External effects of innovation on firm survival:

Evidence from computer and electronic product manufacturing and healthcare

Alexandra Tsvetkova, University of North Carolina at Charlotte

Introduction

In the last two decades, geography came into prominence as an important consideration in the study of knowledge accumulation, firm performance, and economic growth. The role of space as a determinant of economic outcomes comes primarily from the non-uniform distribution of human and social capital across territories. Accumulated knowledge, specific in each region, eventually should translate into productive applications and lead to dissimilar rates of economic growth (Ibrahim, Fallah, & Reilly, 2009). The literature argues that knowledge, innovativeness and entrepreneurship (factors that in the short-run are 'attached' to a region) play definitive role in economic outcomes.

Despite the widely held view echoed by the agglomeration theory that external knowledge and innovation are important for firm performance in general, and business survival in particular, empirical evidence on this issue is lacking. No study has so far provided empirical insights into the relationship between regional innovative environment and firm longevity, although the perspective of regional innovative systems seems to have gained popularity in the last few years (Rodriguez-Pose & Crescenzi, 2008; Uyarra, 2010).

External effects of innovation on firm survival are not straightforward. The agglomeration literature suggests that accumulated stock of knowledge should translate into profitable market applications, and contribute to business productivity and innovation. If this is true, empirical analysis would reveal positive effects of innovative environments. This perspective, however, does not take into account increased competition in the regions with more innovative economies, the so-called 'creative destruction' regime. According to this approach, the net effect of innovation in a region on firm survival within the same region is expected to be negative.

Using duration analysis, this paper empirically tests the impact of innovative environments on survival likelihood of individual non-patenting firms in two industrial sectors, namely computer and electronic product manufacturing and healthcare. These sectors represent the high technology end of the U.S. economy, which is crucial for economic competitiveness of the country.

Why is space important?

Over the last decades, an extensive body of literature has emphasized the importance of geography as an important determinant of industrial performance. Regions influence innovation, firm entry, learning, and economic growth (Scott, 2006). The importance of space for regional economic performance is not a new idea. Back in the 1920s, Marshall articulated the advantages of locational externalities associated with geographically dense networks of suppliers and customers, the character of local labor pool, and pure spillovers from one business to another, which allow firms to become more innovative by employing (modified) designs and concepts of their peers (Ibrahim et al., 2009).

Much in line with Marshall's argument on co-location, the new economic geography, or NEG (Krugman, 1991), explains the emergence and persistence of large urban agglomerations that rely on reduced transportation costs, increasing returns to scale, and benefit from interactions of closely related suppliers and consumers (Schmutzler, 1999). The intra-industry economies of localization, elaborated in the NEG, may occur through (1) economies of specialization; (2) labor market economies; and (3) knowledge spillovers (Breschi & Lissoni, 2001). The agglomeration effects are hypothesized to increase labor productivity and innovativeness of individual firms (Henderson, Shalizi, & Venables, 2001; Porter, 1990) in two possible ways. The first assumes that nearness is able to influence economic outcomes, by the sole virtue of being a part of a spatial business concentration leading to economies of scale (Gordon & McCann, 2000). The second views proximity as a facilitating condition for the exchange of resources among firms (Knoben & Oerlemans, 2006).

Extensive empirical literature supports positive effects of agglomeration on the regional and firm-level economic performance. Rodriguez-Pose and Comptour (2012) argue that industrial clusters, if they happen to form in the areas with highly trained and educated labor force, promote economic growth in the European regions. Lehto (2007) finds that closeness has a positive impact on productivity in a sample of Finnish firms. In a study of U.S. service sector firms across labor market areas (LMAs), the number of establishments per 1,000 residents increases firm survival chances (Acs, Armington, & Zhang, 2007). Wennberg and Lindqvist (2010) discover positive association between population density and firm longevity in Swedish knowledge-intensive manufacturing and services. Location close to the national capital appears to promote firm survival in Greece (Fotopoulos & Louri, 2000).

A related concept that links regions and economic growth through knowledge creation and innovation is regional innovation systems or RIS (Uyarra, 2010). It starts with the proposition that '[i]nnovation is a territorially embedded process and cannot be fully understood independently of the social and institutional conditions of every space' (Rodriguez-Pose & Crescenzi, 2008, p. 54).

Here, territorial actors and institutions are hypothesized to play an important role in regional growth. The competitive advantages of regions are also related to the institutional characteristics such as the level and structure of education and R&D activities, available financial services and so forth (Cassia, Colombelli, & Paleari, 2009). Regional innovative systems should enable regions to adjust to the existing conditions in a way that promotes sustained regional growth. Recent literature emphasizes that production and utilization of knowledge is a primary way to do so. Therefore, knowledge and the spillovers associated with it are essential for the regional economic development process (Stough & Nijkamp, 2009).

Firm survival: the role of knowledge and space

It is almost a convention in the literature that accumulated local knowledge should translate into superior firm performance. Extensive research on local knowledge spillovers describes in great detail why this relationship is expected to hold and provides vast empirical evidence to support this claim. Ibrahim and co-authors (2009, p. 412, italics in original) define knowledge spillovers as the 'useful *local sources* of knowledge found in a region, that were obtained beyond the recipient's organization, and that affected the innovation of the recipient'. The spillover literature models knowledge as a public good, which is at least partially non-rival and non-excludable. Knowledge tends to accumulate in spatially bounded areas and requires some sort of interactions to spread. The intensity of knowledge spillovers declines with distance (Adams & Jaffe, 1996; Bottazzi & Peri, 2003; Rodriguez-Pose & Crescenzi, 2008; Wang, Ma, Weng, & Wang, 2004). Firms located in the areas with intensive research by business and/or universities are expected to be more inventive and productive even when they do not participate in research activities (Koo, 2005; Zachariadis, 2003).

With respect to firm survival, though, presence of LKS is likely to have divergent effects. On the one hand, firms may learn from others in order to become more productive and efficient. In this case, businesses operating in the areas with greater stock of knowledge should live longer. On the other hand, agglomeration of businesses in a particular geographical area may be detrimental to firms located in these areas. Congestion and increased competition, common attributes of agglomerations, may lead to firm failure and exit[1]. In the U.S., Buss and Lin (1990) find no difference between firm survival rates in urban and rural areas; while Renski (2009) concludes that firms located in urban core are more likely to exit. Yet, firms located close to the capital in Austria enjoy survival rates comparable to those of other firms throughout the country (Tödtling & Wansenböck, 2003). An alternative measure of urbanization, population density, is negatively related to economic growth (Funke & Niebuhr, 2005) and survival in manufacturing and business services (Brixy & Grotz, 2007) in West German regions. The latter result is in line with the evidence from U.S. LMAs (Acs et al., 2007).

Competition is likely to intensify in the geographical areas, which enjoy higher levels of innovation. Schumpeter argued that entrepreneurs are those who combine available resources in new ways in order to create novelty in products, production, markets, supply, and organization (Dodgson, 2011). New combinations, in turn, when introduced to the market, cause discontinuity in the steady state or, as Schumpeter calls it, development. Firms are usually innovative during the initial stages of their development when they try to find their niche in the market. After a firm stops creating new combinations and settles in to running its business just like others, it looses entrepreneurial character and is likely to exit as new innovative firms keep introducing new combinations to the economy (Schumpeter, 1934). The process of driving less entrepreneurial firms out of business by more entrepreneurial ones is the nature of 'creative destruction'. More innovative regions should experience greater 'creative destruction': more firms entering and exiting the market, which is a necessary condition for development and increased productivity (Bosma, Stam, & Schutjens, 2011).

Logical model and hypotheses

The discussion in the previous subsection explains two potential mechanisms of the relationship between innovation in a region and firm survival. On the one hand, knowledge

---

[1] Abstracting from LKS, greater stock of knowledge in a region may contribute to a greater likelihood of exit via at least two other routes. The knowledge spillover theory of entrepreneurship postulates that more knowledge being produced (and unutilized) in a region should increase firm formation, thus increasing competition. At the same time, in the localities where more knowledge is generated, the incumbent firms are likely to be exposed to more business ideas. Firm owners might choose to sell off or to shut down their business in order to start something new that looks more promising.

spillovers may increase business productivity in general and to increase its survival chances. Figure 1 shows schematically a simplified logical model of the relationship.

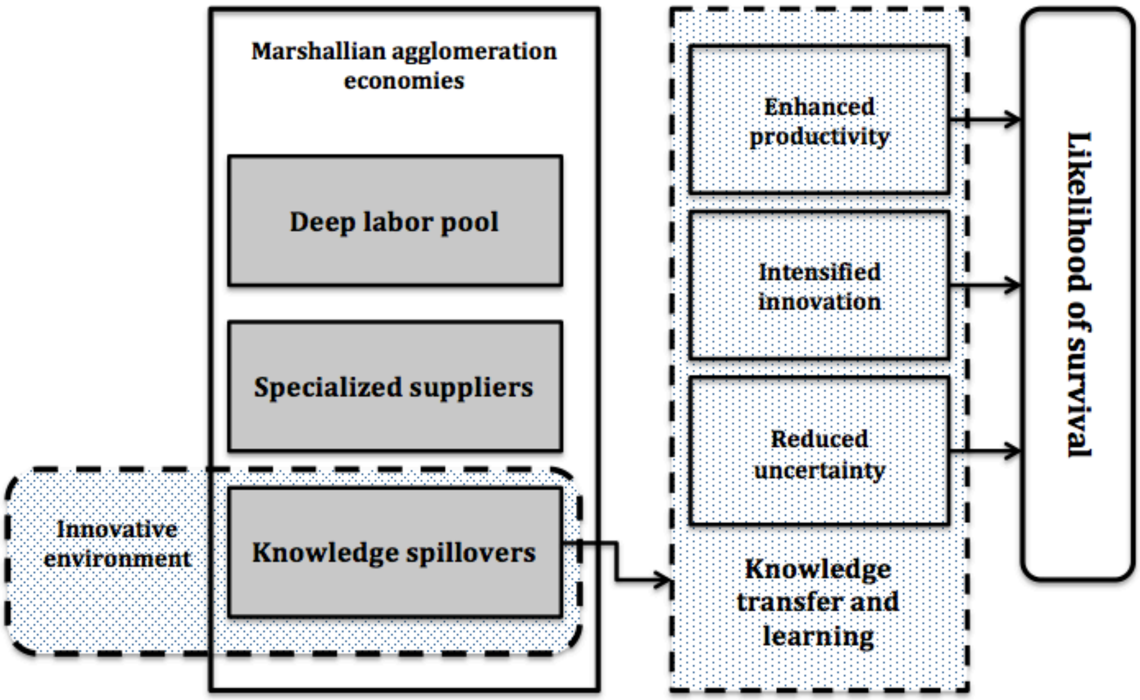

Figure 1. Schematic representation of the relationship between innovation and firm survival as it follows from the literature on knowledge spillovers

On the other hand, innovation in a region may intensify competition and facilitate exit. This mechanism, described by Schumpeter as 'creative destruction', is presented in Figure 2.

