THE IMPACT OF EXTERNAL KNOWLEDGE ON PATENT PRODUCTION AMONG SBIR WINNERS

A thesis

Presented to

The Faculty of the Graduate School

At the University of Missouri

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science

By

Dante Pozzi

Nicholas Kalaitzandonakes, Thesis Supervisor

MAY 2011

The undersigned, appointed by the dean of the Graduate School,

have examined the thesis entitled

THE IMPACT OF EXTERNAL KNOWLEDGE ON PATENT PRODUCTION
AMONG SBIR WINNERS

presented by Dante Pozzi,

a candidate for the degree of Master of Science,

and hereby certify that, in their opinion, it is worthy of acceptance.

____

Professor Nicholas Kalaitzandonakes

____

Professor Yin Xia

____

Professor Xinghe Wang

ACKNOWLEDGEMENTS

I would like to express my gratitude to my thesis supervisor, MSMC Endowed Professor of Agribusiness and Director of the Economics and Management of Agrobiotechnology Center (EMAC), Nicholas Kalaitzandonakes, Ph.D. His knowledge, encouragement, and guidance have been of great value to my thesis and professional development. I am also grateful to Christos Kolympiris, Ph.D., for his detailed and constructive comments and for his important support throughout this work. I would also like to thank Lauren Jackson for helping me edit and submit all thesis materials.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Figure Page

LIST OF TABLES

LIST OF ABBREVIATIONS

AIC	Akaike Information Criterion
CIS	Finnish Community Innovation Survey
CPI	Consumer Price Index
CRENoS	Center for North South Economic Research
CSO	Central Statistical Office
EPO-CESPRI	European Patent Office, Centre for Research on Innovation and Internationalisation Processes
FAME	Financial Analysis Made Easy
GEE	Generalized Estimating Equation
ISIC	International Standard Industrial Classification
ISTAT	Instituto Nazionale de Statistica
LLS	Local Labour Systems
LSFs	Life Science Firms
ML	Maximum Likelihood
MSA	Metropolitan Statistical Area
NIH	National Institute of Health
NSF	National Science Foundation
NUTS	Nomenclature of Units for Territorial Statistics
OECD	Organisation of Economic Cooperation and Development
OLS	Ordinary Least Square
PHID	Pharmaceutical Industry Database
QIC	Quasi-Likelihood Value
R&D	Research and Development
SBIDB	Small Business Innovation Database
SBIR	Small Business Innovation Research

SDC	Securities Data Company
SEC	Securities Exchange Committee
SMSA	Standard Metropolitan Statistical Area
STAN	Structural Analysis dataset
USPTO	United States Patent Office
VC	Venture Capital

ABSTRACT

This paper provides empirical evidence related to the factors which allow SBIR winner firms to be more innovative. The analysis is carried out by investigating the influence of firm-specific characteristics (such as research intensity and the scope of external knowledge sources), together with regional-specific characteristics (such as the spatial proximity to successful firms, the amount of research dollars awarded to universities, and a location quotient based on the number of biochemist and biophysicist employees) on the cumulative number of patents awarded by SBIR to life science firms from 1983 to 2006. Based on a dataset of SBIR winners that operate in the life science field, the data were analyzed under a fitted Poisson count data model with an estimated marginal effect for each of the continuous explanatory variables. We find evidence that larger and older firms are more prolific in generating more patents. We also find evidence that firms that have received venture capital funds produce more than three times the number of patents, as opposed to SBIR awards, which did not show strong explanatory power. Additionally, the idea of positive effects arising from spatial collocation with successful firms is supported from the significance influence that the number of firms located within a 20-mile radius from a given firm has on patent production. Finally, the econometric analysis also shows that firms that employ knowledge bases close to their core science are more abundant in producing patents, which highlights the economic significance of the type of external knowledge sourced by a particular firm.

CHAPTER ONE

INTRODUCTION

Federal and state governments around the world often seek to support innovation and increase national competitiveness through programs that help firms in high-technology areas to overcome what is often called the "Valley of Death" (Wessner, 2005) which refers to the limited financing options that such firms have during the initial stages of their research. The largest innovation program in the United States is the Small Business Innovation Research (SBIR) program, under which 11 federal agencies use a two-phase process to allocate 2.5% of their annual budget to fund private firms' projects with commercial potential. Phase 1 grants are used to explore the scientific and commercial feasibility of an idea/technology and typically do not exceed $100,000; Phase 2 grants are considerably larger and are given to the most meriting Phase 1 winners.[1] SBIR grants are the prominent source of seed funding in the United States since they are mainly used to finance early-stage projects that private investors rarely include in their investment portfolios (Wessner, 2002). As a case in point, in 2007 the SBIR program awarded $2.3 billion for research and seed funding, while the corresponding figure for private-venture markets was $1.2 billion (Wessner, 2009b).

The explicit goals of the program include stimulation of technological innovation and increase of private-sector commercialization of innovations (Wessner, 2009a). With regard to commercialization efforts, researchers have documented that the SBIR program leads to increased commercialization of innovations (Audretsch, Link, & Scott, 2002;

[1] There is also a Phase 3, but federal agencies do not provide funds for these projects.

Cooper, 2003), fosters entrepreneurship (Audretsch, Weigand, & Weigand, 2002; Elston & Audretsch, 2011), and—under certain conditions—attracts funds from private markets (Lerner, 1999). On the other hand, some scientists see the contribution of the SBIR program in stimulating innovation with skepticism (Basu, 2004). To add to this skepticism, academic empirical research has not focused on innovative (compared to commercialized) outcomes that emanate from the SBIR program. This lack of empirical evidence is rather surprising when considering the significance of the SBIR program, its stated goal of boosting innovation,[2] and the established record of the most innovative SBIR winners such as Apple Computer, Chiron, Compaq, and Intel (Audretsch, 2003) in bringing welfare-enhancing products to the market. Against this background, assessing which selected factors allow SBIR winners to be more innovative can inform the grant-awarding process of federal agencies, assist the SBIR program to meet its goals, and potentially help SBIR agencies to identify new innovative firms beforehand. Accordingly, in this paper we study which features have helped SBIR winners to become more innovative.

Following previous literature (e.g., Autant-Bernard, 2001; Boix & Galletto, 2009), we use patents as a measure of innovation.[3] In order to empirically analyze the patenting rate of firms in the SBIR program, we employ a novel dataset of SBIR winners that operate in the life sciences field, which is arguably among the most vibrant high-technology industries in the United States and elsewhere. In the empirical part, we

[2] Link and Ruhm (2009) note the general lack of econometric analysis of the SBIR program related to its objectives.

