CESIS Electronic Working Paper Series

Paper No. 189

The Firm and the Region as Breeding Grounds for Entrepreneurs

Apostolos Baltzopoulos*

(*CESIS and the Division of Economics, KTH)

August 2009

The Royal Institute of Technology
Centre of Excellence for Science and Innovation Studies (CESIS)
http://www.cesis.se

The firm and the region as breeding grounds for entrepreneurs.

Apostolos Baltzopoulos
Division of Economics, Royal Institute of Technology (KTH), Stockholm, Sweden
(e-mail: apobal@infra.kth.se)

Abstract

The present study carries out a mutli-level analysis of entrepeneurship by considering the choice of the individual to leave his job position to become self-employed. A comprehensive data-set matching the individual to his place of work allows controlling for the characteristics of both the firm and the region he worked in before starting his own firm. The results suggest that small firms spawn entrepreneurs more frequently and individuals working in larger regions that are characterized by larger local markets, higher accumulation of knowledge resources and higher population density are more likely to transcend into entrepreneurship. I also find evidence that people are more likely to select the path of self-employment in the face of weak local competition.

Keywords: Entrepreneurship, self-employment, externalities

JEK classification: D01, O12, O18, R10

1. Introduction

Research in the organizational and regional context in which entrepreneurs emerge is largely based on either aggregate or extremely selective data (Gompers et al., 2008; Klepper and Sleeper, 2005; Reynolds et al., 1994). Hyytinen and Maliranta (2008) propose an individual-level approach in which nation-wide census data is used to identify entrepreneurs through transitions to self-employment. Controlling for individual and firm level characteristics, they show that in the case of the Finnish economy small firms with low R&D intensity spawn entrepreneurs more frequently than their larger, R&D intensive counterparts.

Shane and Venkataraman (2000) point out that the creation of new enterprises is the outcome of the discovery, evaluation, and exploitation of entrepreneurial opportunities. Shane (2000) and Casson (2003) describe the individual dimension of this process while Malecki (1997) and Davidsson et al. (1994) point out the significance of the regional dimension as well. Taking stock of an extensive body of literature that identifies a relationship between innovation, entrepreneurship and the attributes of the local regional economic milieu (Acs et al., 2006; Audretsch and Lehmann, 2005; Freeman, 1974; Klein et al., 1978; McLaren, 2000; Davidsson et al., 1994; Reynolds et al., 1994), the present paper extends the model of Hyytinen and Maliranta (2008) to explore the effects of a region's size, knowledge resources and degree of industry specialization/diversity on an individual's decision to become self-employed, heeding the call of Low and MacMillan (1988) for a multi-level approach in the research of entrepreneurship, a call that has to-date remained mostly unanswered (Davidsson et al., 2001).

While robust results exist to support that urbanization/agglomeration and the presence of relevant knowledge resources have a positive effect on firm birth rates (Reynolds et al., 1994; Andersson and Hellerstedt, 2009) research about whether regional industrial specialization or diversity better facilitate entrepreneurial activity remains largely inconclusive (Feldman and Audretsch, 1999; van Oort and Atzema, 2004; van der Panne and van Beers, 2006; Desrochers and Sautet, 2008).

What sets this paper apart from past studies on the effects of regional attributes on entrepreneurial spawning is that it does not utilize data on aggregate business start-up rates but instead considers the decision of the individual using highly detailed micro-level data.

This approach allows controlling for characteristics of both the founder of the new firm and the incumbent firm when considering the effects of the regional context. Following Hyytinen and Maliranta (2008) this method allows contrasting the decision to become an entrepreneur to the counterfactual and examines start-ups across all industries of the economy rather than a particular industry branch (e.g. Klepper and Sleeper, 2005, focus on the laser industry alone). Prior studies that use start-up rates as the unit of observation fail to take into consideration people of similar backgrounds and characteristics as those of the founders of new firms that choose to stay in their current job or switch to a different employer rather than become self-employed.

Besides extending the model of Hyytinen and Maliranta (2008) I also improve their approach by using a more straightforward measure of entrepreneurial spawning. In this paper I use a methodology similar to the one developed by Nås et al. (2003) to identify individuals that transcend into self-employment in a linked employers-employees data set that completely describes the Swedish economy in the period between 1999 and 2005. The data includes information on individuals' education, occupation, places of origin and residence, as well as organizational, location and financial data for the firms and establishments they own or are employed in, which allows for a multi-level analysis.

The results show that while urbanization externalities have a strong positive effect on individual's choice to become self-employed, the results for Marshallian versus Jacobian externalities are less clear cut. The combined evidence support that lack of local competition will increase the probability of transitions to entrepreneurship. Furthermore, smaller firms are more likely to spawn entrepreneurs and while it is in generally less productive firms that individuals will leave to become self-employed, in the case of high-tech industries more productive and more innovative firms are more likely to spawn entrepreneurs.

The rest of the paper is organized in five more sections. Section 2 outlines the theoretical background upon which the analysis that follows builds. Section 3 describes the data used in the study and section 4 describes the empirical framework. Section 5 displays and discusses the results of the estimations while section 6 summarizes and concludes.

2. Theoritical discussion

Entrepreneurship research suffers from a high degree of fragmentation and a lack of a coherent theoretical foundation. Gartner (2001) describes the field of entrepreneurship as a 'cacophony of results and ideas', and Shane and Venkataraman (2000) as a 'hodgepodge'. Low (2001) expresses a sincere concern for the legitimacy of entrepreneurship as an academic discipline. Low and MacMillan (1988) in an important contribution to the field take stock of past shortcomings of entrepreneurial research and call for approaches with clearer purpose and focus. They also encourage scholars to combine multiple levels of analysis in their research recognizing entrepreneurship as the outcome of actions of individuals that act in and are influenced by the organizational and regional environment in which they live and work (Gartner, 1985; Aldrich and Zimmer, 1986).

More than a decade after its publication, Davidsson et al. (2001) assess the extend to which Low and MacMillan's (1988) call for micro/aggregate mix approaches has been heeded and report rather disappointing results, recognizing however the difficulties behind gathering appropriate data for such an analysis. It is this gap that the current paper tries to breach.

In this section a short discourse on defining entrepreneurship is followed by a presentation of the theories concerning primarily the regional and secondly the firm characteristics affecting entrepreneurial spawning.

2.1. The entrepreneur

Much of the early literature focused on the personality of the entrepreneur and sought to portray his figure. Gartner (1988) carries out a very thorough inspection of the literature implementing this approach and points out the lack of coordination and comparability that led to a list of traits that you couldn't possibly hope to find in the same person. According to Gartner the entrepreneur is not a fixed state of existence, rather entrepreneurship is a role that individuals undertake to create organizations. Besides, it is part of the nature of the entrepreneur to be largely irrational, operating on hunches and intuitions. Entrepreneurs are usually the exceptions to the rule and tend to reside at the tails of population distributions in a

heterogeneous enough manner to make any effort of profiling the typical entrepreneur almost impossible (Low and MacMillan, 1988).

Even if we assume that such a thing as entrepreneurial skill exists and is present in just a few gifted individuals it is impossible to observe it before (and unless) it is actually realized. The behavioural approach has successfully identified that the entrepreneur needs to be defined based on his actions rather than who he is. See Hébert and Link (1989) and Davidsson (2004) for an extensive list of definitions applied in the relevant literature. There are however, as Baumol (1993) stresses out, two uses of the term entrepreneur which, though both legitimate are entirely different in their substance. The first is the innovator-entrepreneur as described by Schumpeter (1934) whose function is to introduce novel products and processes in the economy. The second one is Kirzner's (1973, 1985) entrepreneur who operates as an alert discoverer of market opportunities whose efforts lead to the creation of new enterprises without the introduction of innovations.

