the range of UV.

To some extent, the central results of the paper depend on the manner in which a spin-off is defined. Appendix B provides a series of sensitivity analyses based on narrower definitions, which filter out *inter-industry*, *individual*, and *non-corporate* spin-offs. The first category includes cases in which entry has taken place in an industry that is different (in the 2-digit level) than the industry of the parent firm. However, even in this case our data do not allow us to check whether the founder of the spin-off performed a task (in the mother firm) that is unrelated to the industry of the entry. The second category includes cases in which the spin-off firm did not hire any employees. The last category contemplates cases in which the new firm did not enter the stock market. The sensitivity analyses show that results based on our (relatively broad) definition are qualitatively robust when these alternative definitions are considered.[37] Apart from sensitivity analyses, Appendix B provides an analytic description of how a spin-off firm is identified in this study.

---

[37]Qualitatively similar results are also obtained from an alternative estimation that excludes inter-regional spin-offs, i.e., cases in which entry has taken place in a different functional labor market. Indicatively, the marginal

## 6. Conclusions

In this paper, we examined the underlying factors that trigger spin-offs, focussing on regional variables and implementing controls at the individual, firm, and industry levels. A unique data set cross-tabulated on individuals, firms, and regions made this analysis possible. Theoretically, we adopted the Klepper and Thompson disagreement model in which employees choose to leave a firm because of their divergent and idiosyncratic opinions about firms' future strategies, which was augmented to include variables at the levels mentioned above.

In the empirical segment of our study, we used two discrete choice techniques (MNL and NL). After implementing several robustness tests, we conclude that geography influences an individual's decision to undertake a spin-off. Jacobian externalities, defined as related variety, have a positive effect on spin-offs, whereas the opposite effect was found for unrelated variety. Moreover, the prevalence of a regional culture of small firms was found to have a positive but modest effect on spin-offs, which was significant in the analysis of all industries but insignificant for knowledge-intensive sectors. Regional size appeared to be significant in the benchmark model but failed to retain its significance when rigorous robustness tests were applied. Similar results were obtained for the specialisation index (MAR externalities), which at first appeared to be significant for knowledge-intensive sectors. These findings provide some support for the sub-market hypothesis, i.e., that spin-offs may induce clustering in related industries. Finally, both firm-level performance and individual characteristics seem to be important for the individual decision to migrate into entrepreneurship.

In addition to statistical significance, we examined the economic significance of the indices described above. Rough calculations suggested that a one-standard-deviation increase (decrease) in related (unrelated) variety increases the spin-off rate by 25%. Previous research (e.g., Andersson et al., 2012) suggests that entry tends to intensify competition and raise productivity in incumbents, indicating that entry has potentially strong welfare effects. The same should hold for spin-offs, in particular, because they have been shown to be the most successful entrants.

An elaborate welfare analysis would require the specification of precise policy interventions and a measure of their social costs, in addition to a monetised measure of the social benefit associated with a representative spin-off. Such an exercise is beyond the scope of this paper

---

effects of the regional variables of interest (using the base MNL model) are the following: Log of regional size 0.046***, PS index 0.107***, UV index -0.219***, RV index 0.080***, and ASIZE index 0.213***.

and constitutes a topic for future research. However, conceivable instruments might include *spatially varying* tax and subsidy schemes that generate incentives for firms and individuals to locate in certain regions. Such policies have been implemented at least partially in a number of countries and through international agreements (e.g., within the EU). Combined with measures to improve the educational system, such policies might facilitate knowledge diversification in production in regions with a revealed comparative advantage (related variety) and could thus encourage the growth of knowledge-intensive industries and spin-offs. However, it should be stressed that the capability of a local government to design and implement such policies is generally limited and that the above policy interventions depend on political decisions made at the national level.

## References

Agarwal, R., Raj, E., Franco, A.M. and Sarkar, M.B. (2004). Knowledge transfer through inheritance: Spinout generation, development and survival. *Academy of Management Journal*, 47: 501-522.

Andersson, M. and Klepper, S. (2013). Characteristics and Performance of New Firms and Spinoffs in Sweden, *Industrial and Corporate Change*, 22: 245-280.

Andersson, M., Braunerhjelm, P. and Thulin, P. (2012). Entrepreneurs, Creative Destruction and Production. Entry by type, sector and sequence. *Journal of Entrepreneurship and Public Policy*, 1: 125-146.

Anton, J. and Yao, D. (1995). Start-ups, spin-offs and internal projects. *Journal of Law, Economics and Organization*, 11: 362-378.

Arrow, K. (1962). The Economic Implication of Learning by Doing. *Review of Economics and Statistics*, 80: 155-173.

Audretsch, D.B. (2002). Entrepreneurship: A survey of the literature, Prepared for the European Commission, Enterprise Directorate General.

Audretsch, D.B., and Feldman, M. (1996). R&D spillovers and the geography of innovation and production. *American Economic Review*, 86: 630-640.

Baptista, R. and Swann, P. (1999). A comparison of clustering dynamics in the US and UK computer industries. *Journal of Evolutionary Economics*, 9: 373-399.

Beaudry, C. and Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy*, 38: 318-337.

Berglann, H., Moen, E., Roed, K. and Skogstrom, J. (2009). Entrepreneurship: Origins and returns. IZA DP No. 4250, Bonn.

Boschma, R.A. and Wenting, R. (2007). The spatial evolution of the British automobile industry. Does location matter? *Industrial and Corporate Change*, 16: 213-238.

Braunerhjelm, P.,Faini, R., Norman, V., Ruane, F. and Seabright, P. (2000). Integration and The Regions in Europe: How The Right Policies Can Prevent Polarization? *Monitoring European Integration* 10, London: CEPR.

Braunerhjelm, P. (2011). Entrepreneurship, innovation and economic growth: Interdependencies, irregularities and regularities. In Audretsch, D., Falck, O. and Heilbach, P. (eds.) *Handbook of Innovation and Entrepreneurship*, Northampton: Edward Elgar.

Buenstorf, G. and Klepper, S. (2010a). Submarkets dynamics and innovation: the case of the US tire industry. *Industrial and Corporate Change*, 19: 1563-1587.

Buenstorf, G. and Klepper, S. (2010b). Why does entry cluster geographically? Evidence from the US tire industry. *Journal of Urban Economics*, 68:103-114.

Carlton, D. (1985). The location and employment choices of new firms: An econometric model with discrete and continuous endogenous variables. *The Review of Economics and Statistics*, 65: 440-449.

Chatterjee, S. and Rossi-Hansberg, E. (2007). Spin-offs and the markets for ideas, NBER WP 13198, Cambridge, MA.

Cheng, S. and Long, J.S. (2007). Testing for IIA in the multinomial logit model. *Sociological Methods Research*, 35: 583-600.

Dahl, M. S., Reichstein, T. (2013) Are you experienced? Prior experience and the survival of new organizations. *Industry and Innovation*, 14: 497-511,

Dahl, M.S., Sorenson, O. (2009) The embedded entrepreneur. *European management Review*, 6:172-181.

Desrochers, P. and Sautet, F. (2008) Entrepreneurial policy: The case of regional specialization vs. spontaneous industrial diversity. *Entrepreneurship Theory and Policy*, 32: 813-832.

Dunn, T. and Holtz-Eakin, D. (2000). Financial capital, human capital, and the transition to self-employment: evidence from intergenerational links. *Journal of Labor Economics*, 18: 282-305.

Duranton, G. and Puga, D. (2000). Diversity and specialization in cities: Why, where and when does it matter? *Urban Studies*, 37: 533-555.