[3] Despite their wide use, patents have certain shortcomings as a measure of innovation. For instance, innovative firms may not patent for strategic reasons (Teece, 1986), and some patent citations are added by the patent office (Alcacer & Gittelman, 2006). Albeit a less-than-perfect proxy for innovation, patents are still a reliable measure of innovation (Acs, Anselin, & Varga, 2002).

associate the number of patents produced by SBIR Life Science Firms (LSFs) with firm-specific and regional characteristics. In the firm-specific characteristics, this is among the few studies to include in the analysis the effects of the scope of external knowledge sources a firm uses, which is a largely overlooked factor that needs to complement standard determinants of a firm's performance (Laursen & Salter, 2006). With regard to regional characteristics, we incorporate in the analysis how the presence of successful firms in a narrow radius can influence the number of patents for a particular firm through knowledge spillovers, network externalities, and other channels. In line with previous research (e.g., Wallsten, 2001), the use of modern software (ArcView) affords us the benefit of measuring the impact of firms located in economically relevant spatial units, which do not necessarily coincide with zip code, city, or state boundaries.

The remainder of the paper is organized as follows: the next chapter discusses the theoretical framework and develops the research hypotheses of our study. Chapter 3 specifies the econometric model used to test our research hypotheses and Chapter 4 describes our dataset. Chapter 5 presents the estimation results and concluding comments are in Chapter 6.

CHAPTER TWO

LITERATURE REVIEW AND THEORETICAL EXPECTATIONS

We look for guidance in the extant literature in order to form the theoretical expectations of the present study. In particular, we focus on regional characteristics and firm-specific features that are expected to affect the rate that a firm innovates as measured by its patent production patterns.

FIRM CHARACTERISTICS THAT AFFECT A FIRM'S PATENT PRODUCTION

Schumpeter's work (1942) had arguably the most influence in pointing towards a positive relationship between the size of a given firm and the rate that it innovates. Cohen and Levin (1989) outline that large firms can be more innovative because they may benefit from economies of scale and scope, realize benefits from synergies and spillovers among different departments, and be in a favorable position to attract funds from private markets because they are typically seen as less risky than their smaller counterparts. In this vein, Hall and Ziedonis (2001) narrowed the economies-of-scale advantage to patent production to the fixed costs related to maintaining legal departments that deal with intellectual properties issues. In general, the empirical work that has tested the hypothesized relationship between firm size and patenting patterns has supported the theoretical expectations (Peeters & van Pottelsberghe de la Potterie, 2006).[4] For instance, Beaudry and Breschi (2003) studied the patenting patterns of firms in Italy and in the

[4] Cohen (1995) reviews studies on the topic and identifies some research that has not found a positive relationship between firm size and patent production.

United Kingdom and found larger firms to be more prolific in producing patents. Under different research samples that reflected different industries and time periods, Peeters and van Pottelsberghe de la Potterie (2006), Hall and Ziedonis (2001), and Lunn (1987) also reached similar conclusions. Against this background, we develop the following hypothesis.

Hypothesis 1: Larger firms patent more.

Firms with increased research and development (R&D) expenditures are also expected to produce more patents, and extensive empirical literature (e.g., Blundell, Griffith, & Windmeijer, 2002; Gurmu & Pérez-Sebastián, 2008; Hall & Ziedonis, 2001; Montalvo, 1997) has largely supported the notion in question. It is argued that the first research to investigate the relationship between R&D expenditures and patent applications came from Pakes and Griliches (1980). Pakes and Griliches obtained patent data from the Office of Technology Assessments and Forecasts of the US Patent Office, and company data from Compustat (from 10-K firm Securities Exchange Commission [SEC] reports to the SEC), comprising a total of 121 firms from 1968 to 1975; these were divided in two groups—one with firms in research-intensive industries and another with all other manufacturing firms. The model was specified in a log-log functional form with a firm’s number of patents as the dependent variables and firm age, research expenditures (including lagged research expenditures), and firm-specific industries as dummy variables comprising a set of explanatory variables; results show that much of the patent variance is associated with the variation of research expenditures. Hausman, Hall, and Griliches (1984) extend the Pakes and Griliches analysis of patents and R&D

expenditures through a Poisson specification model using a firm's R&D expenditures, five lagged values of R&D, and a time trend as explanatory variables for number of patents. The sample used included patents from 128 firms for the years 1968-1974 (from the Office of Technology Assessment and Forecasting of the US Patent Office) and R&D expenditure figures extracted from Compustat tape and other sources, hence in the same fashion as Pakes and Griliches (1980). The results of this Poisson model were in line with the Ordinary Least Square (OLS) specification, as the coefficient for R&D expenditures shows a positive and significant estimate. The results, however, also show a U-shape distribution for the lag coefficients. Subsequently, firm-specific variables, including dummy variables for the different scientific sectors (drug, computer, scientific instruments, chemical, and electronic equipment industries) as well as an inflation-adjusted book value of the firm in 1971, were included as explanatory variables; these showed positive and significant effects on the expected number of patents, with a decreasing effect over the R&D estimated coefficient. A set of specification models representing "different ways of adding randomness" are also run, all of which suggest that the introduction of firm-specific variables (in this case, proxies for size and field of work) reduce the importance of R&D expenditure over number of patents. All in all, the rationale behind the relationship between R&D and patents is straightforward: almost by definition, R&D expenditures increase the knowledge base of a given firm, which in turn can result in increased innovative measures such as patents. Albeit not all R&D is necessarily successful in increasing the knowledge base of a particular firm and not all innovations are patented, the relationship between R&D and patents is expected to remain. Against this background, we develop the following hypothesis.

Hypothesis 2: Firms with increased R&D resources patent more.

To complement the aforementioned traditional determinants of patent production, the scientific range that a particular firm uses to source knowledge can also affect the innovative outcomes of the firm. However, whether a broad or a narrow knowledge source is more conducive to innovation is not clear *a priori*. On one hand, firms whose research borrows from different disciplines may be more innovative because they can combine complementary knowledge (Leiponen, 2005), and often innovation emanates from such complementarities (Kogut & Zander, 1992). Further, firms with expanded knowledge sources can enhance their knowledge pool and accordingly employ a larger array of options in order to innovate (March, 1991). On the other hand, firms that use specialized knowledge can be more innovative because drawing knowledge from a narrow set of sources increases familiarity with the knowledge in question, which then can reduce the likelihood of errors, minimize unnecessary steps (Eisenhardt & Tabrizi, 1995) and search costs, increase the reliability of ideas, and facilitate the development of routines (Levinthal & March, 1993). In addition, knowledge from specialized sources may be conducive to innovation because repeated use of familiar knowledge can lead to a deeper understanding of concepts, ideas, and techniques that are difficult to grasp by less exposed users (Katila & Ahuja, 2002). The theoretical contradiction between the relative importance of broad versus deep knowledge sources is also reflected in empirical studies on the issue. For instance, in studies that addressed a firm's performance other than innovative outcomes, the findings of Leiponen and Helfat (2010) and Katila (2002) point toward knowledge breadth as a key determinant of a firm's success. Particularly on the former, Leiponen and Helfat use data collected by the Finnish Community Innovation

Survey (CIS) in 1997 from Finnish manufacturing firms with more than 100 employees to test the impact of innovation breadth in both innovation objectives and knowledge sources. The dependent variables used were: (a) a binary variable for firm's introduction of innovation during the 1994-1996 and (b) a percentage of total firm sales in 1996 originated from the sale of technologies introduced in the same time period. The explanatory variables for this examination included measures of breadth for objectives and sources obtained from the CIS survey questions plus other controlling variables that could also have an impact on innovation success (firm size, R&D spending, firm innovative capabilities, and industry operations). Empirical results support the study's main proposition, i.e., that a greater breadth of knowledge is positively associated to innovation success. Finally, results did not support an additional hypothesis associated with the diminishing returns of knowledge source breadth.