Entrepreneurship owes a large part of its newfound popularity on the evidence that new firm creation is a critical driving force of economic growth and job creation (Birch, 1979; Birley, 1987; Reynolds, 1987) but how many of these firms are actually innovative and to what extend? To date almost no research exist that compares the significance for economic growth of radically innovative new firms to that of firms that came into existence without introducing any novel product or process to the market, largely due to the lack of appropriately detailed data. Aldrich and Martinez (2001) support the distinction between "innovation" and "reproduction" in entrepreneurial activities and argue that innovation and entrepreneurship are not necessarily coupled calling into attention the major role of imitation in entrepreneurial processes as is evidenced by the numerically dominant role of reproducers, rather than innovators

An interesting new approach that manages to include both types of entrepreneurial activities while still focusing on the function of the entrepreneur is the opportunities approach. In Venkataraman (1997) and Shane and Venkataraman (2000), entrepreneurship is the process of discovering, evaluating and exploiting opportunities to create future goods and services. Within the products market Drucker (1985) describes three different categories of opportunities: 1. the creation of new information, as occurs with the invention of new technologies, 2. the exploitation of market inefficiencies that result from information

asymmetry, as occurs across time and geography, and 3. the reaction to shifts in the relative costs and benefits of alternative uses for resources, as occurs with political, regulatory, or demographic changes. Clearly this definition encompasses both Schumpeter's and Kirzner's views of entrepreneurship while bringing into focus the issue of the supply of opportunities as the first link in the chain of occurrence-identification-exploitation of entrepreneurial opportunities that lead to the creation of new firms.

It is therefore important when using the creation of new firms, or as is the case in the present study, the decision to start a new firm, as a measure of entrepreneurship to interpret any results keeping in mind that this can be the outcome of two distinctly different entrepreneurial behaviours.

2.2. The firm and the region as breeding ground for new entrepreneurs

2.2.1. The region

There are several different location attributes identified in relevant literature as having an effect on start-up rates. The most important ones are the region's aggregate knowledge resources, the degree of industrial specialization-diversity, and the size of the market.

Knowledge

Much of the literature focusing on the personality of the entrepreneur assumed opportunities to be exogenously given. The shift in attention towards the recognition and exploitation of opportunities (Venkataraman, 1997) was followed by an interest in identifying the source of opportunities that were previously assumed to be exogenously given. In the knowledge production function model proposed by Griliches (1979) innovation is the outcome that follows from a firm's investments in knowledge inputs. According to Warsh (2006) the technology opportunity set is endogenously created by investments in new knowledge. Acs and Audretsch (1988, 1990) observe however that small and medium sized firms exhibit an innovative output disproportional large compared to their investment in R&D. This discrepancy can be largely explained by the differences between knowledge and the other factors of production. Knowledge is nonexcludable, and nonrivalrous (Arrow, 1962).

Audretsch (1995) shows that the gains from investments in knowledge are not contained within the investing firm but new knowledge and as a result new opportunities spillover to other firms as well. The major contribution of the knowledge spillover theory of entrepreneurship was to make the supply of opportunities endogenous and point to the accumulation of new knowledge and ideas as the source of opportunities (Audetsch et al., 2006). According to the knowledge spillover theory of entrepreneurship a major source of entrepreneurial opportunities is new knowledge not fully appropriated by the source that created it, such as firms' R&D laboratories or universities.

Non-excludability however does not make new knowledge a pure public good. Arrow (1962) also describes the difficulties behind codifying and transmitting knowledge. Polanyi (1967) coins the term tacit knowledge to describe highly specific knowledge that demands high transaction costs. Jaffe (1989), Audretsch and Feldman (1996), and Audretsch and Stephan (1996) provide evidence that knowledge spillovers are geographically bounded and localized within spatial proximity to the knowledge source pointing among other things to the importance of knowledge embedded in regionally immobile agents and face-to-face contacts. Audretsch and Lehmann (2005) show that the knowledge spillover theory of entrepreneurship holds for regions as well as industries.

Market size

The size of the regional market has a direct effect on both the supply of opportunities and the feasibility of exploiting them. First of all a large market implies the existence of a large pool of potential clients. On the one hand this allows spreading out sunk costs over larger scales making the exploitation of particular opportunities economically feasible (Andersson and Hellerstedt, 2009) and on the other one increases the probability of identifying unexplored market niches that target a particular clientele. Depending on the nature of the business opportunity pursued, gains from scale economies in production, low-cost accessibility to various inputs, and lower transportation costs may also accrue. Moreover, larger markets protect new ventures from hold-up costs of customer-specific customizations (McLaren, 2000).

In the case of Sweden the size of a region is almost perfectly correlated with the population density of the region. Urbanization/agglomeration is one of the only two processes Reynolds

et al. (1994) find to have a consistently positive effect on regional start-up rates (the presence of small firms and economic specialization being the other). High density markets facilitate immensely businesses that depend on common or even constant face-to-face interaction (Coffey and Bailly, 1991). Furthermore, a concentration of people and market activities is coupled by a concentration of research activities and more frequent inter-firm labor exchange rates that could, in theory, enhance the intensity of knowledge spillovers.

Specialization versus diversity

Whether specialization or diversity of regional economic activities best promotes innovative performance and regional growth is the topic of an ongoing debate. The dilemma is often expressed as a need to identify whether Marshallian specialization or Jacobian diversification externalities favor regional innovativeness and growth (van der Panne and van Beers, 2006).

According to Marshall (1890) firms operating in a specialized economy can benefit from a large industry-specific labor pool, asset-sharing, and specialized suppliers of intermediate input factors of production. Henderson et al. (1995) refer to these externalities as localization externalities. Moreover, Marshall (1890), Arrow (1962) and Romer (1986) suggest that knowledge spillovers are industry-specific and they may only be appropriated in regions of high industry-specific concentrations (hereafter referred to as MAR externalities). Jacobs (1969) promoted his opposing view advocating that it is the knowledge exchange among a diverse spectrum of industries that facilitates experimentation and innovation by implementing established methods in novel ways. A third type of externality attributed to Hoover (1937, 1948) is urbanization externalities. Similarly to localization externalities urbanization externalities are caused by the agglomeration of economic activity but they are not industry specific. All industries situated in a densely populated region are argued to benefit from an extensive local labor market, strong local demand, and well developed infrastructure irrespective of the focus of their business.

Most studies to date have focused on either overall employment growth (Glaeser et al., 1992; Henderson et al., 1995; Dumais et al., 2002) or innovation output (Feldman and Audretsch, 1992; van der Panne and van Beers, 2006) overlooking the role of spatial externalities in entrepreneurial activity. Feldman and Audretsch (1999) report evidence to support the diversity over the specialization thesis while van der Panne and van Beers (2006) do the exact

opposite. Desrochers and Sautet (2008) review further literature on the matter and criticize it on the inadequacy of the measures of specialization/diversity used as well as the measures of innovative "output" implemented such as new product advertisements in technical literature, patent data, and questionnaires. They proceed by theorizing that purposefully promoting specialization can prove counterproductive and regions should be allowed to develop spontaneous diversity. Paci and Usai (1999) however argue that Marshall and Jacobs externalities are not necessarily opposed, since specialization is a particular feature of a certain sector within a local system whilst diversity is a characteristic of the whole area. They use Italian patent data to support their claim.

Closer to the present study, van Oort and Atzema (2004) use data on new firm formation in the Dutch ICT sector to conclude that new establishments in the ICT sector tend to be concentrated in urban areas that are already relatively specialized in this sector and that are relatively rich in the presence of other industries.

2.2.2. The firm

Hyyitinen and Maliranta (2008) discuss the literature on the effects of a firm's size and innovativeness on its probability to spawn entrepreneurs. Two contrasting views concerning the effect of the size of the firm are identified. On the one hand, according to Gompers et al. (2005), individuals working in small firms have more opportunities to develop entrepreneurial skills and build connections to networks of suppliers and customers by being involved in multiple processes within the firm. On the other large firms are more prone to spawn entrepreneurs because they are incapable to identify or unwilling to diversify on opportunities too far away from their main line of business or that constitute too radical innovations. A firm's innovativeness is also theorized to play an ambiguous role since on the one hand individuals employed in R&D intensive firms stand better chances of identifying new technologies they may commercialize independently but on the other hand such firms are more likely to take legal actions in order to prevent leakages of this sort (Kim and Marschke, 2005).