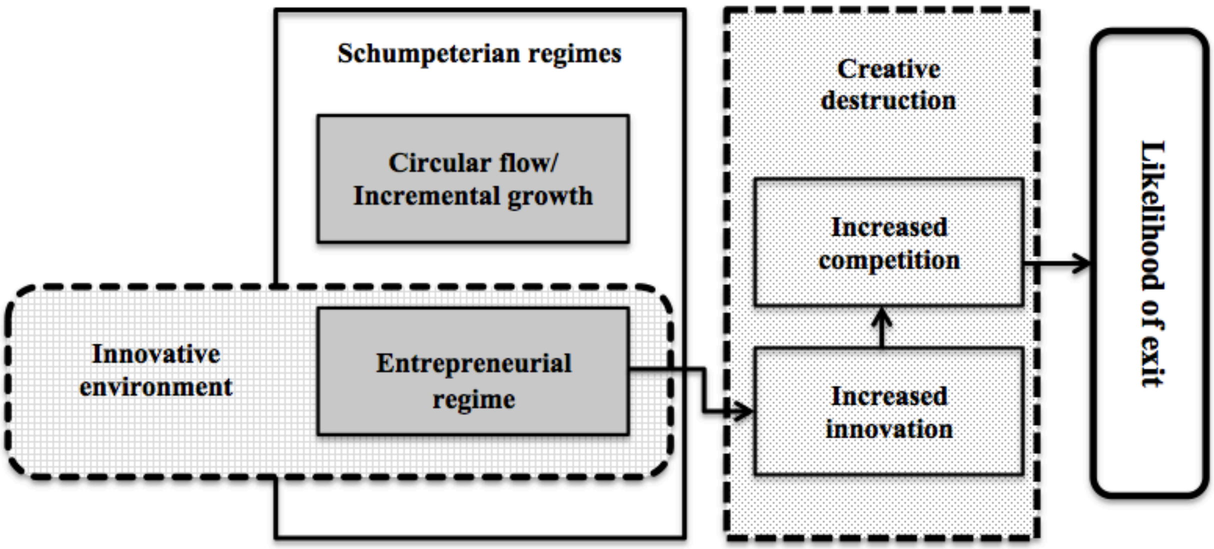

Figure 2. Schematic representation of the relationship between innovation and firm survival as follows from the literature on creative destruction

Figure 3 combines 'spillover' and 'creative destruction' perspectives on the relationship between innovative activities and business longevity in a simplified form, and adds a possibility of exit in order to start another firm, as discussed in the previous subsection.

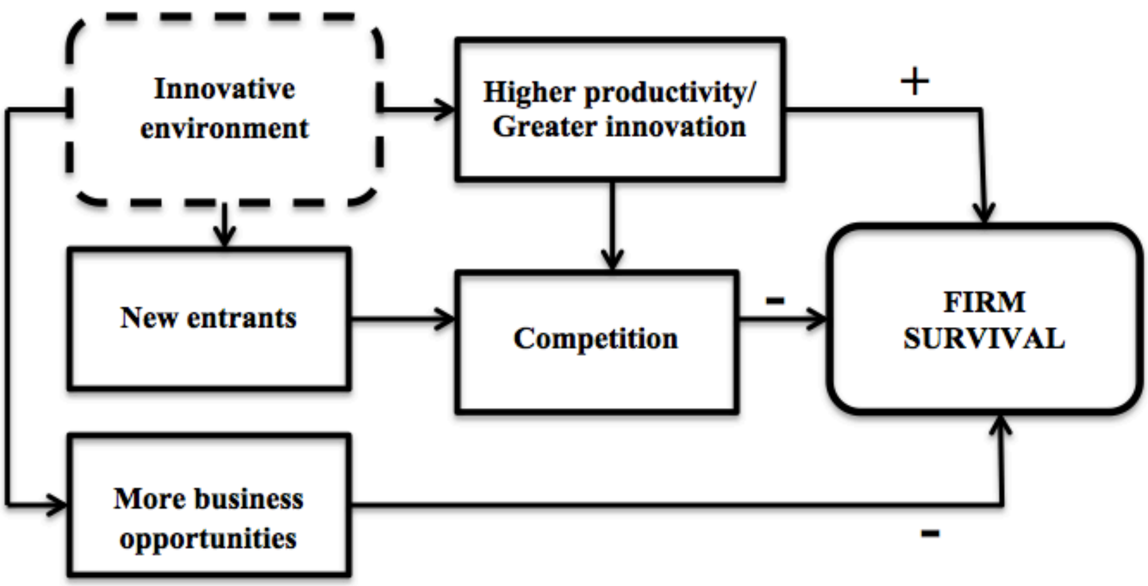

Figure 3. Logical model of the relationship between innovation in a region and firm survival

Two competing hypotheses are tested here:

H0. Innovation in a region promotes firm survival *ceteris paribus*;

HA. Innovation in a region hampers firm survival *ceteris paribus*.

Industries

The effects of the innovative environment on firm survival differ among industries (Audretsch, 1995). Highly innovative industries, which employ people inclined to pick up new ideas from the surrounding and to promptly introduce them into practice, are likely to be more perceptive to the overall innovativeness of a geographic region. Likewise, industries with a production process that allows quick implementation of innovations and experimenting without excessive sunk costs should be expected to benefit more from the level of invention in an area. Other industries due to the individual specifics might be less sensitive to the 'innovative atmosphere'. In general, intensity of spillovers depends on the industry (Glaeser, Kallal, Scheinkman, & Shleifer, 2002).

At the same time, the competition regime is likely to be unique in every industry, as should be the impact of innovation on firm longevity via competition (Fritsch, Brixy, & Falck, 2006; Segarra & Callejón, 2002). This necessitates testing for external effects of innovation on business life expectancy separately by industry. To encompass both manufacturing and service ends of the U.S. economy, we study a high-technology manufacturing sector and a high-technology service sector. Focus on high-technology sectors is determined by their substantial contribution to the national welfare and growth (Koo, 2005). The list of industries in each sector identified by 4-digit NAICS codes is given in Table 1.

Table 1. Industries included in each sector

Code	Industry
High-technology manufacturing sector	
NAICS3341	Computer and Peripheral Equipment Manufacturing
NAICS3342	Communications Equipment Manufacturing
NAICS3343	Audio and Video Equipment Manufacturing
NAICS3344	Semiconductor and Other Electronic Component Manufacturing

NAICS3345	Navigational, Measuring, Electromedical, and Control Instruments Manufacturing
NAICS3346	Manufacturing and Reproducing Magnetic and Optical Media
High-technology service sector	
NAICS6214	Outpatient Care Centers
NAICS6215	Medical and Diagnostic Laboratories
NAICS6216	Home Health Care Services
NAICS6221	General Medical and Surgical Hospitals
NAICS6222	Psychiatric and Substance Abuse Hospitals
NAICS6223	Specialty (except Psychiatric and Substance Abuse) Hospitals

Establishments

One way to approximate business longevity in a region is by calculating the average age of all existing establishments or to classify establishments in age groups (Cefis & Marsili, 2006). A related firm-level approach involves tracing survival of existing businesses through time (Lin & Huang, 2008). Focus on existing firms in survival analysis, the so-called stock sampling, excludes firms that have exited before the study commenced. This may lead to biased results, as long-living firms are overrepresented. Another approach is to divide the number of firms that survived past certain age (usually three years) by the total number of entrants. Several studies use this method (Acs et al., 2007; Brixy & Grotz, 2007). Perhaps the most robust way is to track individual firms from their birth to exit, or until the end of the study period (Renski, 2011).We follow the latter approach.

Using the National Establishment Time Series (NETS) Database[2], we identify all establishments[3] that belong to the selected industries and are created in the continental U.S. MSAs in year 1991. An establishment is assumed to be alive until the last year it is recorded in the Database (YEARLAST). We track the selected companies until the last year in the database or year 2008, whichever happens first. As a result, there are up to 17 observations for each firm. Focusing on seventeen consecutive years of business operation gives plenty of variation in the survival rates, explanatory, and control variables to capture statistically significant patterns and

[2] A more detailed description of the data sources used is given in the next subsection.

[3] The NETS Database includes records of all establishments (not firms or companies) reported by B&D. It has relationship indicators, which identify a headquarter organization for each establishment. Only stand-alone establishments (DUNS Number, primary Database identifier, is the same in ID and HEADQUARTER fields of the NETS Database) are included in the estimation; therefore, the terms ‘establishment’, ‘firm’, and ‘company’ are used interchangeably.

relationships. Most importantly, the study tracks establishments during the years that are, according to the literature, most troubled in terms of survival. The majority of exits occur in the first five to six years of operation. After this, the hazard rate is relatively low and close to being flat, at least for some time. At more 'senior' years, a firm can experience the so-called liability of senescence, which starts in some cases after 50 years of being on the market (Esteve-Perez, Sanchis-Llopis, & Sanchis-Llopis, 2004).