Also supporting this view regarding a positive impact of knowledge breadth, Katila (2002) investigates to what extent the average age and the age diversity of searched knowledge affect innovation production, as it is proposed that the impact of age on innovation will depend on where firms will search knowledge, either internal, intraindustry (among competitors), or extraindustry (among firms from different industries) space. The paper suggests that the higher the number of knowledge elements searched, the higher the chances to have new products (i.e., innovation). Moreover, it is also proposed that old external knowledge will leverage the number of searched knowledge and vice-versa. By using data from 1985 to 1997 from public robotics companies in Europe, Japan, and the United States collected from a variety of sources including databases, trade magazines, and primary research with industry experts, a final sample of 131 firms representing 1,255 firm-years was formed. Regression analysis on the effects of search age were performed using the number of new products as a dependent variable, where ‘new products’ is defined as one showing at least some change in design. Age variables were measured as the age of the cited patents for a given patent for which a firm has applied during a given year. More specifically, the sum of the external cited patents in a year was used to measure the quantity of external search. Regression estimates confirmed this mutual leverage hypothesis between external search age and external search quantity as well as the positive effects on innovation production.

Opposing the empirical evidences favoring a broad knowledge base, Bessen (2008) and Magazzini, Pammolli, Riccaboni, and Rossi (2009) report results which suggest that more narrow knowledge is generally more useful. In Magazzini et al. (2009), the issue regarding knowledge source breadth is examined under the angle of patent disclosure and the related impact on R&D dynamics and, ultimately, on competition among pharmaceutical firms. Based on two large datasets—one comprising all pharmaceutical and biotechnology patents granted by the USPTO office since 1965 and a second one containing data from specific R&D projects drawn from the Pharmaceutical Industry Database (PHID), all together resulting in a nearly 2,000 drug-candidate patent pairs also classified based on the R&D project final outcome—the authors develop a regression framework where the success of a given project (1 if marketed, 0 if discontinued) is used as a dependent variable. The independent variables attempted to address the research on which the patent was built, including the measure of originality index,[5] which works as a proxy for the "breadth of the technology roots of the innovation," following Trajtenberg, Henderson, and Jaffe (1997). The estimated IV-Probit coefficient of originality, however, showed no significant association with R&D success,[6] suggesting that breadth is not significantly correlated to R&D success.

The proposition of a more useful narrow-base technology finds support also in Bessen (2008) and his investigation of the relationship between patent value and a variety of patent characteristics such as patent citations statistics, type of technology, and patent assignee. Additionally, the study also includes a patent's generality and originality measures—as suggested by Trajtenberg et al. (1997)—as explanatory variables to patent value. As for the value of the patent itself (dependent variable), Bessen uses the amount of the renewal fee paid by the owner as a proxy for patent value. The study uses all US utility patents issued in 1991, excluding patents assigned to the governments and foreign individuals. Particularly for the originality variable, the estimates deriving from the regression model indicate a significant, but rather negative, coefficient, suggesting that patent values are significantly and positively affected by a more narrow technology base.

Katila and Ahuja (2002) allege that both knowledge breadth and depth are instrumental in at least considering the number of new products a firm is able to bring to the market. In this study, a two-dimension problem-solving framework is developed; the first assesses a firm's rents in reusing their existing knowledge (search depth) and the second assesses a firm's rents in exploring new knowledge (search scope). It is then

[5] In Magazzini et al. (2009), explanatory variables also included the share of self citation, share of references to non-patent literature, and average time lag between the citing and the cited patents, all of which was built based on Trajtenberg et al. (1997).

[6] The proxy for reference to non-patent literature was the only positive and significant association to R&D success probability.

hypothesized that the way a firm has pursued these two dimensions will affect the firm's ability to develop new products. Particularly on search depth, it is proposed that it positively affects innovation production by reducing the odds of false starts and facilitating routines development, hence increasing the predictability of searches due to the higher familiarity of products requirement and increasing a firm's ability to identify valuable knowledge information (as in-depth understanding emerges from repeated usage). On the other hand, an excessive search depth is proposed to negatively affect innovation due to improvement limitations deriving from diminishing returns over the same knowledge and organization rigidness. From the opposing effects of search depth over time, a curvilinear (inverted U-shape) relationship between search depth and the introduction of new products is hypothesized. Similarly, positive effects of search scope include new choices to solve problems and new knowledge recombination, as new elements added to the set improves the possibilities for new products. Negative effects from high levels of search scope include the increasing integration costs associated with a growing number of new technologies (both technological and organizational) and the decreasing ability to truly absorb and respond to new knowledge accurately. Furthermore, both variables, scope, and depth are proposed to be mutually beneficial, as the existing knowledge enhances a firm's absorption capacity to capture new external knowledge and increases the uniqueness of recombination. In order to test such propositions, Katila and Ahuja (2002) used a sample of 124 firms from a population of industrial robotics companies in Japan, Europe, and North America, averaging 39,000 employees each. The main data resources included new product introduction (the dependent variable) and patent data, from which the firm's patent citation constitutes the backbone for the main independent variables examined; *search depth* was measured as the average number of times a firm repeatedly used the citations from their own patents, while *search scope* was measured as the proportion of previously unused citations. Control variables include the number of a sample firm's factory automation collaborations (as a proxy for collaboration frequency), return over assets (as a proxy of firm performance), R&D expenditure figures (as a proxy of firm's total R&D inputs to the innovation process), and corporate employees (as a proxy for firm size). Dummy variables for nationality (accounting for national technological characteristics) and year dummies between 1985-96 (accounting for market conditions and general economic environment) were also included in the analysis. A GEE Poisson specification was used in order to incorporate the zero values of the dependent variable (Poisson) while accounting for autocorrelation (GEE—Generalized Estimating Equation). The estimated parameters from the regression model confirm the inverted U-shaped relationship between search scope and new products introductions as well as the proposed positive mutual leverage between depth and scope. In most cases, the control variable effects are also confirmed, with R&D expenditure and collaboration frequency increasing the number of robotics production. Finally, a linear relationship between search depth and new products was observed rather than the proposed curvilinear one.