Elfenbein, D.W., Hamilton, B.H. and Zenger, T.R. (2010). The small firm effect and the entrepreneurial spawning of scientists and engineers. *Management Science, Articles in Advance*, 56: 659-681.

Eriksson, T. and Kuhn, J.M. (2006). Firm spin-offs in Denmark 1981-2000 – patterns of entry and exit. *International Journal of Industrial Organization*, 24: 1021-1040.

Figueiredo, O., Guimarães, P. and Woodward, D. (2002a). Home-field advantage: location decisions of Portuguese entrepreneurs. *Journal of Urban Economics*, 52: 341-361.

Figueiredo, O., Guimarães, P., Woodward, D. (2002b). Modeling industrial location in U.S. counties. Working paper series No. 18, Universidade do Minho, Braga.

Franko, A. and Filson, D. (2006). Spin-outs: Knowledge diffusion through employee mobility. *Rand Journal of Economics*, 37: 841-860.

Frenken, K., Van Oort, F. and Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional Studies*, 41: 685-697.

Frenken, K., Cefis, E. andStam, E. (2011). Industrial dynamics and economic geography: a survey. Working paper 11.07, Ecis, Eindhoven.

Fujita, M., Krugman, P. andVenables, A. (1999). The Spatial Economy. Cities, Regions and International Trade. Cambridge (Ma.): MIT Press.

Gartner, W.B. (1985). A conceptual framework for describing the phenomenon of new venturecreation. *Academy of Management Review*, 10: 696-706.

Glaeser, E.L., and Kerr, W. (2009). Local industrial conditions and entrepreneurship: How much of the spatial distribution can we explain? *Journal of Economics & Management Strategy*, 18: 623-663.

Gompers, P., Lerner, J. and Scharfstein, D. (2005). Entrepreneurial spawning: Public corporations and the genesis of new ventures, 1986 to 1999. *Journal of Finance,* 60: 557-615.

Hausman, J., Abrevayab, J. and Scott-Morton, F.-M. (1998). Misclassification of the dependent variable in a discrete-response setting. *Journal of Econometrics*, 87: 239–269.

Hausman, J. and McFadden, D. (1984). Specification Tests for the Multinomial Logit Model. *Econometrica*, 52: 1219-1240.

Hellman, T. (2006). When do employees become entrepreneurs? *Manegement Science*, 53: 919-993.

Henderson, V. and Thisse, J.-F. (2004). *Handbook of Regional and Urban Economics*, Amsterdam: Elsevier.

Hyytinen, A. and Maliranta, M. (2008). When do employees leave their job for entrepreneurship? *Scandinavian Journal of Economics*, 110: 1-21.

Jacobs, J. (1969). *The Economy of Cities*. Random House: New York.

Jovanovic. B. and Nyarko, Y. (1995). A Bayesian learning fitted to a variety of empirical learning curves. *Brookings Papers: Microeconomics*, 247-305.

Kanbur, S. (1979). On risk taking and the personal distribution of income. *Journal of Political Economy*, 87: 769-797.

Keller, W. (2002). Geographic localization of international technology diffusion. *American Economic Review*, 92: 120-142.

Kim, J. and Marschke, G. (2005). Labor mobility of scientists, technological diffusion, and the firm's patenting decision. *Rand Journal of Economics*, 36: 298-317.

Klepper, S. (1996). Entry, exit, growth, and innovation over the product life cycle. *American Economic Review*, 86: 562-583.

Klepper, S. (2001). Employee startups in high-tech industries. *Industrial and Corporate Change*, 10: 639-674.

Klepper, S. (2002). The capabilities of new firms and the evolution of the U.S. automobile industry. *Industrial and Corporate Change*, 11: 645-666.

Klepper, S. (2007). Disagreements, spinoffs, and the evolution of Detroit as the capital of the US automobile industry. *Management Science*, 53: 616-631.

Klepper, S. and Sleeper, S. (2005). Entry by spinoffs. *Management Science*, 51: 1291-1306.

Klepper, S and Thompson, P. (2007). Spinoff Entry in High-tech Industries: Motives and Consequences, in (eds.) Malerba, F. and Brusoni, S., *Perspectives on Innovation*, Cambridge and New York: Cambridge University Press.

Klepper, S. and Thompson, P. (2010). Disagreements and intra-industry spinoffs. *International Journal of Industrial Organization*, 28: 526-538.

Long, J. and Freese, J. (2006) Regression Models for Categorical Dependent Variables Using Stata. Second Edition, Stata Press.

Low, M.B. and MacMillan, I.C. (1988). Entrepreneurship: Past research and future challenges. *Journal of Management*, 14: 139-161.

Machin, S. and van Reenen, J. (1998). Technology and changes in skill structure: Evidence from several OECD countries. *The Quarterly Journal of Economics*, 113: 1215-1244.

Marshall, A. (1890). *Principles of Economics*. MacMillan: London.

McLaren, J. (2000). Globalization and vertical structure. *American Economic Review*, 90: 1239-1254.

Paci, R. and Usai, S. (1999). Externalities, knowledge spillovers and the spatial distribution of innovation. *GeoJournal*, 49: 381-390.

Pakes, A. and Nitzan, S. (1983). Optimum contracts for research personnel, research employment, and the establishment of 'rival' enterprises. *Journal of Labor Economics*, 1: 345-365.

Parker, S. (2009). *The Economics of Entrepreneurship*, Cambridge and New York: Cambridge University Press.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, 94: 1002-1037.

Rosenthal, S.S. and Strange, W.C. (2003). Geography, industrial organization and agglomeration. *The Review of Economics and Statistics*, 85: 377-393.

Rosenthal, S.S. and Strange, W. C. (2010). Small establishments/big effects: Agglomeration, industrial organization and entrepreneurship. In Glaeser, E. (ed.) *Agglomeration Economics*, Chicago: University of Chicago Press.

Small, K.A. and Hsiao, C. (1985). Multinomial logit specification tests. *International Economic Review*, 26: 619-627.

Sorenson, O. and Audia, P.G. (2000). The social structure of entrepreneurial activity: Geographic concentration of footwear production in the United States, 1940-1989. *American Journal of Sociology*, 106: 424-462.

Swann, P., and Prevezer, M. (1996) A comparison of the dynamics of industrial clustering in computing and biotechnology. *Research Policy*, 25: 1139-1157.

Train, K. (2002). *Discrete Choice Methods with Simulation*. Cambridge and New York: Cambridge University Press.

van der Panne, G. and van Beers, C. (2006). On the Marshall-Jacobs controversy: It takes two to tango. *Industrial and Corporate Change*, 15: 877-890.

van Oort, F.G. and Atzema, O. (2004). On the conceptualization of agglomeration economies: The case of new firm formation in the Dutch ICT sector. *The Annals of Regional Science*, 38: 263-290.

Walsh, S., Kirchoff, B. and Boylan, R. (1996). Founder background and entrepreneurial success: Implications for core competence strategy applications to new ventures. In (eds.) Reynolds, P., Birely, S., Butler J., Bygrave, W., Davidsson, P., Gartner, W. and McDougall, P. *Frontiers of Entrepreneurship Research*, Babson College, Wellesey, MA.

Wenting, R. (2008). Spin-off dynamics and the spatial formation of the fashion design industry, 1858-2005. *Journal of Economic Geography*, 8:593-614.

de Wit, G. (1993). Models of self-employment in a competitive market. *Journal of Economic Surveys*, 7: 367-397.