This research focuses on establishments located metropolitan areas. Scott (2006) argues that creativity and innovation manifest themselves most meaningfully at the urban and regional level. External effects of innovation, if present, are likely to be most pronounced in urban areas for a number of reasons. A disproportionate share of innovative activity takes place in metropolitan areas (Acs, Anselin, & Varga, 2002; Bettencourt, Lobo, & Strumsky, 2007). Knowledge spillovers, necessary for positive effects of innovation on non-inventive firms' performance, depend on density and agglomerations (Griliches, 1992; Koo, 2005; López-Bazo, Vayá, & Artís, 2004). In order to identify regional effects, one needs to use a geographic region with meaningful boundaries that encompass economic activity. States, counties, and cities are inappropriate choice for such purpose as their limits are likely to be arbitrary with regard to the existing patterns of economic activity (Acs et al., 2007). U.S. Metropolitan Statistical Areas[4] are a better choice. The definitions (boundaries) of MSAs are constantly re-defined by the Office of Management and Budget (OMB) to reflect the current state of economic linkages in the U.S. urban areas.

A number of other factors may contaminate the estimation results because each of them brings specific dynamics that needs to be studied separately. These factors include exit via M&A (Esteve-Perez, Sanchis-Llopis, & sanchis-Llopis, 2010), effects of firm's own R&D efforts (Cefis & Marsili, 2005), being a subsidiary or a headquarter (Bridges & Guariglia, 2008; Chung, Lu, & Beamish, 2008; Delios & Beamish, 2001), or being an unusually large company (Aldrich & Auster, 1986; Audretsch, Santarelli, & Vivarelli, 1999). These firms were dropped from the analysis. Table 2 presents all categories that were removed from the estimation file, and firm counts by category.

[4] I follow November 2008 definition of MSAs by OMB.

Table 2. Total number of start-ups in 1991, and establishments in the estimation file

NAICS code	High technology manufacturing[5]	High technology services
Total number of start-ups in 1991	**2,658**	**2,162**
Outside of continental USA	11	20
Outside MSAs	229	113
Not independent	261	329
Have at least one patent	172	8
Experienced M&A	12	9
Have more than 100 employees in 1992	6	32
Outliers	1	10
Missing/erroneous data in NETS	325	227
Total establishments in the sample	**1,641**	**1,414**

Source: NETS Database, U.S. PTO, The Deal Pipeline, WRDS, Alacra Store

According to the NETS Database, 2,658 establishments were started in 1991 in high-tech manufacturing sector as defined in this research (5.8% of all 1991 start-ups in the country). Firm formation in high-tech service industries considered in this dissertation was somewhat lower. Only 2,162 establishments started operation in 1991 (4.7% of total number of start-ups in the country). More than 30% of the newly created firms in each sector are excluded from estimation for various reasons. Figures 4 and 5 show the geographic distributions of 1991 start-ups in computer and electronic manufacturing and healthcare respectively.

[5] The next subsection describes high-tech manufacturing and high-tech services.

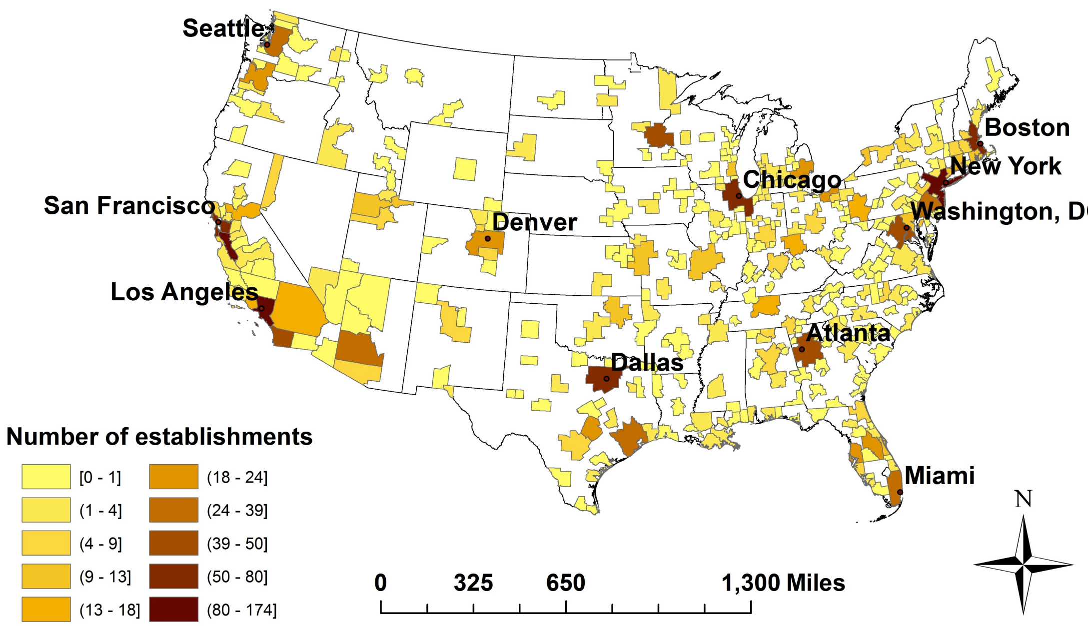

Figure 4. Total number of computer and electronic product manufacturing start-ups in continental MSAs in 1991

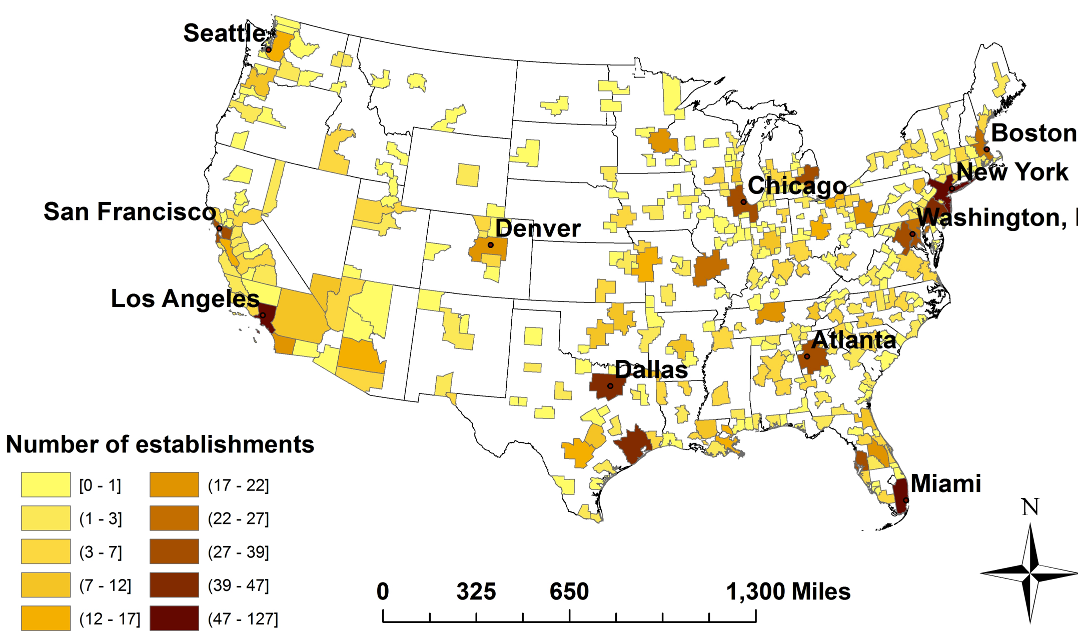

Figure 5. Total number of healthcare services start-ups in continental MSAs in 1991

As the figures show, the distribution patterns are practically identical with the new firms starting predominantly in the largest metropolitan area. This relationship disappears when population-adjusted start-up numbers are considered. Computer and electronic component manufacturing clearly tends to cluster (Figure 6), while healthcare services are more uniformly distributed (Figure 7).

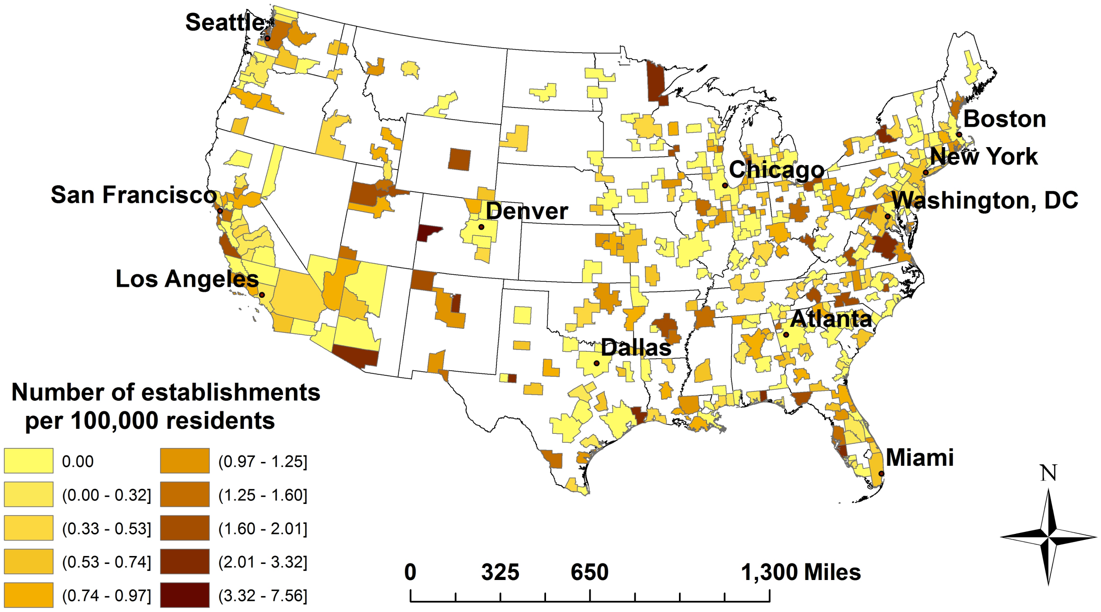

Figure 6. Number of computer and electronic product manufacturing start-ups per 100,000 residents in continental MSAs in 1991