All in all, both theoretical arguments and empirical evidence support the importance of a firm's range of knowledge sources in boosting innovation but provide conflicting predictions and results with regard to the relative significance of a narrow versus a broad knowledge base. Against this background, we propose the following hypotheses.

Hypothesis 3: Firms that employ narrow external knowledge sources patent more.

Hypothesis 4: Firms that employ broad external knowledge sources patent more.

REGIONAL CHARACTERISTICS THAT AFFECT A FIRM'S PATENT PRODUCTION

Besides firm-specific features, characteristics of the regional environment can also affect the pace of innovation achieved by a given firm. A well developed body of literature suggests that positive externalities emanate from the spatial collocation of similar firms (Anselin, Varga, & Acs, 1997; Bottazzi & Peri, 2003; Delaney, 1993; Funke & Niebuhr, 2005; Keller, 2002; Orlando, 2004; Rosenthal & Strange, 2003; Varga, 2000; Wallsten, 2001) particularly in knowledge-based industries which often rely on tacit knowledge. For instance, researchers have reported that collocation of similar firms and other industry participants can create efficiencies in knowledge creation (Coenen, Moodysoon, & Asheim, 2004; Gittelman, 2007; McKelvey, Alm, & Riccaboni, 2003; Moodysson & Jonsson, 2007) which then can lead to increased innovative outcomes. The main argument behind these studies is that geographic proximity assists the transmission of tacit knowledge (Adams & Jaffe, 1996; Aldrich & Wiedenmayer, 1993; Feldman, 1999; Fontes, 2005; Jaffe et al., 2005; Meyer & Rowan, 1977; Thornton, 1999; Zander & Kogut, 1995)[7] because it increases frequent contacts which are considered a rich conduit of knowledge transfer.

[7] There are also contributions to the literature that raise doubts about the importance of tacit knowledge and positive spatial externalities (see Breschi & Lissoni, 2001; Håkanson, 2005).

Feldman (1999) offers a detailed review of landmark empirical studies on the matter of location and innovation. The paper organizes these past researches on two groups. A first one is based on the logic of production function, which seeks to quantify the impact of knowledge spillovers on innovation where explanatory variables from a common geography platform are set against a measure of novelty set as the dependent variable. A second group a second tradition focuses on explaining economic outcomes such as growth or productivity from a geographic perspective by using different factors as explanatory variables, including innovation. Regarding the first stream of research Feldman uses Griliches (1992) definition of knowledge spillover to review four common approaches on the study of knowledge spillover: the innovation production functions; the linkages between patent citations or "paper trails;" the mobility of skilled labor or "start scientists;" and finally knowledge spillover embodied in goods. On the second tradition, a list of three aspects affecting the location on innovation is suggested: the existence of agglomeration externalities, either localization or urbanization economies; characteristics of knowledge; and characteristics of firms. Particularly on the attributes of knowledge, the review argues that it varies on the degree of tacitness, which is related to the uncertain and interpretative features of knowledge (as opposed to the standardized, codified degree of knowledge that can be easily transmitted trough long distances). Hence, it is postulated that the more tacit the nature of a given knowledge, the more it will demand a face to face interaction and ultimately a geographic proximity.

In that vein, Jaffe, Trajtenberg, and Henderson (1993) uses patent citation data in order to investigate the extent of which knowledge spillovers are geographically localized. The research is based on six distinct set of patents, the "originating patents" which included 316 patents granted to universities in 1975 and 482 patents granted to universities in 1985 and four matching samples for "top corporate" (patents granted to the 200 US firms with the greatest R&D spending in 1986) and "other corporate" (drawn from the universe of all other patents assigned to US corporations). All samples are matched in terms of patent classes, hence controlling for variations in citation practices across technological areas, and application year, so that all patents had the same amount of time to be cited. To the extent of geographic localization, patents were assigned to countries, states and MSAs in order to properly address the question of how often the citing patent has the same localization as the originating patent, ounce other sources of agglomeration effects from knowledge spillover were controlled. The regression specification uses a Probit model where the geographic dummy (match/no match) is set as the dependent variable. The explanatory variables included the log of the citation lag, dummy variables for top corporate and other corporate as well as variables related to generality[8] and the fraction of the originating patents that was self-cites. The empirical results indicate that, particularly for citations of 1980, for all datasets (university, top corporate and other corporations) the estimated coefficients are strongly and significantly more localized than controls (even more than the 1975 dataset). Hence, there is a pattern of localization in patents from 1980 is observed in 1989 at the country, state and SMSA level, meaning that citations are 5 to 10 times more likely to come from the same SMSA as the control patents (still 2 to 6 times if self-citations are excluded). Along these findings, Audretsch and Feldman (1996) focus on the relationship between the propensity of innovation activity to cluster and different stages of industry cycle. Based on a dataset

[8] This variable refers to one minus the Herfindhal index across patent classes of the citations received.

of 4.200 commercial innovations from SBIDB (Small Business Innovation Database), Audrestch suggests that tacit knowledge plays a more important role in the early stage rather than mature stage of the industry since, in the later, the standardization of economics aspects would reduce the importance to innovation activity. Based on a geographically bounded nature of tacit knowledge, it is argued that early stage industries would tend to cluster more during mature/decline stages. Different from formal knowledge, which, due to its structured and lucid nature can be codified into formulas and therefore transferred to long distances at low cost, the tacit knowledge, being far more ambiguous and less structured, will ask for a face-to-face contact in order to avoid errors of interpretations. The transmition of tacit knowledge therefore, is facilitated by social bonds provided by geographical proximity in order to overcome the ambiguity and difficulty of "write down in such a way that is meaningful and readily understood" (Teece, 2005).[9] The examination of knowledge transmit from a geographical point of view finds support in Zander and Kogut (1995) as they analyzed to what extent the certain characteristics of manufacturing capabilities can affect the time of transfers and imitation. Based on a questionnaire sent to project engineers regarding 44 out of the top 100 major Swedish innovations, the paper divides the dataset in two parts, a first part addressing mainly factual data (such as the date at which the innovation was first introduced, timing for transfers and first imitations) and a second part addressing the firm's manufacturing of innovation. The questionnaire is based on 43 questions arranged on a seven-item scale where five constructs are used to characterized firm's knowledge; "Codifiability" (if knowledge can be encoded) "Teachability" (if it can be trained)

[9] The importance of social networks and spatial proximity finds support in several studies in Silicon Valley (Saxenian, 1994), Cambridge (Garnsey & Cannon-Brookes, 1993), and Italy (Lazerson, 1995).