**Tables**

*Table 1.*Annual distribution of individuals' occupational choice for next time period

| Choice | Year | | | | | | Total |
|---|---|---|---|---|---|---|---|
| | *1999* | *2000* | *2001* | *2002* | *2003* | *2004* | |
| *Stay* | 1363365 | 1414839 | 1478238 | 1519827 | 1538078 | 1513636 | 8827983 |
| | 79.25% | 78.63% | 81.16% | 81.44% | 82.70% | 81.70% | 80.85% |
| *Switch* | 241946 | 250991 | 225125 | 201530 | 192250 | 203500 | 1315342 |
| | 14.06% | 13.95% | 12.36% | 10.80% | 10.34% | 10.98% | 12.05% |
| *Exit* | 109939 | 128411 | 112083 | 138534 | 121216 | 125746 | 735929 |
| | 6.39% | 7.14% | 6.15% | 7.42% | 6.52% | 6.79% | 6.74% |
| *Spin-off* | 5114 | 5044 | 5939 | 6191 | 8241 | 9723 | 40252 |
| | 0.30% | 0.28% | 0.33% | 0.33% | 0.44% | 0.52% | 0.37% |
| **Total** | 1720364 | 1799285 | 1821385 | 1866082 | 1859785 | 1852605 | 10919506 |

**Table 2.** Distribution of individuals' occupational choice for next time period per industry sector and size of the originating firm

**(SNI 15 - 37) Manufacturing**

| Size of the firm | Status Next Year | | | | |
|---|---|---|---|---|---|
| | Exit | Stay | Switch | Spin-off | Total |
| *0-10 employees* | 24667 | 268439 | 34515 | 1582 | 329203 |
| | (7.49) | (81.54) | (10.48) | (0.48) | (100.00) |
| *11-50 employees* | 40726 | 500495 | 64016 | 1901 | 607138 |
| | (6.71) | (82.44) | (10.54) | (0.31) | (100.00) |
| *51-100 employees* | 23342 | 322857 | 35682 | 856 | 382737 |
| | (6.10) | (84.35) | (9.32) | (0.22) | (100.00) |
| *101-250 employees* | 31297 | 466159 | 46525 | 935 | 544916 |
| | (5.74) | (85.55) | (8.54) | (0.17) | (100.00) |
| *250-500 employees* | 23785 | 370157 | 33998 | 666 | 428606 |
| | (5.55) | (86.36) | (7.93) | (0.16) | (100.00) |
| *501+ employees* | 89924 | 1442143 | 125773 | 2155 | 1659995 |
| | (5.42) | (86.88) | (7.58) | (0.13) | (100.00) |
| | | | | | |
| *Total* | 233741 | 3370250 | 340509 | 8095 | 3952595 |
| | (5.91) | (85.27) | (8.61) | (0.20) | (100.00) |

**Table 2.** (continued)

**(SNI 38 - 64) Low-end services**

| Size of the firm | Status Next Year | | | | |
|---|---|---|---|---|---|
| | Exit | Stay | Switch | Spin-off | Total |
| *0-10 employees* | 100542 | 933493 | 163695 | 8289 | 1206019 |
| | (8.34) | (77.40) | (13.57) | (0.69) | (100.00 |
| *11-50 employees* | 88942 | 953365 | 174102 | 5444 | 1221853 |
| | (7.28) | (78.03) | (14.25) | (0.45) | (100.00 |
| *51-100 employees* | 27833 | 316550 | 55962 | 1319 | 401664 |
| | (6.93) | (78.81) | (13.93) | (0.33) | (100.00 |
| *101-250 employees* | 27915 | 348689 | 56899 | 1351 | 434854 |
| | (6.42) | (80.19) | (13.08) | (0.31) | (100.00 |
| *250-500 employees* | 17914 | 229923 | 37306 | 725 | 285868 |
| | (6.27) | (80.43) | (13.05) | (0.25) | (100.00 |
| *501+ employees* | 91665 | 1077903 | 153027 | 3020 | 1325615 |
| | (6.91) | (81.31) | (11.54) | (0.23) | (100.00 |
| | | | | | |
| *Total* | 354811 | 3859923 | 640991 | 20148 | 4875873 |
| | (7.28) | (79.16) | (13.15) | (0.41) | (100.00 |

*Notes: Percentages in parentheses*

*Table 2.* *(continued)*

### (SNI 65 - 71) Financial and real-estate services

| Size of the firm | Status Next Year | | | | |
| --- | --- | --- | --- | --- | --- |
| | *Exit* | *Stay* | *Switch* | *Spin-off* | *Total* |
| *0-10 employees* | 6069 | 67828 | 9764 | 595 | 84256 |
| | (7.20) | (80.50) | (11.59) | (0.71) | (100.00) |
| *11-50 employees* | 4892 | 58909 | 9536 | 314 | 73651 |
| | (6.64) | (79.98) | (12.95) | (0.43) | (100.00) |
| *51-100 employees* | 2765 | 34732 | 4377 | 126 | 42000 |
| | (6.58) | (82.70) | (10.42) | (0.30) | (100.00) |
| *101-250 employees* | 3467 | 46653 | 6221 | 120 | 56461 |
| | (6.14) | (82.63) | (11.02) | (0.21) | (100.00) |
| *250-500 employees* | 1492 | 26442 | 2531 | 84 | 30549 |
| | (4.88) | (86.56) | (8.29) | (0.27) | (100.00) |
| *501+ employees* | 2483 | 30485 | 4189 | 93 | 37250 |
| | (6.67) | (81.84) | (11.25) | (0.25) | (100.00) |
| | | | | | |
| *Total* | 21168 | 265049 | 36618 | 1332 | 324167 |
| | (6.53) | (81.76) | (11.30) | (0.41) | (100.00) |

*Table 2.* *(continued)*

### (SNI 72 - 74) Knowledge intensive business services

| Size of the firm | Status Next Year | | | | |
| --- | --- | --- | --- | --- | --- |
| | *Exit* | *Stay* | *Switch* | *Spin-off* | *Total* |
| *0-10 employees* | 25434 | 285059 | 51093 | 3341 | 364927 |
| | (6.97) | (78.11) | (14.00) | (0.92) | (100.00) |
| *11-50 employees* | 27272 | 281560 | 69137 | 3153 | 381122 |
| | (7.16) | (73.88) | (18.14) | (0.83) | (100.00) |
| *51-100 employees* | 12001 | 114376 | 30768 | 994 | 158139 |
| | (7.59) | (72.33) | (19.46) | (0.63) | (100.00) |
| *101-250 employees* | 14365 | 146102 | 35578 | 974 | 197019 |
| | (7.29) | (74.16) | (18.06) | (0.49) | (100.00) |
| *250-500 employees* | 9748 | 112260 | 27090 | 603 | 149701 |
| | (6.51) | (74.99) | (18.10) | (0.40) | (100.00) |
| *501+ employees* | 37389 | 393404 | 83558 | 1612 | 515963 |
| | (7.25) | (76.25) | (16.19) | (0.31) | (100.00) |
| | | | | | |
| *Total* | 126209 | 1332761 | 297224 | 10677 | 1766871 |
| | (7.14) | (75.43) | (16.82) | (0.60) | (100.00) |