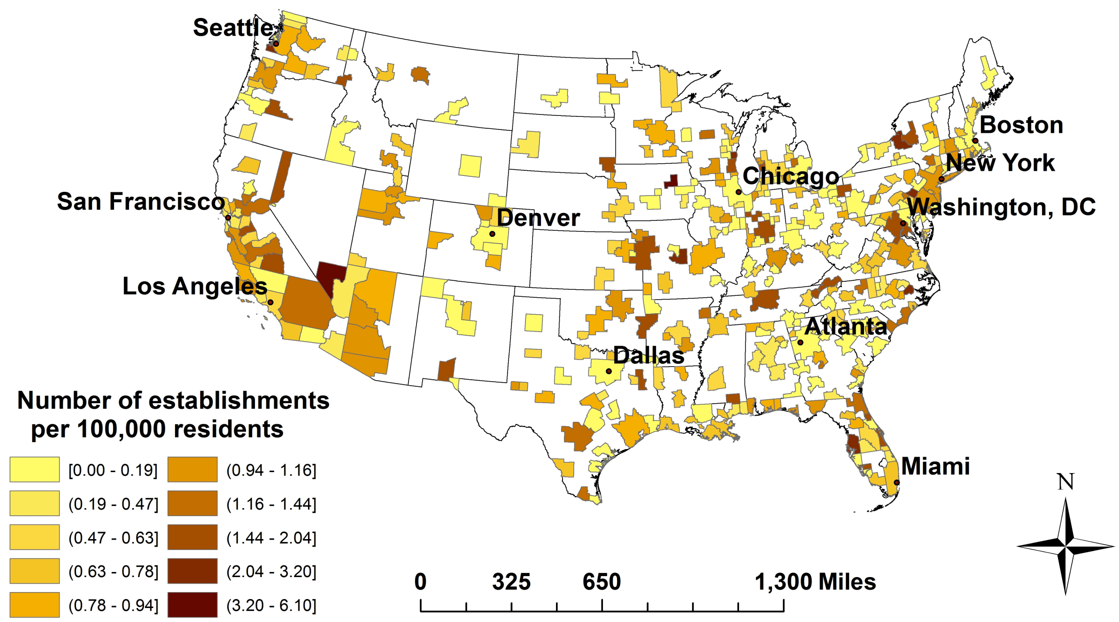

Figure 7. Number of healthcare services start-ups per 100,000 residents in continental MSAs in 1991

The sectors also differ in terms of exit patterns. Figure 8 shows firm formation and firm exit throughout the study period.

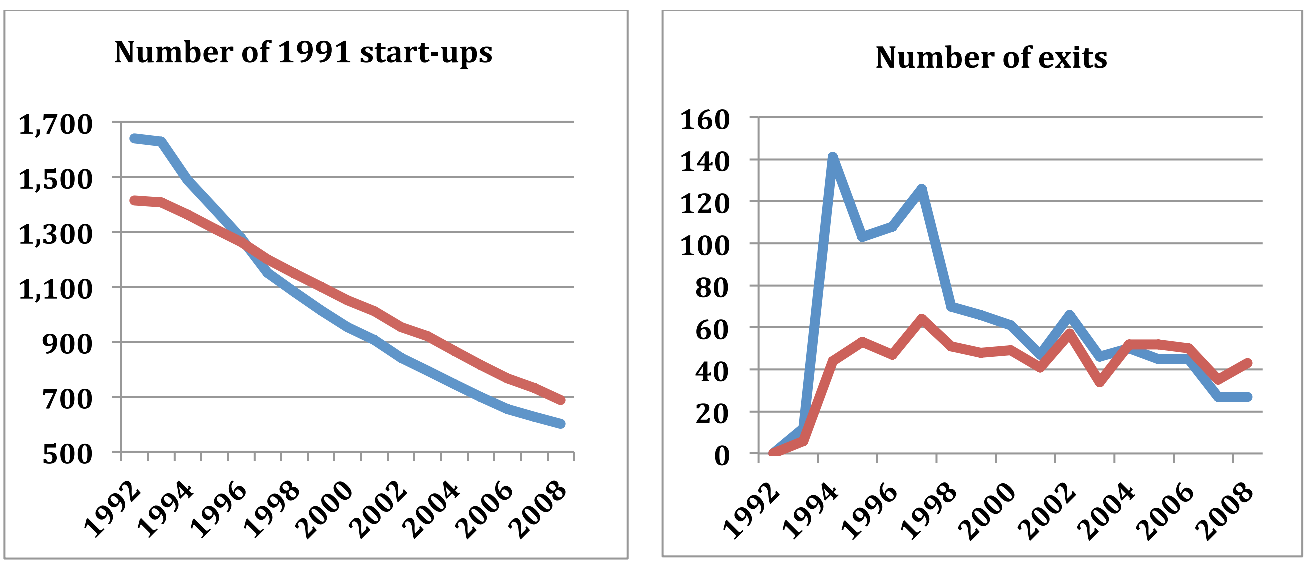

Figure 8. Firm dynamics in computer and electronic product manufacturing (blue line) and healthcare services (red line)

Variables and data sources

The dependent variable in survival analysis is waiting time till the 'event' which is exit in this research. Alternatively, the dependent variable can be defined as hazard rate. It is estimated by the model and requires information about the time of each firm's entry and exit. The time of entry is 1991 for all firms in our sample. The year of exit naturally differs among the businesses. The data come from the NETS Database, which consists of yearly snapshots of the U.S. economy taken every January since 1990. If an establishment goes out of business, its last year of operation is indicated but the record is not removed. This allows for study of active companies, and establishments that have exited. The NETS file available for this research is a subset of the original database. The file contains longitudinal information about each establishment started in 1991 including company name, county FIPS code, years of operation (first and last years in the dataset, year the business has started), industry classification (6-digit NAICS code), type of establishment (standalone, branch, headquarter), and estimated number of employees.

The level of innovativeness in the U.S. MSAs is the explanatory variable in this research. There are numerous ways to measure innovation in a region. The most common ones include R&D expenditures, share of employment in knowledge-intensive industries, and patent counts. When using patent counts as approximation for the innovativeness of a regional economy, one has to understand what exactly this statistic measures and what is does not measure. By definition, patent counts are able to account only for inventions that have been assessed and granted a patent by the U.S. PTO. Innovations that go 'unnoticed' by this governmental authority, and innovations that are denied a patent, are not captured by the patent count variable. In addition, the economic value of each patent (and, thus, its usefulness) differs greatly (Griliches, 1979; Pakes & Griliches, 1980). Despite this fact, patent count is perhaps the best readily available indicator of underlying inventive activity in a region (Acs et al., 2002; Feser, 2002; Griliches, 1990).

Patent count in a MSA per 1,000 residents (*Patents*) is the main explanatory variable in this research. The U.S. PTO maintains a patent database spanning centuries of innovation in this country. For the purposes of this study, each patent was attributed to a MSA on the basis of the inventor's reported address. If inventors residing in different MSAs are listed on a patent, corresponding share was assigned to each metropolitan area. The patent year was determined by the application date. Because of the processing and reporting delay, patents data for the last several years is incomplete. We try to mitigate this problem by adjusting the total patent counts for the years 2006, 2007, and 2008 by 5%, 10%, and 15% respectively[6] using the following formula:

$$\widehat{Patents}_{jt} = Patents_{jt} + y\overline{Patents_j} \qquad (1)$$

where $\widehat{Patents}_{jt}$ is the calculated total number of patents in MSA *j* applied for in year *t*. This number, standardized by population count in a MSA, *Patents*, is used in estimation. $Patents_{jt}$ is a patent count in MSA *j* reported by U.S. PTO for year *t*. $\overline{Patents_j}$ is the average patent count in MSA *j* over years 1992-2005. $t \in [2006, 2008]$; $y = 0.05$ if t=2006, $y = 0.1$ if t=2007, $y = 0.15$ if t=2008. Appendix 1 presents average patent counts before and after adjustment.

The control variables include the number of employees in an establishment, change in the size of the labor force from the previous year, level of entrepreneurship in a metropolitan area, educational attainment, average income, population, business density, industrial diversity, and industry dummies at 4-digit NAICS level.

Firm size is perhaps the most widely studied determinant of firm survival. There are several ways to measure a firm's size. Researchers use the number of employees, assets, or sales volume as an indicator of size. As a rule, the results are not very sensitive to the choice of the size measure (Agarwal, Sarkar, & Echambadi, 2002). Mata, Portugal and Guimaraes (1995) argue that current size is a predictor superior to start-up size. The empirical literature usually finds positive relationship between size and business longevity (Box, 2008; Persson, 2004). We use the current number of employees (*lnSize*) to measure the size of an establishment. Firms with more than 100 employees at start-up are excluded from the analysis. In healthcare, firms tend to be larger than in computer and electronic product manufacturing. The NETS Database is the data source. To ensure a better linear fit, we use a log transformation of the firm's size.

[6] To the best of our knowledge, no study has adjusted the patent counts using a validated and well-defined method. There is no information in the literature as to what fraction of the successful patent applications is left out of the U.S. PTO database each year. It means that the adjustment values have to be picked up at a researcher's discretion. To keep the estimation conservative and not to overinflate innovativeness of the U.S. MSAs, we selected 5%, 10%, and 15% (not, for example, 10%, 15%, and 20%).

Another firm-level control variable is the change in size from the previous year (*Expand*). Post-entry performance is as important for business longevity as the entry process (Audretsch & Mata, 1995). Firms tend to grow during their life span. The relationship between firm the change in size survival can be both positive and negative. On the one hand, expansion is likely to signal that the business has a greater potential. On the other, growth involves new challenges and some firms might be unable to cope with them. Likewise, firm contraction may be a signal of either hardships, or greater efficiency. Variable *Expand* is calculated by subtracting the number of employees in the previous year from the current employee number as reported by the NETS Database.