“Complexity” (if it combines different kinds of competencies) “System Dependence” (if it depends on many different people) and “Product Observability” (if competitors can copy the manufacturing capability). All five constructs therefore access the complexity of knowledge, implying that the more complex is a given knowledge the harder is to transfer or imitate. Two regression models, one with hazard of transfer as a dependent variable, and another with hazard of imitation as a dependent variable are estimated using the five constructs as explanatory variables (with the construction of measures deriving from questionnaire items). The paper shows empirical evidence that codifiability and teachability are positive and significant on predicting hazard of transfer; hence suggesting that the more tacit and difficult to communicate are a capability the harder it is to transfer.

In addition to knowledge spillovers, a rich regional environment can also help local firms to join professional networks that can bring about numerous benefits to economic actors (Huggins & Johnston, 2010) such as the integration of diverse knowledge bases (Dahlander & McKelvey, 2005; Liebeskind, Oliver, Zucker, & Brewer, 1996). In that vein, Sorenson and Stuart (2001) conjecture that economic actors often join localized social networks and develop local professional relationships in order to overcome information asymmetries. By analyzing the SDC’s Venture Economics database from 1986 to 1998 of venture capital investments, including 80,406 investment cases involving 1,025 venture capital firms and 7,590 target companies, Sorenson and Stuart (2001) investigate the probability of a particular venture capital firm to invest in a given firm. The regression model is based on a Logit specification and uses as independent variables data that refers to the dimensions that delimits the flow of information among investors and targets such as; geographic distance (from the venture capital main office and the location of the target), industry distance (proxy by the similarity between the venture capital prior investments and the target firm industry) and the venture capital network position (proxy by measures of mean affiliation, affiliate distance and centrality).The results support the proposal that the likelihood that a venture capital invests in a target reduces as the geographic and industry distance between them increases. Going further, the paper also suggests that the probability of an investment in a far off target increases under strong network position, if the members of previously formed syndicate are also financing the target. Finally, the paper suggests that the probability of an investment decreases as the closest member of the syndicate declines the investment. Along this line, Oliver (1994) argues that network transactions for resources are crucial in the early phase of a firm in order to guarantee survival while a reduction of external alliances is important in a more mature phase when firms look to protect their boundaries. Also Cooke (2002), based in data from biotechnology industry in UK, Germany, and United States, argues that proximity to investors, lawyers and hospital research facilities for clinical tries are crucial for the creation of sectoral innovation systems. Finally Boschma (2005) outlines that it is common to stress the importance of networks as vehicles of knowledge creation and diffusion.

Finally, Huggins and Johnston (2010) provide additional empirical support to the proposition that firms often use local networks to source knowledge. Based on a postal survey of knowledge based firms from sectors such as IT, Computer Technology, and Telecom, Business and Financial Services, Media and Broadcasting from three different regions in England (Yorkshire and Humberside, North East England, and North West

England), the research main objective was to identify how firms enter in collaboration with other firms. The research differentiates two types of networks; social capital, as the networks held by individuals, and network capital, as networks held by firms. It also uses an extension of the RBV to include network resources as potentially one of these resources. The survey included questions addressing which types of partner, how (and if) it is affected by spatial proximity (using the main administrative boundaries of UK as location measure), which type of network is usually invested (hence if interactions occurred outside of work or business environment), how does firm size affects the network, and how does is change and evolve and ultimately performs. Additionally the survey also differentiates between social capital and network capital by identifying if weather or nor such interactions would continue if they were incapable of sourcing the knowledge required. Hence, based in a 10-point scale, issue such as frequency, location, motivations and performance were addressed. The survey was sent to 750 firms draw from the 2,500 firms within FAME database (Financial Analysis Made Easy) from which 74 responses were obtained (a chi-squared goodness of fit test showed the sample to be representative). Responses were than divided in 3 groups, according to their number of employees (large, medium, and small). Due to the small sample size, a non-parametric statistical method was utilized in the analysis. The survey suggests that firms frequently use local networks (except for interactions with competitors, customers and suppliers which get more frequent in the outside regions). Also that, within local networks, social capital investment is the prevalent form of network investment and that network dynamics (changing or adding new collaborators more frequently) is especially important form of innovation for small firms.

Nevertheless, Beaudry and Breschi (2003) warn that what brings about an increase of innovative outcomes for a given firm is not the regional environment *per se* but a regional environment that is already innovative. Based on European Patent data between 1990 and 1998, the paper investigates how (and if) clustering will affect firm's innovative activities. Many authors adopt Porter's view of cluster as a collection of related companies concentrated in a small area. These authors present two arguments to advocate that spatial agglomeration can't fully explain by firm's innovation: The first is that concentration takes place where innovativeness is already present. Second argument is that congestion effect arises from non-innovative firms located in a cluster, in an attempt to absorb knowledge spillover from employees of innovative firms. Based in these two arguments, these authors propose that concentration of firms is neither necessary nor sufficient for innovative activity. A second proposition relates to Jacobs (1969) idea of positive externalities among different industries leading to a higher propensity for innovation. This analysis addresses the relationship between a firm's innovative performance and spatial agglomeration. The data set included 26,055 manufacturing firms from UK and 37,724 manufacturing firm from Italy and combines three sources of data. First, the EPO-CESPRI database with patent information from Italy and UK from the European Patent Office between 1978 and 1998 (with the nomenclature of statistical territorial units—NUTS—adapted as spatial unit). The second source was company economic information from UK (from in Dun Bradstreet's one source vol 1 and vol 2) and Italy (Van Dijk's AIDA). Third source was the employment level for both countries from the Central Statistical Office (CSO) and the Instituto Nazionale de Statistica (ISTAT). The dependent variable in this model was number of patents.

Explanatory variables included firms-specific variables such as average number of employees between 1978 and 1996 (a as proxy for firm size) and firm's depreciated sum of patents from 1978 to 1989 (as a proxy for firm's stock of knowledge, hence accessing firm's individual heterogeneity). Cluster-specific explanatory variables included sector's employment in firm's same industry (accessing regional strength in a given industry) and employment figures in all other industries. The firm-specific estimated coefficients retained a positive and significant coefficient for both Italy and UK showing evidence that firm size and knowledge stock (innovation history) tend to generate a higher number of innovations.

Note that while network effects and spatial externalities are expected to increase the pace of innovation for a particular firm, their effect is largely confined in space. That is, the plausible benefits associated with a rich regional environment are expected to be potent among close proximity of the relevant actors. As a case in point, several empirical studies have corroborated the theoretical notion that spatial externalities and network effects wane with the distance between economic actors (Anselin, Varga, & Acs, 1997; Funke & Niebuhr, 2005; Keller, 2002; Owen-Smith & Powell, 2004) mainly because personal interactions and other knowledge transfer means associated with proximity weaken with increased distance.