*Notes: Percentages in parentheses*

***Table 3.*** *Summary statistics of explanatory variables, 8566321 observations.*

| | Mean | SD | Min | Max |
|---|---|---|---|---|
| Log of regional size | 11.00 | 1.41 | 5.78 | 12.98 |
| PS index | 0.09 | 0.28 | -0.99 | 0.97 |
| UV index | 3.01 | 0.10 | 2.10 | 3.18 |
| RV index | 1.86 | 0.23 | 0.88 | 2.18 |
| ASIZE index | -0.14 | 0.40 | -0.98 | 0.78 |
| Tenure | 5.85 | 5.49 | 0 | 18 |
| Age | 40.39 | 12.31 | 16 | 99 |
| Age$^2$ | 1782.99 | 1028.27 | 256 | 9801 |
| Male | 0.69 | 0.46 | 0 | 1 |
| Log of wage† | 12.25 | 0.66 | 4.60 | 17.07 |
| Education lvl 1 | 0.08 | 0.27 | 0 | 1 |
| Education lvl 2 | 0.14 | 0.35 | 0 | 1 |
| Education lvl 3 | 0.32 | 0.47 | 0 | 1 |
| Education lvl 4 | 0.24 | 0.43 | 0 | 1 |
| Education lvl 5 | 0.12 | 0.32 | 0 | 1 |
| Education lvl 6 | 0.10 | 0.30 | 0 | 1 |
| Education lvl 7 | 0.004 | 0.07 | 0 | 1 |
| Age of firm | 10.48 | 5.21 | 2 | 18 |
| Declining Employment | 0.37 | 0.48 | 0 | 1 |
| Declining Sales | 0.28 | 0.45 | 0 | 1 |
| Log of productivity* | 6.12 | 1.10 | -2.30 | 12.15 |
| Knowledge intensity** | 0.10 | 0.15 | 0 | 1 |
| Size 1-10 | 0.15 | 0.35 | 0 | 1 |
| Size 11-50 | 0.20 | 0.40 | 0 | 1 |
| Size 51-100 | 0.10 | 0.29 | 0 | 1 |
| Size 101-250 | 0.12 | 0.33 | 0 | 1 |
| Size 251-500 | 0.09 | 0.28 | 0 | 1 |
| Size 501 + | 0.34 | 0.47 | 0 | 1 |
| Industry category 1 | 0.40 | 0.49 | 0 | 1 |
| Industry category 2 | 0.44 | 0.50 | 0 | 1 |
| Industry category 3 | 0.03 | 0.17 | 0 | 1 |
| Industry category 4 | 0.13 | 0.34 | 0 | 1 |

*\*Number of observations: 8529729, \*\*number of observations: 8565090(due to loss of observations when taking ratios where the denominator equals zero or taking logs of negative numbers), †wage in 1000s of Swedish Kronas.*

**Table 4.** *MNL Estimation. Marginal effects*

| Variable | Stay | Switch | Exit | Spin-off |
|---|---|---|---|---|
| Log of regional size | -0.414*** | 0.308*** | 0.081*** | 0.025*** |
| | [0.015] | [0.007] | [0.005] | [0.003] |
| PS index | 0.553*** | -0.371*** | -0.196*** | 0.014 |
| | [0.050] | [0.041] | [0.027] | [0.010] |
| UV index | 3.479*** | -2.810*** | -0.379*** | -0.290*** |
| | [0.134] | [0.110] | [0.075] | [0.024] |
| RV index | -0.821*** | 0.343*** | 0.338*** | 0.140*** |
| | [0.100] | [0.082] | [0.056] | [0.018] |
| ASIZE index | -1.705*** | 1.642*** | -0.171*** | 0.234*** |
| | [0.054] | [0.044] | [0.030] | [0.011] |
| Tenure | 1.009*** | -0.796*** | -0.183*** | -0.029*** |
| | [0.003] | [0.003] | [0.002] | [0.001] |
| Age | 1.073*** | -0.309*** | -0.819*** | 0.055*** |
| | [0.007] | [0.006] | [0.004] | [0.001] |
| Age2 | -0.012*** | 0.001*** | 0.010*** | -0.000*** |
| | [0.000] | [0.000] | [0.000] | [0.000] |
| Male† | -1.420*** | 0.864 | 0.286*** | 0.272*** |
| | [0.024] | [0.020] | [0.014] | [0.004] |
| Log of wage | 7.038*** | -2.293*** | -4.444*** | -0.301*** |
| | [0.022] | [0.017] | [0.011] | [0.003] |
| Education lvl 2† | -1.551*** | 2.360*** | -0.914*** | 0.104*** |
| | [0.068] | [0.064] | [0.022] | [0.012] |
| Education lvl 3† | -2.084*** | 2.471*** | -0.556*** | 0.168*** |
| | [0.061] | [0.057] | [0.022] | [0.011] |
| Education lvl 4† | -2.725*** | 2.709*** | -0.215*** | 0.231*** |
| | [0.065] | [0.061] | [0.024] | [0.013] |
| Education lvl 5† | -4.581*** | 3.939*** | 0.279*** | 0.363*** |
| | [0.079] | [0.074] | [0.030] | [0.018] |
| Education lvl 6† | -5.734*** | 5.540*** | -2.370*** | 0.431*** |
| | [0.092] | [0.087] | [0.034] | [0.022] |
| Education lvl 7† | -6.903*** | 5.473*** | 0.876*** | 0.544*** |
| | [0.252] | [0.220] | [0.139] | [0.063] |
| Age of firm | 0.146*** | -0.106*** | -0.039*** | -0.0001** |
| | [0.002] | [0.002] | [0.001] | [0.000] |
| Declining Employment† | -1.348*** | 0.840*** | 0.464*** | 0.044*** |
| | [0.026] | [0.021] | [0.015] | [0.005] |
| Declining Sales† | -2.705*** | 2.020*** | 0.633*** | 0.052*** |
| | [0.029] | [0.024] | [0.016] | [0.005] |
| Log of Productivity | 0.754*** | -0.521*** | -0.224*** | -0.008*** |
| | [0.009] | [0.007] | [0.005] | [0.002] |
| Knowledge intensity | -0.870*** | 0.432*** | 0.274*** | 0.163*** |
| | [0.095] | [0.075] | [0.058] | [0.014] |
| Size 11-50† | -1.646*** | 1.290*** | 0.443 | -0.088*** |
| | [0.040] | [0.033] | [0.023] | [0.005] |
| Size 51-100† | -2.274*** | 1.635*** | 0.784*** | -0.144*** |
| | [0.051] | [0.043] | [0.030] | [0.006] |
| Size 101-250† | -2.183*** | 1.625*** | 0.743*** | -0.184*** |
| | [0.049] | [0.041] | [0.028] | [0.005] |
| Size 251-500† | -2.191*** | 1.557*** | 0.818*** | -0.200*** |
| | [0.055] | [0.046] | [0.032] | [0.006] |
| Size 501 +† | -1.384*** | 0.950*** | 0.724*** | -0.290*** |
| | [0.038] | [0.031] | [0.022] | [0.005] |
| | | | | |
| Additional controls | *Year and industry category dummies* | *Year and industry category dummies* | *Year and industry category dummies* | *Year and industry category dummies* |
| *Y = Pr(Choice = 1)* | *0.86897* | *0.08559* | *0.04117* | *0.00425* |

*Notes: ***significant at 0.1%, **significant at 1%, *significant at 5%, †dummy variable, number of obs: 8527145, marginal effects and standard errors (in brackets) have been multiplied by 100.*

*Table 5A.* MNL estimations, regional variables, individual- and firm-level controls. Marginal effects.