The level of entrepreneurship in a region is potentially a strong predictor of firm performance but its effect is not straightforward. If entrepreneurial spirit translates into a better fit between a firm's activities and market demand, entrepreneurship should facilitate firm survival. On the other hand, a greater number of new firms, a common measure of entrepreneurship, is likely to lead to fiercer competition and thus reduce business survival chances. We approximate entrepreneurship by the number of start-ups per 1,000 residents in a MSA (*Entrepreneurship*). The variable is calculated by dividing the number of new firms in all industries by the estimated population. The data come from the NETS Database and the U.S. Census Bureau.

The average level of education in a MSA serves as a good approximation for the quality of a labor pool and human capital available in an area. We include the total number of graduates with a bachelor degree or higher per 1,000 residents to control for the level of educational attainment. The variable *lnGrad* is calculated using the Integrated Postsecondary Education Data System files available at http://nces.ed.gov/ipeds/datacenter/. The Data System reports total number of completions and their level for each post-secondary educational institution among other indicators. Completions with level five or higher were aggregated using locational information into MSA-level variable. To ensure a better linear fit of the model, a natural logarithm of the variable is used in estimation

We control for size[7] and performance of a metropolitan economy by including estimated average income in the models (*lnInc*). Average income is related to local prosperity and market

[7] Gross Metropolitan Product (GMP), a more traditional approximation for the size of a metro economy, is available for years 2001-2008 only. Presumably GMP and MSA income are calculated using the same methodology and indicators. Correlation between income in the U.S. MSAs and GMP in 2001 – 2008 period is above 0.99. This makes these two measures practically identical for the estimation purposes.

depth. MSAs with higher income may promote firm longevity by ensuring greater demand, more resources, and business possibilities. At the same time, costs of doing business in such metropolitan areas are likely to be higher. The U.S. Bureau of Economic Analysis compiles per capita income estimates, which are available at http://www.bea.gov/regional/downloadzip.cfm for years 1967 - 2009. Estimated per capita income in logarithmic form is used in the models.

Population is an important determinant of firm performance in the agglomeration and knowledge spillover literature. Population may approximate market opportunities as demand for goods and services is likely to be higher in more populated areas. Population size is usually positively related to firm entry (Audretsch & Fritsch, 1994). In our dataset, the pairwise correlation between the number of 1991 start-ups in a metropolitan area and its population is 0.87 in computer and electronic product manufacturing, and 0.92 in healthcare. Apparently, new firms that belong to the sectors of interest tend to locate in more populated regions. We include a log transformation of the estimated metropolitan population size (*lnPop*) in the models[8]. The data come from the BEA.

Business concentration may signal the availability of the resources necessary for business success. At the same time, localized economies are likely to impose competitive pressure and increase hazard rate. The survival literature reports both positive (Renski, 2011; Wennberg & Lindqvist, 2010) and negative (Acs et al., 2007; Sorenson & Audia, 2000; Stuart & Sorenson, 2003) effects of industrial concentration on firm survival. We calculate variable *Density* by dividing the total number of establishments in a sector of interest by the MSA land area. Firm count is derived from the NETS Database by aggregating establishment-level data into MSA-level variables. Land area of the U.S. counties is available from the U.S. Census Bureau. MSA land area is calculated by adding together land areas of counties belonging to a MSA according to the November 2008 definition of MSAs by OMB.

Industrial diversity promotes recombination of ideas and innovation (Feldman & Audretsch, 1999; Jacobs, 1969). Besides, diversification of the economy can alleviate negative economic trends and to promote spillover effects (Stel & Nieuwenhuijsen, 2004). Renski (2011) finds that

---

[8] Alternative, and potentially more relevant, measures such as population density, total employment density, and employment density in corresponding industries produce similar estimation results.

industrial diversity has positive effect on firm survival, especially in knowledge intensive industries. Following a number of studies (Attaran, 1986; Bishop & Gripaios, 2007), we use an entropy measure to approximate the diversification of the economy in each metro area. As suggested by Bishop and Gripaios (2007, p. 1745), the total diversity is calculated using the following formula.

$$TD = \sum_1^n S_i \ln\left(\frac{1}{s_i}\right) \qquad (3)$$

where S_i stands for the share of the 3-digit NAICS category in a MSA employment and there are *n* such categories. The total diversity index is zero if all employment is concentrated in one sector and it is maximized if employment is distributed evenly among the sectors. The measure is also dependent on the total number of sectors with the share of each sector to be weighted by the logarithmic function. We calculate the variable *Diversity* from the NETS Database using NAICS classification and firm level employment that is aggregated into total MSA employment.

In addition to the controls described above, the models include a set of dichotomous variables that identify industry affiliation of each firm at 4-digit NAICS level. In computer and electronic product manufacturing, NAICS 3341, Computer and Peripheral Equipment Manufacturing, is the reference category. The following dummies enter the models: *NAICS3342* (Communications Equipment Manufacturing), *NAICS3343* (Audio and Video Equipment Manufacturing), *NAICS3344* (Semiconductor and Other Electronic Component Manufacturing), *NAICS3345* (Navigational, Measuring, Electromedical, and Control Instruments Manufacturing), and *NAICS3346* (Manufacturing and Reproducing Magnetic and Optical Media). In healthcare, NAICS 6214 (Outpatient Care Centers) is the reference category. Dichotomous variables *NAICS6215* (Medical and Diagnostic Laboratories), *NAICS6216* (Home Health Care Services), *NAICS6221* (General Medical and Surgical Hospitals), *NAICS6222* (Psychiatric and Substance Abuse Hospitals), and *NAICS6223* (Specialty (except Psychiatric and Substance Abuse) Hospitals) are included in the models.

Estimation approach

We use survival analysis[9] to estimate the effect of innovation in a metropolitan area on the likelihood of business survival. Survival analysis is a technique that models time till a certain event happens. It provides tools to analyze time only or time as a function of covariates. For the purpose of this research, the unit of observation is a standalone firm; event means firm exit, and survival time is the time a firm stayed in the market (from its birth to the event). Survival analysis can utilize non-parametric, semi-parametric (Cox regression), or parametric (e.g. Weibull, exponential, log-normal) techniques depending on the peculiarities of the data, and the goal of the analysis.

We use all these approaches to elicit the relationship of interest. We start with non-parametric analysis, which assumes neither a specific distribution of failure times, nor a functional form of the relationship between survival and independent variables. Although non-parametric analysis is unable to adequately deal with complex relationships involving many covariates, it is a powerful tool for preliminary exploration and visualization of the survival patterns observed in the data. If the sample can be divided into groups based on a variable of interest, non-parametric methods compare survival experiences across groups.

In the study of social phenomena, which usually emerge as a result of a multitude of various forces and factors, making assumptions about underlying distributions of events might be a difficult task. For this reason, less parametrization is always preferred to more parametrization (Box-Steffensmeier & Jones, 2007). We use semi-parametric analysis whenever the assumptions of this approach are satisfied. In most general terms, semi-parametric analysis is a combination of separate binary-outcome analyses performed at ordered individual failure times separately. This approach does not assume any distribution of the failure times (only order matters) but it still assumes some distribution of the effects of the covariates, hence the name 'semi-parametric'.

Let us write the hazard $h_j(t)$ as some function of time and covariates:

$$h_j(t) = g(t, \beta_0 + \boldsymbol{x}_j \boldsymbol{\beta}_x) \tag{4}$$

where hazard, $h_j(t)$, is the intensity with which failure occurs for subject j during time t. The Cox proportional hazards regression model assumes that independent variables shift (but do not

[9] A more formal treatment of the survival analysis techniques can be found in Allison (2010), Box-Steffensmeier and Jones (2007), and Cleves et al. (2010)

change) the hazard everyone faces. This hazard is called baseline hazard and is denoted by $h_0(t)$. Incorporating $h_0(t)$ in (4) and assuming the exponential functional form, we get

$$h_j(t) = h_0(t)\exp(\beta_0 + \boldsymbol{x_j}\boldsymbol{\beta}_x) \Leftrightarrow h(t|\boldsymbol{x}_j) = h_0(t)\exp(\beta_0 + \boldsymbol{x_j}\boldsymbol{\beta}_x) \quad (5)$$

Since the shape of the baseline function is presumed to be the same for everyone, in calculations it cancels out and remains unspecified.

The Cox model assumes proportionality of the baseline hazard. We use STATA's `estat phtest` command, which tests for the relationship between Schoenfeld residuals from the Cox regression and a smooth function of time (Cleves et al., 2010). Two conclusions emerge from the test. First, small firms and large firms have different baseline hazard functions regardless of the sector[10]. Besides, the baseline hazard function differs across the NAICS4 codes in both sectors. The proportionality assumption is satisfied when the test is performed separately by 4-digit NAICS industry code.

When hazard functions are not proportional, as it is the case in our pooled dataset, the Cox regression is not appropriate and parametric estimation has to be used. Parametric models assume a specific distribution of failure times. If a choice of distribution is correct, parametric estimators are more efficient. Analysis of the literature together with a visual inspection of exit patterns by year suggests increasing and then decreasing hazard faced by firms in the sectors of interest. Two distributions are able to accommodate non-monotonic hazard functions, namely log-normal and log-logistic[11]. We use the latter as it is less demanding computationally.

The hazard function in log-logistic regression is assumed to be of the form[12]

$$h(t|x_j) = \frac{-\frac{d}{dt}S(t|x_j)}{S(t|x_j)} \quad (6)$$

where $S(t|x_j)$ is the survival function that depends on time t and covariates x. Log-logistic model is an accelerated failure time (AFT) model. Covariates accelerate time by acceleration parameter $exp(-x_j\beta_x)$, thus,

[10] According to Klette and Kortum (2004) different hazard faced by small and large firms is a well-established fact in the literature. The authors state it as a stylized fact.

[11] They are practically indistinguishable for the purposes of estimation.

[12] Discussion in this subsection is based on Cleves et al. (2010)

$$S(t_j|x_j) = S_0\ \{\exp\ (-\boldsymbol{x}_j\boldsymbol{\beta}_x)t_j\} \tag{7}$$

A function of failure time t_j, $\tau_j = exp(-x_j\beta_x)t_j$ is assumed to be distributed as log-logistic with mean β_0 and variance γ (Cleves et al., 2010). STATA package is used to estimate all models.