Anselin et al. (1997) use data from 43 states and 125 metropolitan statistical area (MSA) to further investigate whether or not university research from either public or private might impact the innovative capacity in a certain region. In addition to the MSA itself, the paper applied spatial lag variables in order to capture research activities in concentric rings around the MSA. In accordance to previous research based on knowledge production function, a dataset including 4.200 new product announcements were set as the countable measure of innovation used as dependent variable. As the main explanatory variable, a proxy for R&D was constructed based on data from professional employment in high technology research laboratories (from Bowker directories). Additionally, data from research expenditures (compiled from the NSF Survey of Scientific and Engineering Expenditures at Universities and Colleges in 1982) were included as well as proxy for size, total enrollment at universities and academic quality of high technology departments at universities. Furthermore, in order to capture effects from agglomeration economies, variables such as high technology employment, location quotient for high technology employment, employment in business services and the percent of large firms were also included in the model. The test of local geographic spillover at the state level was based on five measures of local spatial interaction between university and industry research for states, including Jaffe indicator of geographic coincidence (one is a variation of Jaffe's approach and the others are traditional accessibility indices). At the MSA level, the test was based on R&D lab employment figures, working as a proxy for private R&D activity, and spatial lags variables to capture the effect of university research and private R&D in regions around the MSA (50 and 75 miles range from the geographic core center of the MSA). Overall, a positive and significant association between university research and innovative activity is found (both directly and indirectly through private R&D) within a range of 50 miles in the case of spillover from university research.

Funke and Niebuhr (2005) also attempt to quantify a more precise measurement of the geographical extent of R&D spillover, by means of accessibility and degree of interaction, which is assumed to decline as geographical distance increase. Their research is focused in 71 regions in West Germany over the period of 1976-1996 linked by intensive commuting and considerable differences in terms of GDP per capita and R&D density (as it contains rural-peripheral regions as well as agglomerated areas). The paper defines a regression analysis focusing on the relationship between the gross value added per employee (proxy for average annual productivity growth) estimated on a regional scale, and R&D activity and spillovers (altogether R&D potential) set as the self potential of a given region plus the accumulated influence of other regions. A distance decay parameter and a negative exponential function are specified to account for spatial interaction of R&D employees. A significant coefficient for all explanatory variables is found, providing empirical evidence for the geographically bounded hypothesis on R&D spillover. Precisely the paper suggests that intensity of such spillovers decline by 50% over a range of 30 kilometers and on average, decline by 60% between the centers of two neighboring regions. Along this line, Keller (2002) also attempts to quantify the effects of knowledge spillover over nationwide distances by examining the impact on productivity gains from each other's R&D spending. Keller investigates the link between R&D expenditure and R&D productivity. Particularly the paper relates the industry level R&D spending in France, Germany, Japan, UK and United States (referred to as G-5 countries) to the productivity levels in nine other OECD countries (Australia, Canada, Denmark, Finland, France, Italy, The Netherlands, Spain, and Sweden). The analysis is based on manufacturing industries from these fourteen countries between 1970 and 1995 for twelve industries at the two-to-three-digit ISIC level, among low-R&D industries (food, beverages and tobacco, textiles, wood products and furniture, paper and printing, rubber and plastics, non-metallic mineral products, basic metals, non-electrical machinery and instruments and metal products) and high-R&D industries (chemicals and drugs, electrical machinery and transportation equipment). Additionally the paper investigates the effects of globalization over the sample time period, i.e., if the effects of localized knowledge spillover has change over time. The model specified in this study uses productivity as the dependent variable, as it defines as the output divided factor-share weighted capital (number of employees in the STAN dataset multiplied by the average annual hours worked in each country's manufacturing sector). Explanatory variables include measures of cumulative R&D and geographic distance (miles between the capital cities of the countries). The paper provides evidence that international diffusion is in fact localized as the effects of R&D over productivity declines as the geographic distance between send and recipient countries increases. The elasticity estimates of productivity over distance varies between -1% and 2.4% implying that a 10% higher distance to an important R&D producer such as the United States is associated to a 0.15% lower level of productivity. Finally, the average elasticity productivity has decrease by 20% between 1970-1982 and 1983-1985, suggesting that the importance of location for knowledge spillover may be declining.

Like the researchers over firm's relative significance of a narrow versus broad knowledge base, the empirical studies on specialization vs. diversification remain inconclusive. In fact, while several empirical studies support that intra-industry localization externalities are more important in generating innovation and growth (van der Panne, 2004) other researches find that inter-industry urbanization externalities are more predominant and effective. By examining patent data in Italy, Paci and Usai (2000) finds that innovation activity was positively affected by both Marshall and Jacobs externalities. The database used refers to 24.820 observations combining 85 sectors out of 784 Local Labour Systems (LLS) in Italy from the Center for North South Economic Research (CRENoS) between 1978 and 1995. The patent data, initially classified according to the International Patent Classification, was referred to a corresponding industry based on the Yale Technology Concordance. The empirical model included a measure of specialization index (as a proxy for Marshall externalities) and production diversity index based on the reciprocal of the Gini coefficient (as a proxy for diversity externalities). The model also included control variables to account for differences of local labor systems, sectorial characteristics and technological opportunities in each industry. The dependent variable used was the annual average number of patents per capita over the period of 1990-1991. The OLS estimates indicated a positive significant coefficient for both types of externalities. Shefer and Frenkel (1998) investigate the effects of external factors of a "local innovation milieu" in the increase in rate of firm's innovation potential. Particularly, the dataset used in this empirical analysis refers to a sample of 211 firms from Electronics, Plastics, and Metals, three of the fastest-growing industries in Israel, divided in two distinct groups of firms, a first one including all the firm belonging to the Electronics industry, understood as high-tech group, a second with firms belonging to Plastics and Metals, more traditional sectors understood as comparatively low-tech. The paper sets two models for each of these two groups, a first one using as explanatory variable the number of employees in the service industries (as a proxy for agglomeration economies as explanatory variable) and a second with number of employees in the industry concerned (as a proxy for localization economies). Explanatory variables were used equally in all models to control for firms internal characteristics affecting rate of innovation (size, age, ownership type, internal R&D level, and skilled labor force). All models also specified a binary-choice of whether the firm innovates or not as the independent variable. Empirical results for both models indicates internal R&D and skilled labor force to be positively and significantly affecting firms rate of innovation. As for agglomeration vs. localization economies, both variables show positive and strong significant results only in high-tech industries. Particularly on the relationship between high-tech firms and agglomeration externalities, Henderson (1994) argues that "where a pace of technological change is higher and where cross fertilization from outside the core is crucial for breakthroughs in products." Against this background we propose the following hypothesis.

> *Hypothesis 5: Firms located in regions with a knowledge-intensive environment patent more.*

CHAPTER THREE

METHODS AND PROCEDURES

In line with the previous discussion, the empirical part of the present paper associates the number of patents awarded to a given firm with firm-specific and regional characteristics. Formally,

$$y_i = f(X, K),$$

where y_i is a vector that represents the cumulative number of patents awarded to firm i from 1983 to 2006 with firm i being a life science firm that has won at least one SBIR award from 1983 to 2006. X and K are vectors that measure relevant features of firm i and characteristics of the regional environment where firm i is located respectively.