3. Data

The paper utilizes an unbalanced panel data compiled by Statistics Sweden and referred to as FAD which is the acronym of "Firms and establishments dynamics" (Företagens och arebtsställenas dynamik). FAD contains linked information on all firms, establishments and working individuals in Sweden between the years 1985 and 2005. Statistics Sweden uses several sources to build this massive database but the data comes mainly from the Swedish tax office and firms' annual reports. The data set contains information on individuals' education, occupation and places of origin and residence and organizational and financial data for firms and establishments.

One of the biggest strengths of FAD is that it covers all individuals and firms in the Swedish economy. To my knowledge it is only other Nordic countries that collect such information. However, what makes the Swedish data special is the precise linkage among individuals and their working place allowing for incredibly detailed labour mobility analysis. Isolating the information on a single individual and comparing the information available for him at year t and t+1 I can identify whether he has remained at the same job position, switched into a different role or to a new or already existing establishment within the same firm, switched to an altogether different firm, became self-employed or stopped working.

In order to guarantee the highest possible quality of the data the study is limited in the period 1999 - 2005. Furthermore the focus of the research is on individuals that were employed at a firm at time t that still existed at time t+1 and the change (or no change) in their occupation between time t and t+1. That excludes from the part of the population the paper takes into consideration people that were owners of their own firm at time t as well as people that jumped from unemployment at time t into being owners of their own firm at time t+1 and people that were forced into an occupational change because the place of their former employment closed down. These three groups are of considerable interest on their own right but of no consequence for the present study.

One drawback of FAD that needs to be mentioned is imperfect reporting of some variables. While data whose original source is the tax authorities of Sweden is perfectly reported and there are practically no missing values, data derived from firms' annual reports (firms' gross investment, net turnover, value added and others) suffer from imperfect reporting. Approximately 15% of the individuals in the population of interest are employed in firms that

failed to report these variables in one or more time periods. One might fear that there is a trend behind the missing values and excluding the firms with missing data from the analysis would introduce a nontrivial bias. However this is not the case here. Comparing the distributions per size and industry category of firms that reported perfectly their financial data to those that did not they are found to be almost identical making it clear that the missing values are purely random and can be excluded from the analysis without affecting the results.

What is of primary interest is the effect of the characteristics of the firm and the region on an individual's propensity to become self-employed, controlling for several key characteristics of the individual. Similarly to Hyytinen and Maliranta (2008), the advantages of the method used is that it allows the identification of the firms the entrepreneurs were employed in before starting their own firm and the comparison of the entrepreneurs to the individuals in similar environments that chose not to follow the path of self-employment.

The value added of this paper is that unlike in Hyytinen and Maliranta (2008) individuals can be linked to the location of their residence as well as the location of their place of work and therefore one may look past the effects of the organizational context on an individual's occupational choice, to include those of the regional context as well. Past studies have only studied the effects of regional characteristics on aggregate start-up rates (Davidsson et al., 1994; Reynolds et al., 1994). To my knowledge no other research has considered the effect of the regional dimension on the propensity of the individual to transcend into self-employment. Also, the method of identifying the transition to self-employment used in this paper is a lot more exact than Hyytinen and Maliranta's (2008). They characterize as entrepreneurs people that are insured on the basis of the Self-employed Persons' Pension Act (YEL), which includes all people who are partners in a general partnership, assume the role of the general partner in a limited partnership or own 50% or more of the stocks in a limited liability company. I instead am able to identify the moment someone leaves his previous employment to become the owner of his own firm.

At each year FAD contains information on the entire working population and all existing firms and establishments organized in three different matrixes, one for each different level of observation. A firm is a distinct legal entity that could include one or several establishments under its umbrella. Each observation is marked with a unique identification number which makes linking the different observations in the three matrices possible. The observation on

each individual carries his unique identification number as well as that of the establishment and the firm in which he is working at the time period in question. The observation on each establishment carries its establishment identification code as well as that of the firm to which it belongs. Finally the firm carries its own identification code alone since there is no higher authority to which it needs to be linked. Therefore, using the individual as the point of departure one can add information about his working environment by linking his observation to those of the establishment as well as the firm in which he is working in a bottom to top linking process. Of great value for this research is the fact that for each establishment and firm there is information on whether the unit in question is new or pre-existing and whether new units were the result of a merge or a split of other units.

Comparing an individual's current employment status with that of next year it is possible to identify whether she has stayed in the same job, switched to a different job, stopped working or became the owner of her own firm. Exiting the workforce is pretty straightforward; individuals not employed at a certain time period are not included in FAD so people exiting from the panel from one time period to the next are classified as choosing to Exit.

Staying in the same position is surprisingly complicated to identify considering it is the most common occupational choice and should be relatively simple to recognize. However, splits, merges, buy-outs, and internal reorganizing of firms introduce considerable turbulence in the identification codes of firms and establishments even if such changes have no to little effect on the employment of most individuals in the firms involved in these processes. I do not consider an individual to have left his previous employment unless both the establishment and the firm in which he was previously employed have changed taking also in consideration the location of his working place.

Cases in which the establishment an individual was working at was bought or transferred to a different firm I consider as Stay since not much has changed from the point of view of the individual and any change that might have come about was enforced upon him rather than being the result of an individual choice. Switching is simply the act of going from being employed in a certain establishment in a certain firm to being employed in an entirely different one. The choice to become Self-employed is the most interesting one and the focus of this paper. Thanks to the unique detail of the data set used in this paper the choice of

becoming Self-employed is as close to an appropriate measure of entrepreneurial activity as is possible using registry data.

Measuring entrepreneurship has proven to be no easy task and a variety of proxies have been used in the past by many studies. Audretsch (2002) identifies some of the most commonly used proxy measures to be self-employment rates, business-ownership rates, new-firm births or other measures of industry demographics. In this paper I manage to incorporate to a certain degree all of these different measures which to my knowledge constitutes a unique contribution. I presently describe as choosing to become Self-employed individuals that go from being employees in a firm at time t to becoming owners of their own, newly found firm at time t+1. Several drawbacks inherent in the proxies above are overcome this way. First of all using rates of self-employment or business-ownership irrespective of timing grossly overestimates rates of entrepreneurship. For how long after having established a corporation can the founders still be considered to be of an entrepreneurial nature? At which point of its cycle does a firm mature enough to be considered an incumbent rather than a starting firm? Second of all linking start-ups to the founders of these firms and comparing them to their peers that chose a path different than the one of the entrepreneur sets this study apart from all the rest that only consider the success stories of accomplished entrepreneurs.

I do not include in the category of "choosing to become self-employed" people becoming the owners of firms that were the result of splits from incumbent firms. A split is characterized by a group of employees leaving their former employment to form together a new firm. This phenomenon is rather rare and its significance rather doubtful. Firms will often choose to separate part of their operations from their core business creating new legal entities as part of their corporate strategy. Another explanation would be that a branch of a corporation chooses (and also manages, which is not always an easy feat) to severe the ties to the mother company and to pursue to offer its specific services to multiple clients. Either way, a split is the result of corporate strategy or group dynamics and may not be considered in parallel with individual choices.

Summary of Employment Choices

In Table 1 the employment choices for next year of all working individuals at each time period are portrayed. When it comes to the issue of timing I assume that at the end of each year each individual decides his course of action for the following year and the choice made is identified by comparing his employment status at year t and t+1. Therefore, the column labelled "1999" in Table 1 describes the employment status in 2000 of the 1720364 people that were employed in a firm in 1999; 79.3% of those remained in the same job position, 14.06% switched to being employees in a different firm, 6.4% stopped working, and 0.3% transcended into entrepreneurship. That is why although the available data covers the year 2005 as well the analysis in this paper only covers the choices that happened up till year 2004.

Considering the percentages in Table 1 no strong temporal trends seem to emerge. The transition to self-employment exhibits a clear increase in the last two years considered although it remains a relatively rare occurrence at an average of 0.4%. The percentage of people choosing to switch to different forms of employment exhibits an almost steady decline until 2003 only to slightly rise again in 2004. The ratio of people deciding to Stay is slightly higher after 2001 than before while the ratio of those Exiting seems to fluctuate around 7%.