| Variable | Spin-off (I) | Spin-off (II) | Spin-off (III) | Spin-off (IV) |
|---|---|---|---|---|
| Log of regional size | 0.024*** [0.003] | - | - | 0.025*** [0.003] |
| PS index | -0.012*** [0.011] | - | - | 0.014 [0.010] |
| UV index | -0.362*** [0.003] | - | - | -0.290*** [0.024] |
| RV index | 0.182*** [0.002] | - | - | 0.140*** [0.018] |
| ASIZE index | 0.483*** [0.012] | - | - | 0.234*** [0.011] |
| Tenure | - | -0.035*** [0.001] | - | -0.029*** [0.001] |
| Age | - | 0.061*** [0.001] | - | 0.055*** [0.001] |
| Age2 | - | -0.000*** [0.000] | - | -0.000*** [0.000] |
| Male† | - | 0.301*** [0.004] | - | 0.272*** [0.004] |
| Log of wage | - | -0.328*** [0.003] | - | -0.301*** [0.003] |
| Education lvl 2† | - | 0.127*** [0.013] | - | 0.104*** [0.012] |
| Education lvl 3† | - | 0.181*** [0.012] | - | 0.168*** [0.011] |
| Education lvl 4† | - | 0.251*** [0.014] | - | 0.231*** [0.013] |
| Education lvl 5† | - | 0.375*** [0.019] | - | 0.363*** [0.018] |
| Education lvl 6† | - | 0.517*** [0.022] | - | 0.431*** [0.022] |
| Education lvl 7† | - | 0.595*** [0.066] | - | 0.544*** [0.063] |
| Age of firm | - | - | -0.008*** [0.000] | -0.0001** [0.000] |
| Declining Employment† | - | - | 0.028*** [0.005] | 0.044*** [0.005] |
| Declining Sales† | - | - | 0.048*** [0.005] | 0.052*** [0.005] |
| Log of Productivity | - | - | -0.017*** [0.002] | -0.008*** [0.002] |
| Knowledge intensity | - | - | 0.197*** [0.013] | 0.163*** [0.014] |
| Size 11-50† | - | - | -0.142*** [0.005] | -0.088*** [0.005] |
| Size 51-100† | - | - | -0.212*** [0.006] | -0.144*** [0.006] |
| Size 101-250† | - | - | -0.265*** [0.005] | -0.184*** [0.005] |
| Size 251-500† | - | - | -0.289*** [0.005] | -0.200*** [0.006] |
| Size 501 +† | - | - | -0.425*** [0.005] | -0.290*** [0.005] |
| Additional controls | *Year and industry category dummies* | *Year and industry category dummies* | *Year and industry category dummies* | *Year and industry category dummies* |

*Notes: ***significant at 0.1%, **significant at 1%, *significant at 5%, †dummy variable, marginal effects and standard errors (in brackets) have been multiplied by 100.*

**Table 5B.** *Robustness check with stepwise elimination of non-spatial indices (all industries).*

| Variable | Benchmark model | Firm and industry characteristics dropped | Only year dummies and regional indices |
|---|---|---|---|
| Log of regional size | .025*** | .039*** | .043*** |
| | [.003] | [.002] | [.003] |
| PS index | .014 | .045*** | .026** |
| | [.010] | [.009] | [.010] |
| UV index | -.290*** | -.371*** | -.430*** |
| | [.024] | [.024] | [.027] |
| RV index | .140*** | .129*** | .155*** |
| | [.018] | [.018] | [.020] |
| ASIZE index | .234*** | .406*** | .557*** |
| | [.011] | [.007] | [.007] |

*Notes: ***significant at 1%, **significant at 5%, *significant at 10%, marginal effects and standard errors (in brackets) have been multiplied by 100.*

**Table 6.** *MNL estimation on high-tech manufacturing and knowledge intensive services (SNI 29-33 and 72-74). Marginal effects*

| Variable | Stay | Switch | Exit | Spin-off |
|---|---|---|---|---|
| Log of regional size | -0.835*** | 0.645*** | 0.176*** | 0.014* |
|  | [0.034] | [0.029] | [0.018] | [0.006] |
| PS index | 0.738*** | -0.631*** | -0.196*** | 0.089*** |
|  | [0.114] | [0.098] | [0.060] | [0.023] |
| UV index | 4.292*** | -3.218*** | -0.653*** | -0.421*** |
|  | [0.299] | [0.026] | [0.151] | [0.006] |
| RV index | 0.598*** | -0.821*** | 0.012 | 0.210*** |
|  | [0.230] | [0.198] | [0.120] | [0.004] |
| ASIZE index | 2.025*** | -0.969*** | -1.294*** | 0.238*** |
|  | [0.148] | [0.123] | [0.081] | [0.027] |
| Tenure | 1.064*** | -0.851*** | -0.178*** | -0.034*** |
|  | [0.007] | [0.006] | [0.003] | [0.001] |
| Age | 1.249*** | -0.403*** | -0.895*** | 0.049*** |
|  | [0.016] | [0.014] | [0.007] | [0.003] |
| Age2 | -0.014*** | 0.002*** | 0.012*** | -0.000*** |
|  | [0.000] | [0.000] | [0.000] | [0.000] |
| Male† | -1.037*** | 0.525*** | 0.242*** | 0.270*** |
|  | [0.053] | [0.045] | [0.027] | [0.009] |
| Log of wage | 5.794*** | -1.5381*** | -4.010*** | -0.246*** |
|  | [0.043] | [0.040] | [0.021] | [0.006] |
| Education lvl 2† | -2.519*** | 3.117*** | -0.750*** | 0.152*** |
|  | [0.181] | [0.175] | [0.051] | [0.037] |
| Education lvl 3† | -2.725*** | 3.236*** | -0.720*** | 0.209*** |
|  | [0.159] | [0.153] | [0.047] | [0.033] |
| Education lvl 4† | -3.455*** | 3.659*** | -0.527*** | 0.323*** |
|  | [0.164] | [0.158] | [0.050] | [0.038] |
| Education lvl 5† | -4.662*** | 4.644*** | -0.445*** | 0.463*** |
|  | [0.181] | [0.173] | [0.055] | [0.045] |
| Education lvl 6† | -4.858*** | 5.214*** | -0.800*** | 0.444*** |
|  | [0.184] | [0.177] | [0.059] | [0.045] |
| Education lvl 7† | -5.892*** | 4.949*** | 0.355* | 0.587*** |
|  | [0.363] | [0.334] | [0.11] | [0.100] |
| Age of firm | 0.352*** | -0.271*** | -0.078*** | -0.000 |
|  | [0.005] | [0.004] | [0.003] | [0.001] |
| Declining Employment† | -2.385*** | 1.711*** | 0.609*** | 0.065*** |
|  | [0.059] | [0.051] | [0.031] | [0.010] |
| Declining Sales† | -2.747*** | 2.132*** | 0.544*** | 0.071*** |
|  | [0.062] | [0.052] | [0.033] | [0.010] |
| Log of Productivity | 0.577*** | -0.437*** | -0.141*** | 0.000 |
|  | [0.014] | [0.011] | [0.008] | [0.003] |
| Knowledge intensity | 0.108 | 0.008 | -0.228** | 0.112*** |
|  | [0.140] | [0.116] | [0.080] | [0.021] |
| Size 11-50† | -2.610*** | 1.802*** | 0.848*** | -0.039*** |
|  | [0.010] | [0.085] | [0.056] | [0.012] |
| Size 51-100† | -3.247*** | 2.229*** | 1.131*** | -0.114*** |
|  | [0.123] | [0.106] | [0.071] | [0.013] |
| Size 101-250† | -2.936*** | 2.085*** | 1.053*** | -0.204*** |
|  | [0.113] | [0.097] | [0.064] | [0.012] |
| Size 251-500† | -4.088*** | 2.972*** | 1.392*** | -0.276*** |
|  | [0.125] | [0.108] | [0.071] | [0.011] |
| Size 501 +† | -2.729*** | 1.981*** | 1.108*** | -0.359*** |
|  | [0.088] | [0.076] | [0.048] | [0.012] |
|  |  |  |  |  |
| Additional controls | *Year and industry category dummies* | *Year and industry category dummies* | *Year and industry category dummies* | *Year and industry category dummies* |
| *Y = Pr(Choice = 1)* | *0.85362* | *0.10105* | *0.04052* | *0.00476* |