Vector X includes five relevant firm-specific variables. We represent the size of the firm with a variable that increases with the number of employees for each firm *(Size)* and we expect a positive sign for the variable in question. While we expect the number of employees to be an appropriate proxy for the size of a firm, we also recognize that the relationship between employees and size might not be linear especially for firms that employ more automated work techniques. In order to mitigate potential issues that can arise from the issue at hand, we also include in X a variable that measures the number of years in our sample (1983 to 2006) that a firm was active as an independent firm and before potential mergers, acquisitions and bankruptcies (Age).[10] This variable increases

[10] When a firm ceases to be an independent entity, it is excluded from the sample.

with the age of the firm. We assume that older firms typically also become larger and we expect the effect of age to compensate for any residual effects of the size of the firm that are not captured by the *Size* variable.

We account for the effects of R&D intensity with two variables. The first variable measures the total sum of funds from Phase 1 and Phase 2 grants a particular firm has received from SBIR awards *(SBIR)* from 1983 to 2006. The variable captures two features of the SBIR program that can intensify R&D efforts and consequently help a given firm to produce more patents. The first feature refers to the increased research funds available for those firms that secure the most funds from the SBIR program. The second feature refers to a certification effect where in the absence of bank financing for early stage high technology firms, government grants serve as a signal of quality for private investors (Lerner, 1999). It follows that firms with the most grants from SBIR may be better off in attracting private investors.

Collectively, the direct expansion of the R&D pool and the plausible attraction of private investors lead us to expect a positive sign for the variable that reflects the sum of funds from SBIR awards. The second variable we employ to account for the level of R&D performed by a given firm is a dummy variable that takes the value of 1 (and 0 otherwise) for firms that have received funds from venture capital firms[11] *(VC)*. Venture capital firms provide sizeable funds to firms in high technology areas such as life science especially in later financing stages (Carpenter & Peterson, 2002), which then can allow

[11] We acknowledge that a more appropriate measure of the effects of venture capital funds would account for the time of venture financing, so that firms that received funds at an early stage of their operation (or even started backed up with venture capital) would normally be more prolific in producing patents. Unfortunately, data limitations do not allow us to incorporate the time of funding in the analysis.

the funded firms to intensify their R&D efforts and potentially increase the number of patents they produce. Against this background, we expect the dummy variable in question to have a positive sign.[12]

We represent the scope of knowledge sources employed by a given firm with a Herfindahl-type index where, the number of citations (in the 3 digit patent class) is the equivalent of the firm's sale in a industrial context, as it was first proposed by Trajtenberg et al. (1997) and was later adopted among others by Bessen (2008) and Magazzini et al. (2009).

The index ranges from 0 to 1 and is defined for each patent as $Scope = 1 - \sum_{j}^{n_j} s_{kj}^2$, where s_{kj}^2 is the percentage of citations made by patent k that belong to patent class j where there are n_j patent technology classes. If a patent cites patents that belong to a narrow set of technology classes, then the index takes a low value which signals a patent that borrows from a limited number of scientific disciplines different than the discipline of the particular patent. On the contrary, a high index score originates from patents that build on knowledge from a wide range of sources. The scope index for each firm is calculated as the average scope index for each of patents of the firm. The index consists in a backward-looking measure as it derives directly from the relationship between a patent body of knowledge preceded it, as opposed to Forward-looking measures ("generality") which derives from the relationship between a patent and the innovation that is build based on it (Trajtenberg et al., 1997).

[12] A third source of funds that can affect R&D intensity is contractual work performed for other companies. However, we argue that the degree of prevalence of this type of funding source is rather limited for the firms in our sample, who are mostly small, typically do not have an established record that could attract outside contracts and are mostly in the race of establishing their own research.

As explained in the preceding section, both narrow and broad knowledge sources have advantages. As such, the scope variable is largely exploratory and we do not form theoretical expectations with regard to the direction it moves the patent production of a particular firm.

Vector K includes variables used to capture regional characteristics that are expected to influence the patenting patterns of a given firm. In order to incorporate the effects of collocation with other successful firms we employ two variables that represent the presence of such firms in close proximity to the firm. The first variable measures the number of life science firms located within a twenty miles radius[13] from the origin firm that have received at least one SBIR grant from 1983 to 2006 *(SBIRFirms)*. In the second variable we employ a stricter measure of successful firms; we measure the number of life sciences firms in a twenty miles radius that not only have received SBIR grants but have also been granted at least one patent from 1983 to 2006 *(PatentFirms)*. Since both measures reflect successful firms we expect a positive sign for both variables. However, if collocation with the most successful firms is more conducive to increased patenting rate of the origin firm, then the variable that measures the number of firms which have at least one patent will have a larger effect than the *SBIRFirms* variable.

In order to test whether there is a relationship between the rate of innovation for a given firm and the degree that close by firms source knowledge from a narrow or a diverse set of scientific sources we include in K a variable that measures the average

[13] Since the relevant economic radius where spatial externalities and network effects are expected to occur can vary according to geography, social norms, spatial infrastructure and the like we also tried larger and smaller radii to ensure robustness of the results. The use of different size radii did not alter the empirical results presented on Table 2 qualitatively. These results are not presented here for parsimony.

scope index for firms located in a twenty miles radius from the origin firm *(ScopeFirms)*. The variable in question is exploratory because the arguments presented previously with regard to benefits accruing from a narrow compared to a broad knowledge base apply here as well.

Our definition of the regional environment is not confined to the collocation of firms in the relevant geographic space but extends to other factors that can affect the rate at which knowledge is generated and plausibly flows in the area where firms reside. In this vein, we include a variable that measures the total amount of research dollars awarded from the National Institutes of Health (NIH) from 1983 to 2006 to universities located in the same Metropolitan Statistical Area (MSA) with the origin firm *(UnivNIH)*. NIH grants are a major source of funds for life science research performed at university labs, hence one would expect increased funding to be associated with increased research. By its nature, university research can create knowledge that is not perfectly appropriated by its creators and so it is possible for some of it to flow to nearby firms (Varga, 2000). Varga also concluded that this flow from university to nearby firms will be influenced by local agglomeration factors, primarily high technology employment. By developing a regression model consistent with Griliches-Jaffe knowledge production function framework and based on a variety of datasets including the United States Small Business Administration (SBA) innovation citation database (as a proxy of innovation production), R&D employment figures from the 17th edition of the Industrial Research Laboratories of the United States (as a proxy for private research activities) and research expenditures data from the NSF Survey of Scientific and Engineering Expenditures at Universities and Colleges (as a proxy for research activity at academic institutions), Varga investigates the effects of agglomeration on academic knowledge transfers. The model estimates demonstrate that a given amount of university research spending can lead to different innovation outputs, depending on the concentration of economic activity. Moreover, it is found that a “critical mass” of agglomeration is necessary in order to achieve significant local economic effect on academic research spending. Accordingly, we use NIH grants to approximate the degree of knowledge from public institutions that can flow to local firms and as such we expect a positive sign for it.