Table 2 summarizes the choices of the pooled sample per industry category and size of the originating firm. Although the actual values of the percentages change between industry categories some very interesting trends emerge. The percentage of employees choosing to transcend to entrepreneurship consistently declines as the size of the firm increases, in all sectors. So does the percentage of people Exiting although the decline is not that pronounced in the third and fourth category. The propensity to stay in the same firm increases parallel to the size of the firm in all cases but that of the knowledge intensive firms. The propensity to switch seems to exhibit the grater variance among the different industry sectors, dropping in some cases while the size of the firm increases and following the exact opposite trend in some others. Finally it is worth noting that although large firms of more than 501 employees seem to disturb most of the other trends, whether those are upward or downward ones, they fail to do so in the case of the choice to become self-employed. Entrepreneurs seem to be a lot more likely to come from small rather than large firms in all four industry sectors considered.

4. Empirical design

4.1. Setup

Following a setup similar to Hyytinen and Maliranta (2008), at the end of each time period each employee faces the choice of what form of employment she would like to have in the next time period. She can decide to Stay in her current employment, Switch to a different employer, Exit from the workforce, or become Self-employed. These four options make up her choice set C_n. The probability that any element in C_n is chosen is given by:

$$P_n(i) = \Pr(U_{in} \geq U_{jn}, \forall j \in C_n, i \neq j) \qquad (1)$$

Or in words, an individual will choose alternative *i* over the other elements of C_n as long as the utility resulting from that choice exceeds the utility that would have resulted by choosing any of the other alternatives. The utility, *U*, has a deterministic (*V*) and a random component (*e*):

$$P_n(i) = \Pr(V_{in} + \varepsilon_{in} \geq V_{jn} + \varepsilon_{jn}, \forall j \in C_n, i \neq j) \qquad (2)$$

If $U_{in} = V_{in} + \varepsilon_{in}$ for all $i \in C_n$ and the disturbances are (1) independently distributed, (2) identically distributed and (3) Gumbel distributed with a scale parameter $\mu > 0$, then the choice probabilities are given by:

$$P_n(i) = \frac{e^{\mu V_{in}}}{\sum_{j \in C_n} e^{\mu V_{in}}} \qquad (3)$$

The stochastic component in *U* exists due to lack of perfect information available to the modeler. Manski (1977) identifies four different sources of uncertainty when trying to model individuals' utility; non-observable characteristics, non-observable variations in individual utilities, measurement errors, and functional misspecifications. In such a setup where the outcome is nominal and the categories are assumed to be unordered the appropriate model is a multinomial logit (MNL) model that allows the estimation of the probability of choosing self-

employment conditional on the other available choices in each period (Long and Freese, 2006; McFadden, 1974). Taking the decision of the individual to Stay in her current job as the base outcome the MNL model to be estimated is:

$$P(y = j \mid y = j \quad or \quad y = Stay, \mathbf{x}) = \Lambda\left[\mathbf{x}\left(\boldsymbol{\beta}_j - \boldsymbol{\beta}_{Stay}\right) + \varepsilon \right] \qquad (4)$$

Where y = the employment choice of the individual for the next time period

$j \in$ {Switch, Exit, Self-employed}

$\mathbf{x}$ = vector of individual, firm and regional controls

$\boldsymbol{\beta}$ = vector of coefficients

$\Lambda(c) = \exp(c) / \left[1 + \exp(c)\right]$

ε = i.i.d. error term

Although access to panel data allows the use of dynamic modelling approaches that incorporate the time dimension explicitly, Rhody (1998) and Dunn and Holtz-Eakin (2000) suggest limited gains from such methods. I therefore pool the data from all the years and apply a straightforward MNL estimation with the inclusion of time dummies and use corrected standard errors to account for the fact that there are multiple observations for most of the individuals in the data.

A basic assumption of the MNL estimation that needs to be considered is the assumption of independence from irrelevant alternatives (IIA). Applied in this case the IIA assumption requires that when choosing for example between Staying and Exiting, switching to a different employment or becoming an entrepreneur are not issues that will affect that choice. The problem with the IIA assumption is that it cannot be tested with absolute certainty but needs to be considered carefully in each case (Long and Freese, 2006).

Given the nature of the issue addressed, I believe it is reasonable to assume that the IIA assumption holds in this case. It is hard to imagine that people deciding to stop working or retire are considering entrepreneurship or a new job as alternatives affecting their choice (at least in the immediate future). The irrelevance between starting an own firm and switching to a different form of employment might not be so straightforward since both are ways of remaining in the working force while still exiting from the previous employment and the two

"search processes" could in theory be carried out simultaneously by the same individuals. Hyytinen and Ilmakunnas (2007) compared the two search processes and found that most employees are not engaged in both at the same time, nor are the two processes alike. Following Hyytinen and Ilmakunas (2007), we therefore assume that the IIA assumption holds and the MNL estimation is the most appropriate estimator in the present case.

4.2. Control Variables.

Regional controls

The measure used to control for the size of the region is, rather straightforwardly, the log of the number of active workers in the region. The coefficient on this particular conditioning variable however contains a lot more information than just the effect of the size of the market. Most importantly, the size of the region is assumed to capture the effect that the regional knowledge sources play on the propensity of the individual to transcend into self-employment. There are two different measures of regional knowledge resources included in the present research: 1. the region's knowledge intensity, measured as the ratio of employees with a tertiary education over the total number employees in the region, and 2. private firms' and universities' R&D investments. 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 therefore the simultaneous inclusion of any combination of these measures would cause all the coefficients in the specification to be biased. Moreover the size of the region correlates almost perfectly with the population density which has been used in relevant literature to capture the effect of urbanization externalities (Burger et al., 2008).

A series of different indices have been developed and implemented in relevant literature for capturing the effects of industrial specialization and diversity (Feldman and Audretsch, 1999; Paci and Usai, 1999; Duranton and Purga, 1999; van der Panne and van Beers, 2006). In order to facilitate the interpretation of the results and given the number of control variables to be included in the final specification I decided to implement the most straightforward of those.

I include the production specialization index (PS) for each industry branch in each region to capture the degree of relative specialization of these industries. The PS-index measures the extent to which region *j* is specialized towards industry *i* :

$$PS_{ij} = \left[\frac{E_{ij}}{\sum_i E_{ij}} \right] / \left[\frac{\sum_j E_{ij}}{\sum_i \sum_j E_{ij}} \right] \qquad (5)$$

where $i = 1, \ldots, 43$ for each industry branch

$j = 1, \ldots, 81$ for each functional region

E = employment

As for capturing the extent of the Jacobian externalities in each region I implement the inverse of a Hirshman-Herfindhal index which sums over all sectors the square of each sector's share in local employment. Formally,

$$DI_j = 1 / \sum_i \left[\frac{E_{ij}}{\sum_i E_{ij}} \right]^2 \qquad (6)$$

A value equal to 1 for the DI means the economic activity in the particular region is fully concentrated in one sector, and the index increases as the region becomes more diverse.

Firm controls

Following Hyytinen and Maliranta (2008) I mainly wish to control for a firm's size and innovativeness. The size of the firm is almost always defined as the number of employees working in a firm while sales or profits might be used as crude proxies when the number of employees is not known. Controlling for a firm's innovativeness is considerably less straightforward and it's the one point where researchers tend to diverge based on the availability of the data. Ideally one would like to control for both innovative input and output.

The input side would include a description of a firm's investments in R&D as well as the number of researchers dedicated in the development of new products and processes. The

output side would include the outcome of the input efforts, in other words the introduction of new products and/or processes in the market. Data of such fine detail, and in particular a good measure of innovative output, is unfortunately extremely hard to obtain. The most commonly used measure of innovative output is the number of patents issued from a firm, an approach that despite some drawbacks (Pavitt, 1982) has proven to be an accurate way of measuring innovative output (Griliches, 1990). Unfortunately, such data is not available in the data set implemented.