Notes: ***significant at 0.1%, **significant at 1%, *significant at 5%, †dummy variable, number of obs: 2019483, marginal effects and standard errors (in brackets) have been multiplied by 100

**Table 7A.** *MNL estimations, regional variables, individual- and firm-level controls.*

*High-tech manufacturing and knowledge intensive services  (SNI 29-33 and 72-74). Marginal effects.*

| Variable | Spin-off (I) | Spin-off (II) | Spin-off (III) | Spin-off (IV) |
|---|---|---|---|---|
| Log of regional size | 0.023*** [0.007] | - | - | 0.014** [0.006] |
| PS index | 0.023 [0.026] | - | - | 0.089*** [0.023] |
| UV index | -0.535*** [0.064] | - | - | -0.421*** [0.006] |
| RV index | 0.281*** [0.049] | - | - | 0.210*** [0.004] |
| ASIZE index | 0.399*** [0.032] | - | - | 0.238*** [0.027] |
| Tenure | - | -0.039*** [0.001] | - | -0.034*** [0.001] |
| Age | - | 0.053*** [0.003] | - | 0.049*** [0.003] |
| Age2 | - | -0.000*** [0.000] | - | -0.000*** [0.000] |
| Male† | - | 0.289*** [0.009] | - | 0.270*** [0.009] |
| Log of wage | - | -0.263*** [0.007] | - | -0.246*** [0.006] |
| Education lvl 2† | - | 0.201*** [0.042] | - | 0.152*** [0.037] |
| Education lvl 3† | - | 0.268*** [0.037] | - | 0.209*** [0.033] |
| Education lvl 4† | - | 0.423*** [0.042] | - | 0.323*** [0.038] |
| Education lvl 5† | - | 0.589*** [0.052] | - | 0.463*** [0.045] |
| Education lvl 6† | - | 0.631*** [0.051] | - | 0.444*** [0.045] |
| Education lvl 7† | - | 0.766*** [0.144] | - | 0.587*** [0.100] |
| Age of firm | - | - | -0.008*** [0.001] | -0.000 [0.001] |
| Declining Employment† | - | - | 0.052*** [0.011] | 0.065*** [0.010] |
| Declining Sales† | - | - | 0.069*** [0.011] | 0.071*** [0.010] |
| Log of Productivity | - | - | 0.000 [0.003] | 0.000 [0.003] |
| Knowledge intensity | - | - | 0.184*** [0.013] | 0.112*** [0.021] |
| Size 11-50† | - | - | -0.092*** [0.012] | -0.039*** [0.012] |
| Size 51-100† | - | - | -0.179*** [0.013] | -0.114*** [0.013] |
| Size 101-250† | - | - | -0.278*** [0.011] | -0.204*** [0.012] |
| Size 251-500† | - | - | -0.354*** [0.011] | -0.276*** [0.011] |
| Size 501 +† | - | - | -0.466*** [0.012] | -0.359*** [0.012] |
| Additional controls | *Year and industry category dummies* | *Year and industry category dummies* | *Year and industry category dummies* | *Year and industry category dummies* |

*Notes: \*\*\*significant at 0.1%, \*\*significant at 1%, \*significant at 5%, †dummy variable, marginal effects and standard errors (in brackets) have been multiplied by 100.*

*Table 7B.Robustness check with stepwise elimination of non-spatial indices (KIBS).*

| Variable | Benchmark model | Firm and industry characteristics dropped | Only year dummies and regional indices |
|---|---|---|---|
| Log of regional size | .014* | .041*** | .047*** |
| | [.006] | [.006] | [.006] |
| PS index | .089*** | .154*** | .136*** |
| | [.023] | [.021] | [.023] |
| UV index | -.421*** | -.516*** | -.526*** |
| | [.006] | [.055] | [.060] |
| RV index | .210*** | .178*** | .194*** |
| | [.004] | [.041] | [.045] |
| ASIZE index | .238*** | .607*** | .722*** |
| | [.027] | [.014] | [.014] |

*Notes: ***significant at 1%, **significant at 5%, *significant at 10%, marginal effects and standard errors (in brackets) have been multiplied by 100.*

*Table 8. The Hausman-McFadden IIA tests on the full MNL specification.*

| Omitted Alternative | MFH statistic | Degrees of freedom | Threshold value $\alpha=0.01$ | Evidence |
|---|---|---|---|---|
| Switch | 1200.00 | 68 | 97,2866 | Reject IIA |
| Exit | -640.00 | 68 | 97,2866 | Accept IIA |
| Spin | 219.477 | 68 | 97,2866 | Reject IIA |

*Table 9. The suest-basedHausman IIA tests on the full MNL specification.*

| Omitted Alternative | SBH statistic | Degrees of freedom | Threshold value $\alpha=0.01$ | Evidence |
|---|---|---|---|---|
| Switch | 1.8e+04 | 70 | 99,0858 | Reject IIA |
| Exit | 5.0e+04 | 70 | 99,0858 | Reject IIA |
| Spin | 3590.872 | 70 | 99,0858 | Reject IIA |

*Table 10. The Small-Hsiao IIA tests on the full MNL specification.*

| Omitted Alternative | $L_R(\widehat{\beta}_{12U})$ | $L_R(\widehat{\beta}_{2R})$ | SH statistic | Degrees of freedom | Threshold value $\alpha=0.01$ | Evidence |
|---|---|---|---|---|---|---|
| Switch | -9.00e+05 | -8.99e+05 | 313.505 | 70 | 99,0858 | Reject IIA |
| Exit | -1.41e+06 | -1.41e+06 | 650.287 | 70 | 99,0858 | Reject IIA |
| Spin | -2.09e+06 | -2.09e+06 | 87.488 | 70 | 99,0858 | Accept IIA |

*Table 11.Marginal effects of regional indices on Spin-off across a series of robustness checks (all industries).*

| Variable | Benchmark model | Model I | Model II | Model III | Model IV | Model V |
|---|---|---|---|---|---|---|
| Log of regional size | .025*** | .025** | -.010 | .018 | .018 | .001 |
| PS index | .014 | .014 | .003 | -.012 | -.012 | -.019 |
| UV index | -.290*** | -.290*** | -.039 | -.433*** | -.433*** | -.295*** |
| RV index | .140*** | .140** | .159*** | .205*** | .205** | .189** |
| ASIZE index | .234*** | .234*** | .223*** | .218*** | .218*** | .211*** |

*Notes: ***significant at 1%, **significant at 5%, *significant at 10%, marginal effects have been multiplied by 100. Model I: MNL estimated with clustered standard errors. Model II: MNL including agglomeration dummies estimated with clustered standard errors. Model III: Basic Nested logit. Model IV: Nested logit with clustered standard errors. Model V: Nested logit including agglomeration dummies estimated with clustered standard errors.*

*Table 12.*Marginal effects of regional indices on Spin-off across a series of robustness checks (KIBS).