Estimation results and discussion

To perform a non-parametric analysis, we divide all MSAs into more innovative and less innovative based on the patenting intensity. For each year, a median number of patents per 1,000 residents was calculated. Metropolitan areas that enjoyed patenting intensity above the median value were assigned into the more innovative group. All the rest belong to the less innovative group. Figure 9 presents smoothed hazard estimates for both sectors.

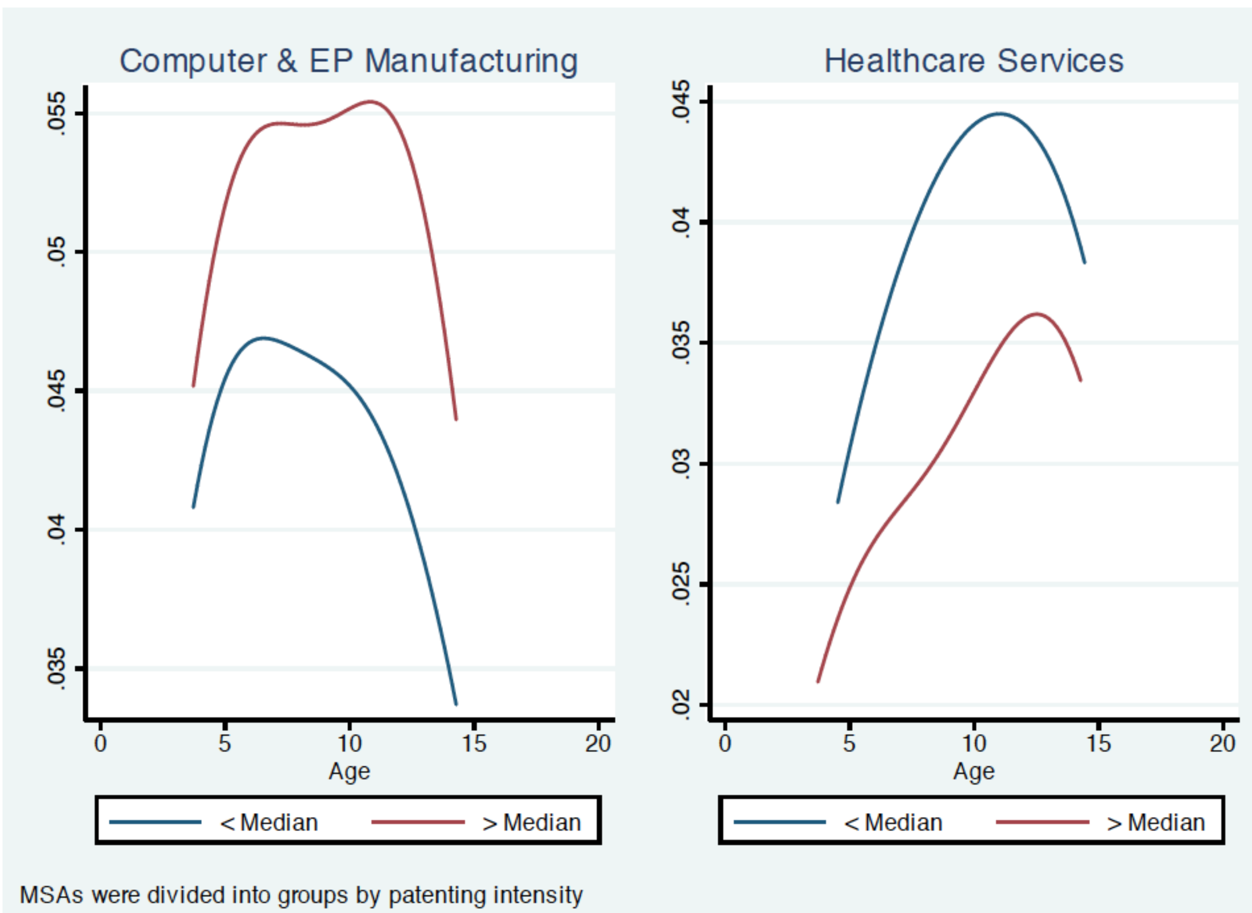

Figure 9. Empirical smoothed hazard estimates for the firms in more innovative (red line) and less innovative (blue line) MSAs

A visual inspection of Figure 9 reveals a negative relationship between innovation in a MSA and firm survival in computer and electronic product manufacturing and a positive one in healthcare services. This result, however, is derived without controlling for other relevant factors.

To overcome this weakness, we perform a multivariate analysis. The models that include all explanatory variables cannot be estimated by the semi-parametric technique due to violation of the proportionality assumption. We estimate the models using log-logistic regression. Tables 3 and 4 contain the results for the two sectors separately. The log-logistic regression models the likelihood of survival, that is, negative coefficients in the two tables below indicate adverse effect of a variable on business longevity.

Table 3. Log-logistic regression estimation results for computer and electronic product manufacturing

Variable	**Coef.**	**Robust Std. Err.**	**P>z**
lnPat	-0.237***	0.042	0.000
lnSize	-0.061**	0.026	0.019
Expand	-0.002***	0.001	0.000
lnGrad	0.066***	0.009	0.000
lnInc	1.445***	0.198	0.000
lnPop	-0.114***	0.041	0.006
Diversity	0.412	0.261	0.114
Density	0.000	0.000	0.879
NAICS3342	0.126	0.095	0.185
NAICS3343	0.274***	0.106	0.010
NAICS3344	0.367***	0.084	0.000
NAICS3345	0.328***	0.075	0.000
NAICS3346	0.698***	0.072	0.000
Constant	-2.552	1.067	0.017
Gamma	0.536	0.016	
# of subjects	1,641		
# of failures	782		
# of obs.	17,493		
# of MSAs	197		
Wald $\chi 2$ (13)	310.38		
Prob > $\chi 2$	0.000		

*** - significant at 0.01 level; ** - significant at 0.05 level; * - significant at 0.1 level.
Note: standard errors (in parentheses) adjusted for MSAs

Innovation in a MSA, measured by total number of patents per 1,000 residents, negatively affects survival likelihood of independent non-patenting firms located in that MSA. This result is in line with the conclusions of non-parametric analysis and confirms Schumpeterian argument of 'creative destruction'. Almost all control variables are significant. The analysis suggests that larger firms, as well as companies that expanded in the previous year, are more likely to exit. This result should be interpreted carefully. Most likely, independent companies usually start small and grow over time. As time passes, hazard first increases and then decreases. The observed negative relationship between size and survival may be an artifact of actual association between the size and the time passed, which is also related to higher exit likelihood during certain period. Another explanation is that larger firms indeed tend to exit sooner as a result of insufficient productivity per worker. Larger firms are more difficult to manage and may require more resources to prosper. The negative effect of expansion indirectly supports the argument of the difficulties associated with growth management that translates into a higher hazard. Average income in a metropolitan area and education level seem to promote survival. The former characteristic approximates the depth of the market and the resources available to the firms located in more affluent MSAs. The latter signals the quality of the labor pool companies may draw upon. High-technology firms tend to live less in more populated metropolitan areas, while neither industrial density, nor diversity are significant.

Significance of industry dummy variables suggests different survival dynamics depending on the NAICS4 industry in focus. Numerous empirical studies demonstrate that firm performance varies by industry, sector, and stage of product life cycle (Agarwal, 1997; Agarwal & Gort, 2002; Renski, 2009). Industries included in this analysis represent a variety of technological processes, market and industry characteristics, and stages of the main products produced. Significance of industry indicators in this case is expected and warrants industry-specific analysis, performed below.

Table 4 presents the results of corresponding analysis in healthcare services. In this sector, innovation appears to be unrelated to firm survival. After controlling for a number of individual, industrial, and regional characteristics, the coefficient of innovation measure is negative, contrary to the non-parametric analysis results.

Table 4. Log-logistic regression estimation results for healthcare services

Variable	Coef.	Robust Std. Err.	P>z
lnPat	-0.017	0.050	0.738
lnSize	-0.062***	0.020	0.002
Expand	0.005	0.004	0.230
lnGrad	0.002	0.013	0.871
lnInc	0.915***	0.254	0.000
lnPop	-0.127***	0.040	0.002
Diversity	-0.322	0.212	0.129
Density	0.001**	0.000	0.036
NAICS6215	-0.294***	0.083	0.000
NAICS6216	-0.198***	0.063	0.002
NAICS6221	0.151	0.141	0.284
NAICS6222	0.076	0.084	0.369
NAICS6223	-0.105	0.085	0.219
Constant	1.404	1.154	0.224
Gamma	0.478	0.020	
# of subjects	1,414		
# of failures	562		
# of obs.	18,017		
# of MSAs	254		
Wald $\chi2$ (13)	113.33		
Prob > $\chi2$	0.000		

*** - significant at 0.01 level; ** - significant at 0.05 level; * - significant at 0.1 level.
Note: standard errors (in parentheses) adjusted for MSAs

Consistently with the high-tech manufacturing estimation results, size is negatively related to longevity. Larger firms are more likely to exit. Average income in a metropolitan area seems to promote business survival, and population size to hinder it. In contrast to high-technology manufacturing, industrial diversity appears to be a significant predictor of increased longevity, although the coefficient is very small. Only two out of five industrial indicators are significant. Medical and diagnostic laboratories and home healthcare services have a survival dynamics different from outpatient care centers, the reference category. All other industries, which include various types of hospitals, are identical to outpatient care centers in terms of longevity.

Tables 3 and 4 imply that the industry a firm belongs to is significantly related to business survival. We therefore proceed with separate semi-parametric analyses at industry level. Tables 5 and 6 report estimation results for computer and electronic product manufacturing and healthcare services respectively. The Cox regression used in estimation, unlike log-logistic regression, models hazard ratios. Coefficients greater that one in the tables below indicate increased probability of exit (higher hazard) and coefficients below one suggest the opposite.