In the present research the size of the firm is measured with the usual convention of number of employees by creating 6 size-category dummy variables. In order to control for the innovativeness of the firm I use the knowledge intensity of the firm as a proxy, defined as the ratio of employees with a tertiary education over total firm employment.

Moreover, I control for the age of the firm, the log of productivity, where productivity is defined as value added per employee, declining employment and declining sales dummy and the industry category. The Declining Employment dummy equals one if there has been a drop in a firm's employment between time t and t-2 and the Declining Sales dummy equals one if there has been a drop in the sales of the firm between t and t-2. These two are rather important controls that help capture any push-out effects caused by a decline in the business of the incumbent firm. Especially the choice to Exit can very often be involuntary and this is a very convenient way to control for whether the firm is downsizing or not. The four industry categories into which the firms are broken down, based on the two-digit SNI codes[1], are manufacturing (SNI codes 15 to 37), low end services (SNI codes 38 to 64), financial and real estate services (SNI codes 65 to 71), and knowledge intensive business services (SNI codes 72 to 74).

Individual controls

I control for the individual characteristics most commonly referred to in the literature of entrepreneurship and labour mobility as important determinants of employment choices.

[1] The SNI – Swedish Standard Industrial Classification codes used by Statistics Sweden correspond almost perfectly to the **NACE** -- Classification of Economic Activities in the European Community codes.

These are tenure[2], age, age squared, male (equal one if male, zero otherwise) and a set of seven education dummies corresponding to the highest educational level attained as categorized by Statistics Sweden. The seven levels are: 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 max 3 years. 5. post secondary education, shorter than 3 year, 6. post secondary education, 3 years or longer (excluding PhD), 7. PhD. FAD contains no family background information but these have been found to have no significant effect on entrepreneurial intentions (Kolvereid, 1996).

Table 3 provides a summary of the variables included in this paper. Note that the summary statistics were taken after pooling the data in the period 1999-2004. Some observations concerning the data are that approximately 70% of the individuals included in the study are males; about a third of the individuals were employed in a firm exhibiting declining sales/employment; the percentage of people with a tertiary education is roughly 1%; around a third of the population works in firms of fewer than 50 employees, a third in firms employing between 51 and 500 individuals and the last third in firms of more than 501 employees; the majority of the individuals (83.5%) work in the two first industry categories.

5. Results

The focus of this paper is to determine how the characteristics of the region and the firm in which an individual is employed affect his or her choice to become an entrepreneur. To this end a MLN estimation is run on a rich matched employee-firm dataset describing the Swedish economy in the period 1999-2005. The dependent variable is the choice of one of the following alternatives: stay in the same job, switch to a new job, stop working, or transcend to self-employment.

Carrying out the MNL estimation requires that one of the alternatives is used as the base outcome for normalization (here the base outcome was taken to be Stay) and the results are

[2] Since the FAD data dates back to 1985 the Tenure variable has a maximum value of 18. It is possible that some individuals have been in the same job for a lot more than that but the significance of the variable is so high that undermining its importance because of misspecification is not an issue.

the effects of a change of the independent variables on the probability that the individual will choose each one of the other three alternatives over the base outcome.

In order to ease the interpretation of the results I report the marginal effects of the estimation. A small increase in one of the continuous control variables (or a discrete change from 0 to 1 of one of the dummy variables) will increase the probability of an individual choosing one or more of the options while decreasing the probability of choosing the rest. All in all, choosing to become self-employed remains a rare occurrence and being able to explain even the slightest variations in the individual's propensity to become self-employed is important. Table 4 reports these marginal effects along with their standard errors.

Note that in each case a set of dummy variables were used to divide the population into subgroups the first category in each division was selected as the reference group. These are Education level 1, firms of size of 1 to 10 employees, industry category 1 (manufacturing) and year 1999.

Although the focus of this paper is the effect of the control variables on the propensity of individuals to transcend into self-employment and therefore the marginal effects reported in the fourth and last column, one can derive a host of interesting observations from the rest of the analysis. I will refer to the most interesting ones in parallel with the discussion on the results reported in the fourth column.

The results of the MNL estimation on the effect of regional size on the propensity of individuals to transcend into entrepreneurship agree with those from studies looking into aggregate firm birth rates. The effect is positive and significant. Coupled with the positive influence on Switching and the negative effect on Staying and Exiting the results suggest a much more mobile work force in large and dense markets. Using Hoover's (1937, 1948) terminology, urbanization externalities are found to have a strong positive effect on individuals' choice to become self-employed.

Turning to the controls used to capture the effects of MAR and Jacobian externalities the results are at first glance surprising with the coefficients being negative and significant for both. An interesting picture emerges from these combined results. It appears that *ceteris paribus* individuals are more likely to select self-employment when facing a weak local

competition since they are more likely to transcend into entrepreneurship when living in a region that is concentrated in few industries while at the same time they are not likely to originate from an industry the region is specialized in. This would suggest that entrepreneurship adds to the development of regional industrial diversity by exploiting market gaps resulting from industrial concentration. Any effort to compare these results to other studies that focus on the effects of externalities must take into consideration the different methodologies applied and most importantly the fact that the most commonly used dependent variable is some measure of innovative output (patent output or new product announcements in technical magazines). Great care was taken in this paper into underlining the difference between entrepreneurial activity as measured by new firm creation as opposed to innovation. These results however are in line with van der Panne (2004) and van der Panne and van Beers (2006) who find that competition among firms negatively affects regional innovativeness, in support of Marshallian externalities. They also provide further evidence to the results of Dumais et al. (2002) according to which the location choices of new firms play a de-agglomerating role since they are more likely to start away from current geographic centers of that particular industry. Further and more careful research in this issue is warranted. Some interesting questions that arise from these results are the extent to which individuals leaving their jobs to start new firms remain in the same industry branch, as well as the significance of regional externalities for potential switches. Also, does regional entrepreneurial activity translate into future industrial diversification?

When considering the firm-level control variables, the results only partly agree with those of Hyytinen and Maliranta (2008) but are closer to what might be intuitively expected. Similarly to Hyytinen and Maliranta (2008) smaller firms are found to be more likely to spawn entrepreneurs since the marginal effect on Self-employed diminishes significantly as the size of the firm increases. As for the other choices there seems to be an increase in the probability to Exit as the firm gets larger but there is no clear trend as far as Stay and Switch is concerned. However, in the case of the Finnish data less productive and less innovative firms seem to be consistently more likely to spawn entrepreneurs, which is not the case in Sweden. When considering the economy as a whole and according to Table 4, although productivity has a negative effect on the choice to become self-employed, a firm's innovativeness, measured by its knowledge intensity, has no significant effect. People working in firms with high knowledge intensity are however more likely to Stay rather than Switch or Exit. One of several robustness checks is to focus on individuals employed in high-tech sectors (SNI codes

29-33 and 72-74) where productivity and especially the innovativeness of the firm can be argued to be more important in the occurrence and identification of entrepreneurial opportunities. In contrast to Hyytinen and Maliranta (2008) but in line with Gompers et al. (2005) more productive high-tech firms that exhibit a higher degree of knowledge intensity are more likely to spawn entrepreneurs (see Table A2 in the Appendix).

Everything else considered the age of the originating firm seems to play a positive and significant but extremely weak effect on the propensity to become Self-employed. The effect of the age of the firm is considerably stronger in the other three choices with older firms being more likely to hold their employees rather than let them switch to different ones or stop working. Declining employment and sales obviously have a strong, negative effect on Staying. This push-out effect is mainly translated into people Switching to a different job or Exiting. The gains of entrepreneurship are certainly weaker than those of between-firms mobility but significantly positive and certainly not to be overlooked.