| Variable | Benchmark model | Model I | Model II | Model III | Model IV | Model V |
|---|---|---|---|---|---|---|
| Log of regional size | .014* | .014 | -.022** | -.011 | -.011 | -.027 |
| PS index | .089*** | .089** | .069 | .041 | .041 | .032 |
| UV index | -.421*** | -.421*** | -.162 | -.728*** | -.728*** | -.650** |
| RV index | 0.210*** | .210** | .232*** | .307** | .307** | .334** |
| ASIZE index | 0.238*** | .238*** | .224*** | .125 | .125 | .121 |

*Notes: ***significant at 1%, **significant at 5%, *significant at 10%, marginal effects have been multiplied by 100. Model I: MNL estimated with clustered standard errors. Model II: MNL including agglomeration dummies estimated with clustered standard errors. Model III: Basic Nested logit. Model IV: Nested logit with clustered standard errors. Model V: Nested logit including agglomeration dummies estimated with clustered standard errors.*

*Figure 1.*Predicted probabilities calculated at the mean vector of the rest of independent variables for KIBS industries in 2004
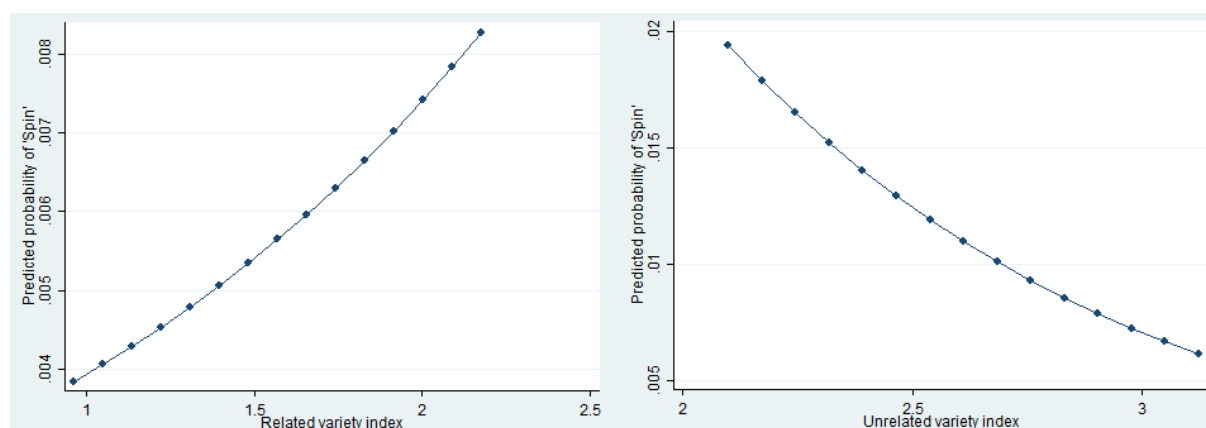
**Appendix A**

**An augmented disagreement model**

Following Klepper and Thomson (2007; 2010), the basic theoretical structure to capture entrepreneurial entry through spin-offs is as follows. By definition a spin-off ($s$) occurs from an already existing firm enrolling $n$ employees, where the firm is undertaking activity $y$. For simplicity we assume that all spin-offs locate in the same region as the parent firm, but a distance decay effect could be introduced (see below). Based on individual ($i$) heterogeneous abilities, originating in $i$'s experience and education, $i$ believes that $y$ is the optimal strategy at time $t$ for firm $j$ to reach a target $\theta_{it}$, $y_{jt} = \theta_{it}$. Employees aim at maximizing the value ($v$) of the firm $j$, where $v = -(\theta - y)^2$.

The actual activity undertaken by the firm is a weighted average of all employees' beliefs, $y_{jt} = \bar{\theta}_t = \sum_{i=1}^{n} \varphi_i \theta_{it}$ where $\varphi_i$ are the employee's decision weights assumed invariant over time, $\sum_{i=1}^{n} \varphi_i = 1$. The target ($\theta$) at time $t$ is unknown to all $n$ employees but assumed to be drawn from a normal distribution with mean $\theta_{it}$ and variance $\sigma_{it}^2$. Hence, based on the respective employee's beliefs, the expected pay-off from a strategy $y_{jt} = \theta_{it}$ is $v_{it} = -\sigma_{it}^2$. Taking into account the other $n$ employees expected value of firm $j$ yields the following expression for employee $i$,

$$E_{it}(v) = -E_{it}[(\theta - \bar{\theta}_{it})^2] = -E_{it}[((\theta - \theta_{it}) + (\theta_{it} - \bar{\theta}_t))^2]$$
$$= -E_{it}[(\theta - \theta_{it})^2] - (\theta_{it} - \bar{\theta}_t)^2 = -\sigma_{it}^2 - (\theta_{it} - \theta_t)^2 \qquad (A1)$$

using the fact that any Bayesian posterior is unbiased, i.e. $E_{it}(\theta - \theta_{it}) = 0$.

If employee $i$'s optimal strategy deviates from the weighed actual strategy decision taken by the firm, the likelihood of a spin-off, or moving to another firm, increases. But disagreement with regard to strategies is not sufficient to alter an individual's occupational strategy, other costs associated with changing occupation must also be taken into account. For simplicity, assume an employee that is considering to set up her own firm, either alone or together with individuals that have been identified to share the exact same beliefs. Let $c$ represent a vector of costs associated with establishing a firm in region $r$. If the expected value of the spin-off ($v^s$), $E_{it}(v^s) = -c_{rt} - \sigma_{it}^2$, a new firm will be spun off from an incumbent when,

$$v^2{}_{it}^s = (\theta_{it} - \bar{\theta}_t)^2 \geq c_{rt} \qquad (A2).$$

Hence, the employee's decision to set up a new firm depends on the subjective evaluations of future pay-off and costs associated with alternative occupations, as well as regional, industrial and firm factors. More generally, the emergence of spin-offs can be attributed heterogeneous individual capabilities such as education and tenure ($H$), firm- and industry-specific factors ($Y$) related to age and previous performance, as well as regional factors ($X$) stemming from the size distribution of firms, agglomeration economies and market size. Hence, the individual decision to spin off ($s$) a new venture can be expressed more generally as,

$$s_{it} = f(H_{it}, Y_t, X_{rt}) \qquad\qquad\qquad \text{A3}$$

where *H, Y* and *X* represents vectors of explanatory variables.[38] Dynamics are introduced as heterogeneous individuals receive noisy information that will alter their expectations over time.[39]

---

[38]We abstain from formally introducing a distance decay effect since 85% of spin-offs locate in the same region as the parent firm. However, an alternative specification, where regional advantages diminishes with distance to the home region, could be $s_h = X e_r^{-\alpha d}$, where d represent distance between home (h) and other regions (r)and $0 < \alpha < 1$. If the spin-off locates in the home region, then $e_r^{-\alpha d} = 1$ and the spin-off can reap the full benefits of home regional advantages.

[39] More precisely, $i_{it} \equiv \theta + \varepsilon_{it}$ ($\varepsilon_{it} = N(o, \sigma_\varepsilon^2)$), where each individual believes that received information have the variance $\gamma \sigma_\varepsilon^2$ while other individuals' information have variance $\gamma \zeta \sigma_\varepsilon^2$.

**Appendix B.**

**Founder identification and sensitivity analyses based on alternative definitions**

To identify a spin-off, we combine information from two different databases. The first contains information regarding all individuals active in the Swedish labour market (both employers and employees). The second contains information regarding all active firms in Sweden. Several primary data sources, including tax records, social security records, health and education records, have been combined to construct these databases. The use of a single and unique personal identification number in every function of an individual's life enables us to link information among these diverse data sources.