Table 5. Cox regression results (hazard ratios) for computer and electronic product manufacturing by NAICS4 industry code

Var.\NAICS	**3341**	**3342**	**3343**	**3344**	**3345**	**3346**
lnPat	1.286* (0.169)	1.699*** (0.334)	1.189 (0.270)	1.225 (0.170)	0.969 (0.123)	1.206 (0.160)
lnSize	1.099* (0.061)	1.002 (0.096)	1.488*** (0.173)	0.951 (0.059)	1.115 (0.076)	1.184* (0.121)
Expand	0.995 (0.006)	0.992 (0.008)	0.979 (0.027)	0.998 (0.006)	1.004*** (0.001)	0.996 (0.0108)
lnGrad	0.967 (0.034)	0.998 (0.055)	1.191 (0.177)	0.865*** (0.040)	0.921* (0.042)	0.860*** (0.047)
lnInc	0.450 (0.305)	0.041*** (0.035)	0.100** (0.098)	0.229** (0.136)	1.469 (0.858)	0.642 (0.476)
lnPop	1.241** (0.123)	0.903 (0.131)	1.074 (0.195)	1.152 (0.138)	0.967 (0.081)	0.978 (0.158)
Diversity	0.588 (0.328)	0.496 (0.478)	1.770 (2.257)	0.364 (0.224)	1.069 (0.679)	0.966 (0.778)
Density	1.000 (0.001)	1.003** (0.001)	1.001 (0.001)	0.999 (0.001)	1.000 (0.001)	1.00 (0.001)
# of subjects	373	162	92	345	366	303
# of failures	216	87	47	161	177	94
# of obs.	3,304	1,589	1,023	3,725	4,054	3,798
# of MSAs	91	63	41	95	122	100
Wald $\chi2$ (8)	28.88	22.29	20.52	28.22	75.25	18.59
Prob > $\chi2$	0.000	0.004	0.009	0.000	0.000	0.017

*** - significant at 0.01 level; ** - significant at 0.05 level; * - significant at 0.1 level.
Note: standard errors (in parentheses) adjusted for MSAs

A quick inspection of Table 5 suggests that effects of explanatory and control variables differ by industry. Innovation appears to be a significant predictor of firm exit only in Computer and Peripheral Equipment Manufacturing, and Communications Equipment Manufacturing. It is not statistically related to the hazard faced by companies in the other industries. In Computer and Peripheral Equipment Manufacturing, one unit increase in patent count per 1,000 residents, that is, 100 additional patents in a MSA with a population of 100,000, is associated with almost 30%

higher probability of exit. The same change in explanatory variable in Communications Equipment Manufacturing is associated with almost 70% increase in hazard.

Larger companies are more likely to exit only in two industries, Computer and Peripheral Equipment Manufacturing and Audio and Video Equipment Manufacturing. Expansion is negatively related to survival in Navigational, Measuring, Electromedical, and Control Instruments Manufacturing. These two variables are insignificant in the other industries. Firms in NAICS3344, NAICS3345, and NAICS3346 enjoy somewhat greater expected longevity in metropolitan areas with higher level of education. Business establishments in NAICS3342, NAICS3343, and NAICS3344 appear to be sensitive to average income. They survive much longer in more affluent MSAs. The size of metropolitan population has negative effect on firm longevity in NAICS3341, while industrial density slightly promotes survival in NAICS3342.

Table 6 shows that in healthcare services survival dynamics is driven by factors other than localization economies or knowledge spillovers. The model estimated in the table below is significant only in the case of Medical and Diagnostic Laboratories. The results suggest a positive relationship between firm survival and average per capita income in a MSA, and a negative relationship between survival and population size.

Table 6. Cox regression results (hazard ratios) for healthcare by NAICS4 industry code

Var.\NAICS	**6214**	**6215**	**6216**	**6221**	**6222**	**6223**
lnPat	0.942 (0.147)	1.066 (0.166)	0.907 (0.106)	0.590 (0.251)	0.915 (0.192)	0.950 (0.219)
lnSize	1.039 (0.074)	1.107 (0.087)	1.074 (0.055)	1.333 (0.317)	1.080 (0.175)	1.207 (0.134)
Expand	0.995 (0.011)	0.983 (0.016)	0.999 (0.005)	0.816 (0.119)	0.989 (0.0169)	0.949 (0.048)
lnGrad	1.033 (0.048)	1.076 (0.084)	0.998 (0.039)	1.055 (0.808)	1.140** (0.072)	0.915 (0.065)
lnInc	0.330 (0.283)	0.159** (0.129)	0.631 (0.447)	3.605 (11.511)	0.510 (0.554)	3.403 (3.319)
lnPop	1.022 (0.135)	1.330** (0.178)	1.153 (0.118)	0.798 (0.297)	0.949 (0.229)	1.146 (0.195)
Diversity	2.026 (1.451)	0.453 (0.346)	2.482 (1.758)	11.466 (40.396)	4.884 (5.843)	1.203 (1.204)
Density	1.000 (0.001)	1.000 (0.001)	1.000 (0.001)	1.000 (0.002)	1.000 (0.001)	1.000 (0.001))
# of subjects	390	280	395	49	122	178
# of failures	134	127	184	12	42	63
# of obs.	5,278	3,261	4,977	558	1,760	2,183

# of MSAs	147	95	140	30	73	98
Wald $\chi 2$ (8)	8.45	17.59	7.07	5.74	10.14	12.12
Prob > $\chi 2$	0.391	0.025	0.529	0.676	0.2554	0.1458

*** - significant at 0.01 level; ** - significant at 0.05 level; * - significant at 0.1 level.
Note: standard errors (in parentheses) adjusted for MSAs

Conclusion

The purpose of this paper is to estimate the external effects of innovation, one of the major determinants of regional and business performance, on firm survival. Justification for the relationship between regional innovation and business longevity comes primarily from the agglomeration literature. It postulates that in the areas of business concentration, knowledge spillovers are likely to happen leading to increased productivity and innovativeness of all companies, even those not engaged in purposeful generation of the new knowledge. Empirical studies relate both productivity and innovativeness to a greater likelihood of survival.

The preliminary non-parametric analysis reveals that higher patenting intensity in a metropolitan area is related to increased hazard in computer and electronic product manufacturing, while it appears to promote business longevity in healthcare services. The results of the multivariate duration analysis suggest that, if statistically significant relationship between innovation and firm survival exists, it is negative. Overall innovativeness in a metropolitan area seems to impose competitive pressure and to force independent non-patenting companies to exit sooner in Computer and Peripheral Equipment Manufacturing, and Communications Equipment Manufacturing. In all other industries, both manufacturing and services, the relationship is not statistically meaningful, although the coefficients in the majority of the cases indicate increased hazard faced by firms that locate in more innovative MSAs.

From the standpoint of agglomeration theory, findings in this dissertation imply that knowledge spillovers, if present, do not translate into increased survival chances for standalone non-patenting establishments. These, results, however, are nicely in line with the Schumpeterian perspective on innovation. According to the economist, more innovative environments should impose greater competitive pressure on all businesses, and should stimulate firm exit. The analysis in this paper supports such a view. Insignificance of innovation in many instances, though, could be possibly explained by the counteraction of the mechanisms simultaneously promoting firm longevity and hindering it, which results in a zero net effect.

References:

Acs, Z. J., Anselin, L., & Varga, A. (2002). Patents and innovation counts as measures of regional production of new knowledge. *Research Policy, 31*, 1069–1085.

Acs, Z. J., Armington, C., & Zhang, T. (2007). The determinants of new-firm survival across regional economies. *Papers in Regional Science, 86*(3), 367-391.

Adams, J. D., & Jaffe, A. (1996). Bounding the Effects of R&D: An Investigation Using Matched Establishment-Firm Data. *RAND Journal of Economics, 27*(4), 700-721.

Agarwal, R. (1997). Survival of firms over the product life cycle. *Southern Economic Journal, 63*(3), 571-585.

Agarwal, R., & Gort, M. (2002). Products and firm life cycles and firm survival. *American Economic Review, 92*(2), 184-190.

Agarwal, R., Sarkar, M., & Echambadi, R. (2002). The conditioning effect of time on firm survival: An industry life cycle approach. *Academy of Management Journal, 45*(5), 971-994.

Aldrich, H., & Auster, E. R. (1986). Even dwarfs started small: liabilities of age and size and their strategic implications. *Research in Organizational Behavior, 8*, 165-198.

Allison, P. D. (2010). Survival Analysis. In G. R. Hancock & R. O. Mueller (Eds.), *The Reviewer's Guide to Quantitative Methods in the Social Sciences* (pp. 413-425). New York: Routledge.

Attaran, M. (1986). Industrial diversity and economic performance in U.S. areas. *The Annals of Regional Science, 22*(2), 44-54.

Audretsch, D. (1995). Innovation, growth and survival. *International Journal of Industrial Organization, 13*, 441-457

Audretsch, D., & Fritsch, M. (1994). The geography of firm births in Germany. *Regional Studies, 28*(4), 359-365.

Audretsch, D., & Mata, J. (1995). The post-entry performance of firms: Introduction. *International Journal of Industrial Organization, 13*(4), 413-419.

Audretsch, D., Santarelli, E., & Vivarelli, M. (1999). Start-up size and industrial dynamics: some evidence from Italian manufacturing. *International Journal of Industrial Organization, 17*, 965–983.

Bettencourt, L. M. A., Lobo, J., & Strumsky, D. (2007). Invention in the city: Increasing returns to patenting as a scaling function of metropolitan size. *Research Policy, 36*, 107-120.

Bishop, P., & Gripaios, P. (2007). Explaining spatial patterns of industrial diversity: An analysis of sub-regions in Great Britain. *Urban Studies, 44*(9), 1739–1757.

Bosma, N., Stam, E., & Schutjens, V. (2011). Creative destruction and regional productivity growth: Evidence from the Dutch manufacturing and services industries. *Small Business Economics, 36*(4), 401-418.

Bottazzi, L., & Peri, G. (2003). Innovation and spillovers in regions: Evidence from European patent data. *European Economic Review, 47*(4), 687-710.

Box, M. (2008). The death of firms: Exploring the effects of environment and birth cohort on firm survival in Sweden. *Small Business Economics, 31*, 379–393.