As for the marginal effects of changes in the control variables of the individual characteristics the results agree with the existing literature. Better educated males with shorter tenure exhibit the highest propensity to become entrepreneurs. Long tenure has a very strong and significant positive effect on the probability of staying in one's current employment. This was of course expected since long tenure usually means high wages and better positions in a firm. Age has a non-linear impact on all four choices however it is interesting to notice that the effect on Self-employed is significantly weaker than the effect on the other three choices. Furthermore, men are a lot more likely to stay in the same job position while women will decide to stop working more often than men. It is also very interesting to note how the chance of Staying in the same position drops as the level of education increases in favor mostly of Switching and to a much lesser degree of becoming Self-employed. This increased between-firms mobility of highly educated individuals is noteworthy given that the knowledge-spillovers theory of innovation hypothesizes that labor mobility is one of the major sources of tacit knowledge transferences.

Robustness checks[3]

Thanks to the richness of the data set I was able to carry out several important robustness checks. The MNL estimation was carried out for each year independently as well as for 20% and 33% samples of the population without significant differences. I also used a slightly different specification of the model where I included only two instead of seven education dummies splitting the population in those with a tertiary education and those without, and I used the log of the number of employees in a firm as a firm size control instead of the six size category dummies. The knowledge intensity of the firm was excluded as a control in this alternative specification. Table A1 in the Appendix displays the results of this alternative specification that agree with those of the original one. A specification that did not include the regional control variables was also found consistent with the results on individual and firm level controls presented in this paper. Finally, I run the MNL estimation on the individuals that were employed in high-tech sectors alone and this was the only robustness check that led to the differences discussed above. Keeping in mind that the founding of a new firm can be the result of two distinctively different processes (the exploitation of either an innovation or of a market disequilibrium) these differences are by no means a weakness of the present paper but rather add merit to the validity of the results.

6. Conclusions

This paper applies a multi-level approach in order to identify the characteristics of the regions and the firms that are more likely to breed entrepreneurs using a dataset describing the Swedish economy in the period 1999-2005. The unit of analysis is the decision of the individual to transcend into self-employment in contrast to other employment alternatives, controlling for characteristics of the individual, the firm she is currently employed in, and the region she is currently living and working in.

Transitions to entrepreneurship are found to be rare with individuals choosing that path at a probability of just 0.44%. In summary, the main findings of the empirical analysis are that individuals working in larger regions that are characterized by larger local markets, higher

[3] Only the results of two of the robustness checks that were carried out can be found in the Appendix. The rest were excluded to save space but are available upon request from the author.

accumulation of knowledge resources and higher population density are more likely to transcend into entrepreneurship. I also find evidence that people are more likely to select the path of self-employment in the face of weak local competition as expressed by a negative effect of both industrial specialization and regional diversity.

Moreover, my results only partly agree with a similar research by Hyytinen and Maliranta (2008) that focuses on just the characteristics of the individual and the firm in the case of Finland. In the case of both Sweden and Finland the size of the incumbent firm is found to have an inverse effect on the probability that its employees will transcend into entrepreneurship. However while the Finnish study reports that a firm's innovativeness and productivity are also inversely related to the probability that it will spawn entrepreneurs, that is not precisely the case of Sweden. When considering the whole economy I find that the productivity of the firm is indeed negatively related with the probability of self-employment but a firm's innovativeness plays no significant role. When focusing on high-tech sectors the results are reversed and I find a positive effect for both the productivity and innovativeness of the firm.

Note that this analysis concerns the choice of the individual and one needs to take into consideration the demographic characteristics of a region in order to translate these results into entrepreneurial output. The present analysis may be extended in several ways by matching the emerging entrepreneurs to their start-up firms and examining their survival and growth rates. Looking into the extent to which an individual starts a new firm in the same industry he was formerly employed in can also shed further light on the importance of Marshallian over Jacobian externalities.

References

Acs, Z. J., and Audretsch, D. B. (1988) Innovation in small and large firms: An empirical analysis. *American Economic Review*, 46, 5-23.

Acs, Z. J., and Audretsch, D. B. (1990) *Innovation and Small Firms.* Cambridge: MIT Press.

Acs, Z. J., Audretsch, D. B., Braunerhjelm, P., and Carlsson, B. (2006) The knowledge spillover theory of entrepreneurship, *CESIS Electronic Working Paper Series.*

Aldrich, H. E., and Martinez, M. A. (2001) Many are called, but few are chosen: An evolutionary perspective for the study of entrepreneurship, *Entrepreneurship: Theory and Practice*, vol.25, issue 4, p41-57.

Aldrich, H. E., and Zimmer, C. (1986) Entrepreneurship through social networks. In D. L. Sexton and R. W. Smilor (Eds.), *The art and science of entrepreneurship* (pp. 2-23). Cambridge, MA: Ballinger Publishing.

Andersson, M., and Hellerstedt, K. (2009) Location attributes and start-ups in knowledge-intensive business services. *Industry and Innovation*, Vol. 16, No. 1, 103-121.

Arrow, K. (1962) Economic welfare and the allocation of resources for invention. In *The Rate and Direction of Inventive Activity,* pp. 609-626. Princeton: Princeton University Press.

Audretsch, D. B. (1995) *Innovation and Industry Evolution*, Cambridge: MIT Press.

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

Audretsch, D. B., Keilbach, M. C., and Lehmann, E. E. (2006) *Entrepreneurship and Economic Growth*, Oxford University Press.

Audretsch, D. B., and Lehmann, E. E. (2005) Does the knowledge spillover theory of entrepreneurship hold for regions? *Research Policy*, 34: 1191-1202.

Audretsch, D. B., and Stephan, P. E. (1996) Company-scientist locational links: The case of biotechnology. *American Economic Review*, 86(3), 641-652.

Baumol, W. J. (1993) *Entrepreneurship, Management and the Structure of Payoffs*. Cambridge, MA: MIT Press.

Birch, S. (1979) *The Job Generation Process*. Cambridge, MA: MIT Program on Neighborhood and Regional Change.

Birley, S. (1987) New ventures and employment growth. *Journal of Business Venturing*, 2(2), 155-165.

Burger, M., van Oort, F., and van der Knaap, B. (2008) A treatise on the scale-dependency of agglomeration externalities and the modifiable areal unit problem. ERIM Report Series *Research in Management.*

Coffey, W., and Bailly, A. (1991) Producer services and flexible production: an explanatory analysis. *Growth and Change*, 22, 95-117.

Casson, M. C. (2003) *The entrepreneur: an economic theory* (2nd ed.) Cheltenham: Edward Elgar.

Davidsson, P. (2004) *Researching Entrepreneurship*. International Studies in Entrepreneurship.

Davidsson, P., Lindmark, L., and Olofsson, C. (1994) New firm formation and regional developments in Sweden. *Regional Studies*, 28(4): 395-410.

Davidsson, P., and Wiklund, J. (2001) Levels of analysis in entrepreneurship research: Current research practice and suggestions for the future. *Entrepreneurship Theory and Practice*, 25(4): 81-100.

Desrochers, P., and Sautet, F. (2008) Entrepreneurial policy: The case of regional specialization vs. spontaneous industrial diversity. *Entrepreneurship Theory and Policy*, 32(5), pp. 813-832.

Drucker, P. (1985) *Innovation and Entrepreneurship*. New York: Harper & Row.

Dumais, G., Ellison, G., and Glaeser, E. L. (2002). Geographic concentration as a dynamic process. *The Review of Economics and Statistics*, 84(2): 193-204.

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*, vol.18, no.2.

Duranton, G., and Purga, D. (2000) Diversity and specialization in cities: Why, where and when does it matter? *Urban Studies, vol. 37, no. 3, 533-555.*

Feldman, M. P., and Audretsch. D. B. (1999) Innovation in cities: Science-based diversity , specialization and localized competition. *European Economic Review*, 43: 409-429.

Freeman, C. (1974) *The Economics of Industrial Innovation*. Harmondsworth: Penguin Education.

Gartner, W. B. (1985) A conceptual framework for describing the phenomenon of new venture creation. *Academy of Management Review*, 10(4): 696-706.

Gartner, W. B. (1988) 'Who is an entrepreneur?' is the wrong question. *American Journal of Small Business*, 12, 4 (Spring): 11-32.

Gartner, W. B. (2001) Is there an elephant in entrepreneurship? Blind assumptions in theory development. *Entrepreneurship Theory and Practice*, 25(4): 27-39.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., and Schleifer, A. (1992) Learning in cities. *Journal of Urban Economics*, 46: 254-277.