The paper defines a spin-off fairly broadly. A spin-off occurs when *all* of criteria (i)-(v) below hold (see further below for the role of criteria (vi)-(viii)).

i) Individual $x$ was employed during period $t$ at the firm with organisation number $m$;
ii) During period $t+1$, individual $x$ *owned* the firm with organisation number $g$;
iii) The mappings $(x,t) \rightarrow m$ and $(x,t+1) \rightarrow g$ are made by matching the individual and firm databases described above;
iv) Organisation number $g$ did not exist in the database during any period before $t+1$;
v) Organisation number $m$ continued to exist during period $t+1$;
vi) The firms that correspond to organisation numbers $m$ and $g$ in the firm registry database belong to the same sector;
vii) The firm that corresponds to organisation number $g$ employed more than E employees at time $t+1$; and
viii) The firm that corresponds to organisation number $g$ is recorded as a corporate firm during time $t+1$.

Note that this is one of the few registry data-based studies in which all the above criteria are directly observable and identified, i.e., without resorting to methods involving indirect inferences, such as those employed in Hytinnen and Maliranta (2008), where firm ownership (criterion (ii)) is assumed based on whether individuals are insured under the Self-employed Persons' Pension Act. Adding one or more of criteria (vi)-(viii) results in a much narrower definition of spin-off. However, the alternative definition is not necessarily more refined, given the general absence of a broadly accepted definition for spin-off.

**Table A1.** *Relative frequency of spin-offs under alternative definitions.*

| Sectorial aggregation | Intra-industry spin-offs | Inter-industry spin-offs |
|---|---|---|
| Branch | 47.3% | 38.9% |
| 2nd digit SNI | 26.6% | 59.5% |
| 5th digit SNI | 16.5% | 69.7% |

Notes: In 13.8% of the cases, our data do not allow a definite conclusion regarding the sector in which entry took place.

Furthermore, the use of a narrow definition drastically reduces the relative frequency of a spin-off in the data, rendering it extremely rare. Table A1 illustrates this phenomenon under a series of alternative and gradually narrower definitions that filter out an increasing portion of spin-offs that are recorded as *inter-industry*. Similar patterns are observed when gradual filtering addresses the size of the spin-off firm.

Given the significant variation in the relative frequencies reported in Table A1, we explore the qualitative and quantitative consequences of adopting a narrower definition of spin-off in the context of this paper. Because the primary interest focuses on the marginal effects of regional indices, we focus on the changes occurring in these estimates under alternative definitions. Table A2 plots the results from two such alternative MNL estimations. In the first (alternative definition 1), only start-ups that became active in the same 2-digit industry are classified as spin-offs, whereas the remaining entrepreneurial activity is classified as a simple job change. Therefore, we now adopt a narrower definition that includes an additional requirement, i.e., condition (vi) above. In this case, only 26.6% of the spin-off cases survive the re-definition, as shown in Table A1. This frequency may appear surprising, but it is the logical outcome of a highly refined classification in relation to the scope of this study. For example, consider the neighbouring industries 26 and 27 (in the SNI code, these category numbers denote manufacturers of computers, electronic and optical products and manufacturers of electric apparel, respectively) that are logically close enough for their cross spin-offs to be recorded as intra-industry (it is common that cross innovation is generated by such neighbouring industries) but are, under this definition, filtered out.[40] In the second estimation (alternative definition 2), only start-ups that hired employees were considered spin-offs. Therefore, all entrepreneurial activity at the individual level was recorded as a job switch.[41] This definition adds condition (vii) to conditions (i)-(v). The results displayed show that the qualitative changes are minimal, i.e., the direction and significance of the marginal effects remain intact. The quantitative changes can be partially explained by the nature of the maximum likelihood estimation in the MNL settings, which presses the predicted share of each alternative to equal the actual share of that alternative in the sample.

Another alternative and narrower definition may consider only corporate firms as spin-offs (i.e., the definition in which, apart from conditions (i) to (v), condition (viii) is also required to hold). Using this definition results in relabeling the vast majority (83.9%) of cases initially considered spin-offs. This very large percentage results from the considerable entry costs a newly established firm must undertake to be classified as a corporate firm.[42] Table A3 displays the marginal effects of the most notable variables from an MNL model in which corporate firm spin-offs are included as a separate (fifth) alternative. As in the previous cases, the estimates of the related variety, unrelated variety and average firm size, i.e., of the variables that have been shown to be significant throughout the study, are shown to be qualitatively robust.

The exact implications of these findings for the proposed policy interventions (discussed in the concluding remarks) depend on the differences in performance indicators (survival, export activity, employment, etc.) among the various types of spin-offs and the exact relation of these indicators to the underlying objectives of the policy maker. For instance, the policy implications from the estimation made under alternative definition 2 might be preferable for a policy maker that prioritises job creation,

---

[40] We have chosen to undertake the associated sensitivity analyses (included in the new appendix) at this level, i.e., to employ the strictest robustness tests possible by filtering out cases such as those described above without altering the this study's research question, such as if an extremely high resolution, e.g., at the 5th digit, was used to filter out inter-industrial entry. For example, consider industries 31011 and 31012, which both involve manufacturing furniture specifically for commercial use (hotels, restaurants etc.) with minimal differences between the industries. Filtering at this level removes 69.7% of documented spin-off activity: only 16.5% of the cases survive this definition, as in 13.8% of the cases (independent of the aggregation level) our data do not allow a definite conclusion regarding the sector in which entry occurred.

[41] Note that the underlying problem addressed is not misclassification, but redefinition. Under misclassification, i.e. in the presence of a prevalent, widely-accepted definition of spin-off that the research design failed to adopt, alternative estimation techniques would be required (see Hausman et al., 1998).

[42] During the period of this study, these costs, which account for share capital, were approximately SEK 100000.

but the respective implications from the base model might better serve policies that favour inter-sectorial mobility and the creation of new entries that will intensify competition across several sectors. Isolating the precise spin-off definition that might serve best for each policy objective extends beyond the scope of this study.

***Table A2.*** *MNL marginal effects under the base and two narrower definitions: Inter-industry spin-offs and firms with no employees.*

|  | Base model | Alternative definition 1 | Alternative definition 2 |
|---|---|---|---|
| Log of regional size | .025*** | .014*** | .010*** |
| PSindex | .014 | .046*** | .005 |
| Unrelated variety | -.290*** | -.024** | -.100*** |
| Related variety | .140*** | .018** | .040*** |
| ASIZE index | .234*** | .201*** | .087*** |

Notes: In the estimation under *alternative spin-off definition 1*, all cases for which the spin-off firm became active in a 2-digit sector that differs from that of the parent firm at period t+1 are classified as job changes. In the estimation under *alternative spin-off definition 2*, all cases for which the spin-off firm hired no employees during period t+1 are classified as job changes. *significant at 10%, **significant at 5% and ***significant at 1%.

***Table A3.*** *MNL marginal effects under an alternative estimation: Corporate versus non-corporate firms.*

|  | Base model | Corporate firms | Non-corporate firms |
|---|---|---|---|
| Log of regional size | .025*** | .000 | .022*** |
| PSindex | .014 | .002 | .012 |
| Unrelated variety | -.290*** | -.045*** | -.218*** |
| Related variety | .140*** | .010* | .124*** |
| ASIZE index | .234*** | .040*** | .187*** |

Notes: Non-corporate spin-offs comprise 83.9% of all spin-off cases. *significant at 10%, **significant at 5% and ***significant at 1%.