Box-Steffensmeier, J. M., & Jones, B. S. (2007). *Event history modeling. A guide for social scientists*. New York: Cambridge University Press.

Breschi, S., & Lissoni, F. (2001). Knowledge spillovers and local innovation systems: a critical survey. *Industrial and Corporate Change, 10*(4), 975-1005.

Bridges, S., & Guariglia, A. (2008). Financial constraints, global engagement, and firm survival In the United Kingdom: evidence from micro data. *Scottish Journal of Political Economy, 55*(4), 444-464.

Brixy, U., & Grotz, R. (2007). Regional patterns and determinants of birth and survival of new firms in Western Germany. *Entrepreneurship & Regional Development: An International Journal, 19*(4), 293 – 312.

Buss, T. F., & Lin, X. (1990). Business survival in rural America: A three state study. *Growth and Change, 21*(3), 1-8.

Cassia, L., Colombelli, A., & Paleari, S. (2009). Firms' growth: Does the innovation system matter? *Structural Change and Economic Dynamics, 20*(3), 211-220.

Cefis, E., & Marsili, O. (2005). A matter of life and death: innovation and firm survival. *Industrial and Corporate Change, 14*(6), 1-26.

Cefis, E., & Marsili, O. (2006). Survivor: the role of innovation in firm's survival. *Research Policy, 35*(5), 626-641.

Chung, C. C., Lu, J. W., & Beamish, P. W. (2008). Multinational networks during times of economic crisis versus stability. *Management International Review, 48*, 279-295.

Cleves, M., Gutierrez, R., Gould, W., & Marchenko, Y. (2010). *An introduction to survival analysis using Stata*. College Station, TX: Stata Press.

Delios, A., & Beamish, P. W. (2001). Survival and profitability: The roles of experience and intangible assets in foreign subsidiary performance. *The Academy of Management Journal, 44*(5), 1028-1038.

Dodgson, M. (2011). Exploring new combinations in innovation and entrepreneurship: Social networks, Schumpeter, and the case of Josiah Wedgwood (1730-1795). *Industrial and Corporate Change, 20*(4), 1119-1151.

Esteve-Perez, S., Sanchis-Llopis, A., & Sanchis-Llopis, J. A. (2004). The determinants of survival of spanish manufacturing firms. *Review of Industrial Organization, 25*, 251–273.

Esteve-Perez, S., Sanchis-Llopis, A., & sanchis-Llopis, J. A. (2010). A competing risks analysis of firms' exit. *Empirical Economics, 38*(2), 281-304.

Feldman, M. P., & Audretsch, D. (1999). Innovation in cities:: Science-based diversity, specialization and localized competition. *European Economic Review, 43*(2), 409-429.

Feser, E. J. (2002). Tracing the sources of local external economies. *Urban Studies, 39*(13), 2485-2506.

Fotopoulos, G., & Louri, H. (2000). Location and survival of new entry. *Small Business Economics, 14*, 311-321.

Fritsch, M., Brixy, U., & Falck, O. (2006). The effect of industry, region, and time on new business survival – a multi-dimensional analysis. *Review of Industrial Organization, 28*, 285–306.

Funke, M., & Niebuhr, A. (2005). Regional geographic research and development spillovers and economic growth: Evidence from West Germany. *Regional Studies, 39*(1), 143-153.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (2002). Growth in cities. *Journal of Political Economy, 100*(1), 125-152.

Gordon, I., & McCann, P. (2000). Industrial clusters: Complexes, agglomeration and/or social networks? *Urban Studies, 37*(3), 513-532.

Griliches, Z. (1979). Issues in assessing the contribution of R&D 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*(4), 1661-1707.

Griliches, Z. (1992). The search for R&D spillovers. *Scandinavian Journal of Economics, 94*(0), S29-S47.

Henderson, J. V., Shalizi, Z., & Venables, A. J. (2001). Geography and development. *Journal of Economic Geography, 1*(1), 81-105.

Ibrahim, S. E., Fallah, M. H., & Reilly, R. R. (2009). Localized sources of knowledge and the effect of knowledge spillovers: an empirical study of inventors in the telecommunications industry. *Journal of Economic Geography, 2009*, 405–431.

Jacobs, J. (1969). *The economy of cities*. New York: Random House.

Klette, T. J., & Kortum, S. (2004). Innovating firms and aggregate innovation. *Journal of Political Economy, 112*(5), 986-1018.

Knoben, J., & Oerlemans, L. A. G. (2006). Proximity and inter-organizational collaboration: a literature review. *International Journal of Management Reviews, 8*(2), 71-89.

Koo, J. (2005). Technology spillovers, agglomeration, and regional economic development. *Journal of Planning Literature, 20*, 99-115

Krugman, P. (1991). *Geography and trade*. Cambridge: MIT Press.

Lehto, E. (2007). Regional Impact of Research and Development on Productivity. *Regional Studies, 41*(5), 623 - 638

Lin, P.-C., & Huang, D.-S. (2008). Technological regimes and firm survival: evidence across sectors and over time. *Small Business Economics, 30*(2), 175-186.

López-Bazo, E., Vayá, E., & Artís, M. (2004). Regional externalities and growth: evidence from European regions. *Journal of Regional Science, 44*(1), 43-73.

Mata, J., Portugal, P., & Guimaraes, P. (1995). The survival of new plants: Start-up conditions and post-entry evolution. *International Journal of Industrial Organization, 13*(4), 459-481.

Pakes, A., & Griliches, Z. (1980). Patents and R&D at the firm level: A first report. *Economics Letters, 5*(4), 377-381.

Persson, H. (2004). The survival and growth of new establishments in Sweden, 1987-1995. *Small Business Economics, 23*(5), 423-440.

Porter, M. (1990). *The competitive advantage of nations*. New York: Free Press.

Renski, H. (2009). New firm entry, survival, and growth in the United States: A comparison of urban, suburban, and rural areas. *Journal of the American Planning Association, 75*(1), 60–77.

Renski, H. (2011). External economies of localization, urbanization and industrial diversity and new firm survival. *Papers in Regional Science, 90*(3), 473-502.

Rodriguez-Pose, A., & Comptour, F. (2012). Do clusters generate greater innovation and growth? An analyisis of european regions. *The Professional Geographer, 64*(2), 211-213-231.

Rodriguez-Pose, A., & Crescenzi, R. (2008). Research and development, spillovers, innovation systems, and the genesis of regional growth in Europe. *Regional Studies, 42*(1), 51-67.

Schmutzler, A. (1999). The new economic geography. *Journal of Economic Surveys, 13*(4), 355-379.

Schumpeter, J. (1934). *The Theory of Economic Development*. New York: McGraw Hill.

Scott, A. (2006). Entrepreneurship, innovation and industrial development: Geography and the creative field revisited. *Small Business Economics, 26*(1), 1-24.

Segarra, A., & Callejón, M. (2002). New firms' survival and market turbulence: new evidence from Spain. *Review of Industrial Organization, 20*(1), 1-14.

Sorenson, O., & Audia, P. (2000). The social structure of entrepreneurial activity: Geographic concentration of footwear production in the United States, 1940–1989. *The American Journal of Sociology, 106*(2), 424-461.

Stel, A. J. v., & Nieuwenhuijsen, H. R. (2004). Knowledge spillovers and economic growth: an analysis using data of Dutch regions in the period 1987-1995. *Regional Studies, 38*(4), 393-407.

Stough, R., & Nijkamp, P. (2009). Knowledge spillovers, entrepreneurship and economic development. *The Annals of Regional Science, 43*(4), 835-838.

Stuart, T., & Sorenson, O. (2003). Geography of opportunity: Spatial heterogeneity in founding rates and the performance of biotechnology firms. *Research Policy, 32*(2), 229-253.

Tödtling, F., & Wansenböck, H. (2003). Regional differences in structural characteristics of start-ups. *Entrepreneurship and Regional Development, 15*, 351–370.

Uyarra, E. (2010). What is evolutionary about 'regional systems of innovation'? Implications for regional policy. *Journal of Evolutionary Economics, 20*(1), 115-137.

Wang, Z., Ma, C., Weng, G., & Wang, Y. (2004). A study on temporal and regional process of knowledge spillover. *The Annals of Regional Science, 38*(4), 595-606.

Wennberg, K., & Lindqvist, G. (2010). The effect of clusters on the survival and performance of new firms. *Small Business Economics, 34*(4), 221-241.

Zachariadis, M. (2003). R&D, innovation, and technological progress: a test of the Schumpeterian framework without scale effects. *Canadian Journal of Economics, 36*(3), 566-586.

Appendix 1

Average patent counts before (Table 1A) and after (Table 2A) adjustment

Table 1A

Year	Number of observations	Mean	Standard Deviation	Min	Max
1992	366	158.52	442.45	0.00	4627.98
1993	366	167.9	465.25	0.33	4714.02
1994	366	190.4	530.99	0.25	5528.93
1995	366	226.68	647.28	0.33	6388.23
1996	366	218.03	626.58	0.00	5934.29
1997	366	254.76	756.7	0.00	7195.75
1998	366	253.91	746.13	0.00	6984.04
1999	366	268.53	799.21	0.00	7655.39
2000	366	285.91	852.94	1.50	8572.2
2001	366	288.78	866.98	0.00	8943.53
2002	366	289.27	877.8	0.00	9173.12
2003	366	279.8	830.05	0.00	8479.73
2004	366	263.23	801.52	0.00	8394.73
2005	366	248.75	763.73	0.00	7989.95
2006	366	219.36	669.01	0.00	6978.2
2007	366	180.43	543.13	0.00	5563.03
2008	366	112.21	332.46	0.00	2992.39

Table 2A, only 2005 - 2008

Year	Number of observations	Mean	Standard Deviation	Min	Max
2005	366	248.75	763.73	0.00	7989.95
2006	*366*	*231.48*	*703.86*	*0.02*	*7304.53*
2007	*366*	*204.68*	*613.13*	*0.30*	*6215.69*
2008	*366*	*148.58*	*437.6*	*0.51*	*3971.38*