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

Griliches, Z. (1979) Issues in assessing the contribution of research and development to productivity growth. *Bell Journal of Economics*, 10, 92-116.

Griliches, Z. (1990) Patent statistics as economic indicators: A survey. *Journal of Economic Literature*, 28: 1661-1707.

Henderson, V., Kuncoro, A., and Turner, M. (1995) Industrial development in cities. *The Journal of Political Economy*, 103(5): 1067-1090.

Hébert, R. F., and Link, A. N. (1989) In search of the meaning of entrepreneurship. *Small Business Economics* **1**, 39-49.

Hoover, E. M. (1937) *Locatinon Theory and the Shoe and Leather Industry*. Harvard University Press, Cambridge.

Hoover, E. M. (1948) *The Location of Economic Activity*. McGraw Hill, New York.

Hyytinen, A., and Ilmakunnas, P. (2007) Entrepreneurial aspirations: Another form of job search? *Small Business Economics* (2007) 29:63-80.

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

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

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.

Kirzner, I. M. (1973) *Competition and Entrepreneurship*, Chicago: University of Chicago Press.

Kirzner, I. M. (1985) *Discovery and the Capitalist Process*, Chicago: University of Chicago Press.

Klein, B., Crawford, R., and Alchian, A. A. (1978) Vertical integration, appropriable rents and the competitive contracting process. *Journal of Law and Economics*, 21: 297-326.

Klepper, S., and Sleeper, S. (2005) Entry by spinoffs. *Management Science*, Vol. 51, No. 8, pp. 1291-1306.

Kolvereid, L. (1996) Prediction of employment status choice intentions. *Entrepeneurship Theory and Practice*, 21, 47-57.

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

Low, M. B. (2001) The adolescence of entrepreneurship research: Specification of purpose. *Entrepreneurship Theory and Practice*, 25(4): 17-25.

Low, M. B., and MacMillan, I. C. (1988) Entrepreneurship: Past research and future challenges. *Journal of Management*, **14**(2), 139-161.

Malecki, E. J. (1997) Entrepreneurs, networks and economic development: a review of recent research. In *Advances in entrepreneurship, firm emergence and growth* (Vol. 3, pp. 57-118) Greenwich, CT: JAI Press.

Manski, C. F. (1977) The structure of random utility models. *Theory and Decision*, 8: 229-254.

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

McFadden, D. L. (1974) Conditional logit analysis of qualitative choice behavior, in P.Zarembka (ed.), *Frontiers in Econometrics*, Academic Press, New York.

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

Nås, S. O., Sandven, T., Eriksson, T., Andersson, J., Tegsjö, B., Lehtoranta, O., and Virtaharju, M. (2003), High-Tech spin-offs in the Nordic countries, Main Report, STEP Report 23-2003.

van Oort, F. G., and Atzema, O. A. L. C. (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.

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

van der Panne, G. (2004) Agglomeration externalities: Marshall versus Jacobs. *Journal of Evolutionary Economics*, 14: 593-604.

van der Panne, G., and van Beers, C. (2006) On the Marshall-Jacobs controversy: It takes two to tango. *Industrial and Corporate Change*, 15(5), 877-890.

Pavitt, K. (1982) R&D, patenting and innovative activities. A statistical exploration. *Research Policy*, 11: 33-51.

Polanyi, M. (1967) *The Tacit Dimension*. New York: Anchor Books.

Reynolds, P. D. (1987) New Firms: Societal contributions versus survival potential. *Journal of Business Venturing*, 2(3), 231-246.

Reynolds, P., Storey, D. J., and Westhead, P. (1994) Cross-national comparisons of the variation in new firm formation rates. *Regional Studies*, 28:4, 443-456.

Rhody, Stephen (1998). *Three Essays on Labor Market Transitions.* Ph.D. dissertation, Syracuse University, Department of Economics.

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

Schumpeter, J. A. (1934) *Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle*. Cambridge: Harvard University Press.

Shane, S. (2000) Prior knowledge and the discovery of entrepreneurial opportunities. *Organization Science*, 11, 448-469.

Shane, S., and Venkataraman, S. (2000) The promise of entrepreneurship as a field of research. *The Academy of Management Review*, vol. 25, No. 1, pp. 217-226.

Venkataraman, S. (1997) The distinctive domain of entrepreneurship research: An editor's perspective. In J. Katz & R. Brockhaus (Eds.), *Advances in Entrepreneurship, Firm Emergence, and Growth*, vol. 3: 119-138. Greenwich, CT: JAI Press.

Warsh, D. (2006) *Knowledge and the Wealth of Nations.* New York: W.W. Norton & Company.

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%
Self-employed	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	Entrepr	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

(SNI 38 - 64) Low-end services					
Size of the firm	Status Next Year				
	Exit	Stay	Switch	Entrepr	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

Table 2. (continued)

(SNI 65 - 71) Financial and real-estate services					
Size of the firm	Status Next Year				
	Exit	Stay	Switch	Entrepr	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

(SNI 72 - 74) Knowledge intensive business services					
Size of the firm	Status Next Year				
	Exit	Stay	Switch	Entrepr	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
Total	126209	1332761	297224	10677	1766871
	7.14	75.43	16.82	0.60	100.00

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

	Mean	SD	Min	Max
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
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
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
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
Knowledge intensity*	0.10	0.15	0	1
Log of productivity**	6.12	1.10	-2.30	12.15
Log of regional size	11.00	1.41	5.78	12.98
PS index	1.67	1.67	0.004	58.29
HH index	14.15	2.16	4.65	18.39

*Number of observations: 8565090, **number of observations: 8529729 (due to loss of observations when taking ratios where the denominator equals zero or taking logs of negative numbers)

Table 4. MNL Estimation. Marginal effects

Variable	**Stay**	**Switch**	**Exit**	***Self-employed***
Tenure	1.184***	-0.823***	-0.323***	*-0.037****
	[0.003]	[0.003]	[0.002]	*[0.001]*
Age	1.998***	-0.507***	-1.522***	*0.030****
	[0.007]	[0.006]	[0.004]	*[0.001]*
Age2	-0.022***	0.003***	0.018***	*-0.00002****
	[0.000]	[0.000]	[0.000]	*[0.000]*
Male†	1.933***	0.024	-2.130***	*0.173****
	[0.025]	[0.019]	[0.016]	*[0.004]*
Education lvl 2†	-1.583***	2.241***	-0.757***	*0.099****
	[0.067]	[0.063]	[0.026]	*[0.013]*
Education lvl 3†	-1.153***	2.116***	-1.109***	*0.132****
	0.060]	[0.055]	[0.024]	*[0.011]*
Education lvl 4†	-0.473***	1.972***	-1.638***	*0.138****
	[0.062]	[0.058]	[0.023]	*[0.012]*
Education lvl 5†	-2.561***	3.095***	-0.767***	*0.233****
	[0.073]	[0.069]	[0.028]	*[0.016]*
Education lvl 6†	-2.497***	4.128***	-1.875***	*0.244****
	[0.082]	[0.078]	[0.027]	*[0.018]*
Education lvl 7†	-2.302***	3.475***	-1.387***	*0.214****
	[0.206]	[0.190]	[0.088]	*[0.043]*
Age of firm	0.120***	-0.094***	-0.026***	*0.00001*
	[0.002]	[0.002]	[0.001]	*[0.000]*
Declining Employment†	-1.599***	0.858***	0.690***	*0.051****
	[0.026]	[0.021]	[0.016]	*[0.005]*
Declining Sales†	-2.655***	1.943***	0.659***	*0.052****
	[0.029]	[0.023]	[0.017]	*[0.005]*
Log of Productivity	0.899***	-0.548***	-0.335***	*-0.016****
	[0.009]	[0.007]	[0.005]	*[0.001]*
Knowledge intensity	3.641***	-1.187***	-2.444***	*-0.009*
	[0.097]	[0.073]	[0.066]	*[0.015]*
Size 11-50†	-0.797***	0.923***	0.012	*-0.138****
	[0.039]	[0.032]	[0.024]	*[0.005]*
Size 51-100†	-1.226***	1.137***	0.287***	*-0.198****
	[0.050]	[0.040]	[0.030]	*[0.005]*
Size 101-250†	-1.088***	1.039***	0.292***	*-0.243****
	[0.047]	[0.038]	[0.029]	*[0.005]*
Size 251-500†	-0.974***	0.907***	0.328***	*-0.260****
	[0.052]	[0.042]	[0.032]	*[0.005]*
Size 501 +†	-0.540***	0.420***	0.489***	*-0.369****
	[0.038]	[0.030]	[0.023]	*[0.005]*
Industry category 2†	-1.191***	1.597***	-0.497***	*0.090****
	[0.028]	[0.023]	[0.017]	*[0.006]*
Industry category 3†	-1.071***	1.516***	-0.471***	*0.026**
	[0.074]	[0.064]	[0.039]	*[0.013]*
Industry category 4†	-2.577***	2.303***	0.074**	*0.199****
	[0.043]	[0.036]	[0.025]	*[0.009]*
Log of regional size	-0.126***	0.126***	-0.019***	*0.020****
	[0.009]	[0.007]	[0.005]	*[0.002]*
PS index	0.318***	-0.232***	-0.066**	*-0.019****
	[0.008]	[0.007]	[0.004]	*[0.002]*
HH index	0.099***	-0.108***	0.018**	*-0.009****
	[0.005]	[0.004]	[0.003]	*[0.002]*
Y = Pr(Choice = 1)	*0.86276*	*0.08451*	*0.04827*	*0.00443*

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.*

APPENDIX

Table A1. Alternative specification (robustness check) MNL Estimation – marginal effects

Variable	Stay	Switch	Exit	***Self-employed***
Tenure	1.190***	-0.831***	-0.320***	*-0.038****
	[0.003]	[0.003]	[0.002]	*[0.001]*
Age	1.929***	-0.442***	-1.520***	*0.033****
	[.006]	[0.005]	[0.004]	*[0.001]*
Age2	-0.021***	0.002***	0.018***	*-0.00003****
	[0.000]	[0.000]	[0.000]	*[0.000]*
Male†	1.931***	0.053**	-2.157***	*0.171****
	[0.025]	[0.019]	[0.016]	*[0.004]*
Educational Group 2†	-0.025	1.300***	-1.347***	*0.073****
	[0.038]	[0.031]	[0.022]	*[0.007]*
Age of firm	0.108***	-0.083***	-0.024***	*-0.00008**
	[0.002]	[0.002]	[0.001]	*[0.000]*
Declining Employment†	-1.705***	0.910***	0.744***	*0.050****
	[0.026]	[0.021]	[0.016]	*[0.005]*
Declining Sales†	-2.593***	1.924***	0.616***	*0.052****
	[0.029]	[0.023]	[0.017]	*[0.005]*
Log of Productivity	0.884***	-0.536***	-0.330***	*-0.017****
	[0.009]	[0.007]	[0.005]	*[0.001]*
Log of size of firm	0.019***	-0.008*	0.053***	*-0.064****
	[0.005]	[0.004]	[0.003]	*[0.001]*
Industry category 2†	-1.208***	1.561***	-0.444***	*0.090****
	[0.028]	[0.023]	[0.016]	*[0.006]*
Industry category 3†	-1.056***	1.546***	-0.505***	*0.015*
	[0.078]	[0.065]	[0.039]	*[0.013]*
Industry category 4†	-2.138***	2.167***	-0.234***	*0.206****
	[0.041]	[0.034]	[0.023]	*[0.009]*
Log of regional size	-0.082***	0.129***	-0.068***	*0.022****
	[0.009]	[0.007]	[0.005]	*[0.002]*
PS index	0.328***	-0.239***	-0.069***	*-0.019****
	[0.008]	[0.007]	[0.004]	*[0.002]*
HH index	0.093***	-0.105***	0.021***	*-0.010****
	[0.005]	[0.004]	[0.003]	*[0.001]*
Y = Pr(Choice = 1)	*0.86228*	*0.08472*	*0.04879*	*0.00239*

Notes: ***significant at 0.1%, **significant at 1%, *significant at 5%, †dummy variable, number of obs: *8509291, marginal effects and standard errors (in brackets) have been multiplied by 100.*

Table A2. MNL estimation of original specification on high tech sector (SNI 29-33 and 72-74)

Variable	**Stay**	**Switch**	**Exit**	***Self-employed***
Tenure	1.183***	-0.840***	-0.302***	*-0.040****
	[0.007]	[0.006]	[0.004]	*[0.001]*
Age	2.122***	-0.572***	-1.573***	*0.024****
	[0.017]	[0.014]	[0.009]	*[0.003]*
Age2	-0.023***	0.004***	0.019***	*-0.0001****
	[0.000]	[0.000]	[0.000]	*[0.000]*
Male†	2.184***	-0.346***	-1.985***	*0.147****
	[0.056]	[0.044]	[0.036]	*[0.009]*
Education lvl 2†	-1.1842***	2.479***	-0.761***	*0.125***
	[0.172]	[0.163]	[0.061]	*[0.037]*
Education lvl 3†	-1.244***	2.435***	-1.331***	*0.140****
	[0.151]	[0.142]	[0.054]	*[0.031]*
Education lvl 4†	-0.912***	2.560***	-1.864***	*0.216****
	[0.154]	[0.146]	[0.052]	*[0.035]*
Education lvl 5†	-2.247***	3.485***	-1.571***	*0.333****
	[0.171]	[0.163]	[0.056]	*[0.042]*
Education lvl 6†	-1.543***	3.821***	-2.559***	*0.281****
	[0.175]	[0.166]	[0.055]	*[0.040]*
Education lvl 7†	-2.406***	3.697***	-1.540***	*0.249****
	[0.338]	[0.316]	[0.131]	*[0.073]*
Age of firm	0.360***	-0.258***	-0.099***	*-0.003***
	[0.005]	[0.004]	[0.003]	*[0.001]*
Declining Employment†	-2.407***	1.527***	0.803***	*0.075****
	[0.062]	[0.050]	[0.037]	*[0.011]*
Declining Sales†	-2.311***	1.705***	0.556***	*0.049****
	[0.064]	[0.053]	[0.038]	*[0.011]*
Log of Productivity	0.617***	-0.435***	-0.187***	*0.005**
	[0.015]	[0.011]	[0.009]	*[0.003]*
Knowledge intensity	2.671***	-0.333**	-2.435***	*0.096****
	[0.153]	[0.121]	[0.100]	*[0.023]*
Size 11-50†	-1.406***	1.020***	0.507***	*-0.121****
	[0.102]	[0.083]	[0.064]	*[0.011]*
Size 51-100†	-1.569***	0.976***	0.797***	*-0.204****
	[0.124]	[0.100]	[0.079]	*[0.012]*
Size 101-250†	-1.704***	1.031***	0.953***	*-0.280****
	[0.115]	[0.093]	[0.073]	*[0.011]*
Size 251-500†	-2.946***	1.922***	1.373***	*-0.350****
	[0.129]	[0.105]	[0.082]	*[0.010]*
Size 501 +†	-1.420***	0.702***	1.202***	*-0.484****
	[0.090]	[0.072]	[0.057]	*[0.013]*
Log of regional size	-0.375***	0.329***	0.016	*0.028****
	[0.020]	[0.017]	[0.012]	*[0.004]*
PS index	0.145***	-0.124***	0.001	*-0.023****
	[0.018]	[0.015]	[0.010]	*[0.005]*
HH index	0.078***	-0.088***	0.024***	*-0.013****
	[0.013]	[0.010]	[0.007]	*[0.002]*
Y = Pr(Choice = 1)	*0.85226*	*0.09388*	*0.04906*	*0.00477*

Notes: ***significant at 0.1%, **significant at 1%, *significant at 5%, †dummy variable, number of obs: *1750476, marginal effects and standard errors (in brackets) have been multiplied by 100*