As well, we construct location quotients for each firm’s MSA in order to account for the intensity of life sciences activity in the region where the firm resides. Because most of the employees of life science firms typically have a degree in biological sciences we approximate life science activity with the number of employed biochemists and biophysicists in the MSA. We estimate the location quotient for each MSA m as $LQ_m = \frac{BS_m / BS_{US}}{EM_m / EM_{US}}$, where BS_m is the number of biochemists and biophysicists employed in the MSA of the origin firm, BS_{US} is the number of biochemists and biophysicists employed in the United States, EM_m is the total number of employees in the MSA of the origin firm and EM_{US} is the total number of employees in the United States. It follows that a high value of the location quotient signals increased concentration of life science activity in the MSA when compared with the rest of the country. The location quotient is used to capture the effects of a favorable business climate (including policies, capital sources, infrastructure, etc.) towards life science that is typically associated with regions exhibiting above average concentration of life science activity. For instance, much of the life science activity in the United States occurs in the infamous San Diego, CA, and

Boston, MA, clusters which are characterized by an ample pool of venture capital (Powell, Koput, Bowie, & Smith-Doerr, 2002), favorable tax regimes, supporting institutions performing complementary research etc. It is in these types of areas that the location quotient is high; hence we anticipate a positive sign for it. The final proposition is than summarized below.

CHAPTER FOUR

DATA SOURCES AND PRESENTATION

InKnowVation, Inc. provided a dataset that reported all the Phase 1 and Phase 2 SBIR grants awarded to LSFs from 1983 up to 2006. The dataset included information for each grant as well as for each LSF that won grants. The information included in the dataset was used to construct the dependent variable and the *Age*, *Size*, *VC*, *Scope* and *SBIR* variables. For the *SBIR* variable we converted the nominal amount for each grant to 2006 dollars using the CPI. The data from InKnowVation, Inc. included the address of each SBIR winner which we converted to coordinates with the ArcView software to develop the *SBIRFirms*, *PatentFirms* and *ScopeFirms* variables. In order to construct the location quotients we used employment statistics available from the US Bureau of Labor Statistics. The *UnivNIH* variable was built with data collected from NIH and reflected life science grants from the National Institutes of Health. Similar to the SBIR variable, the nominal amounts of the NIH grants were converted to real 2006 dollars using the CPI.

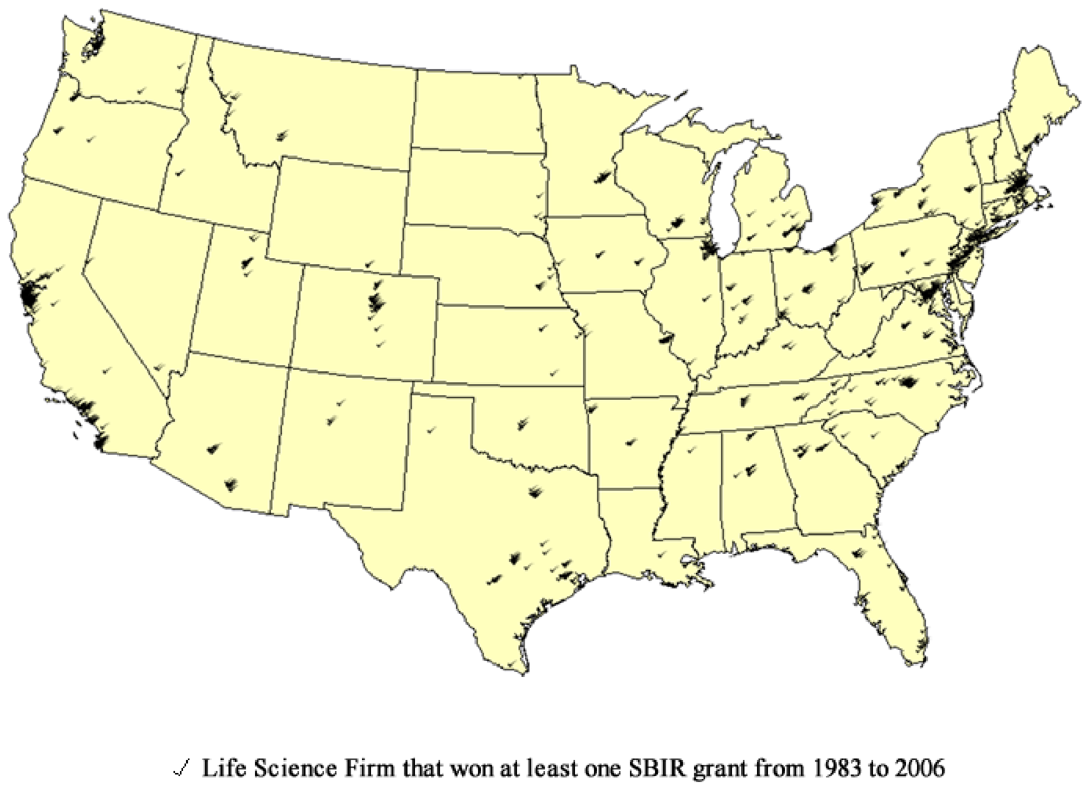

Figure 1. Life Science Firms Included in the Dataset (1,671 Firms).

Figure 1 is a map of the life science firms included in the dataset. Most of the firms that received SBIR grants are located at the East and West Coast but roughly one third of the firms are located in urban and rural interior regions of the United States. The wide geographical distribution of the firms in the dataset suggests that our empirical estimates are not potentially specific to a certain region of the United States, hence the policy implications of the present work can apply regardless of location.

An additional observation that deserves attention because it relates to the variables that measure the density of successful firms in a twenty miles radius from the origin firm (*SBIRFirms*, *PatentFirms*) is the very close proximity observed among SBIR winners. The majority of the firms in the dataset are located within walking distance from at least one SBIR winner, and mostly within the twenty miles radius used for the aforementioned variables. The close proximity among firms is not typical only among firms in the most densely populated regions of the country (i.e., New York, NY, Boston, MA, San Diego, CA, San Francisco, CA, etc.) but also in interior cities like Fayetteville, AR and Lincoln, NE. This similarity of location trends across firms situated in diverse regions of the country, largely emanates from the fact that regardless of region many of those firms are located in science parks, university campuses, business incubators and other facilities that either belong to research universities or are in the vicinity of universities. The implication of the previously described location patterns for the current work is that the twenty miles radius we use for the *SBIRFirms* and *PatentFirms* variables seems to be capture potential spatial externalities and network effects regardless of the geographic location of firms in the dataset.

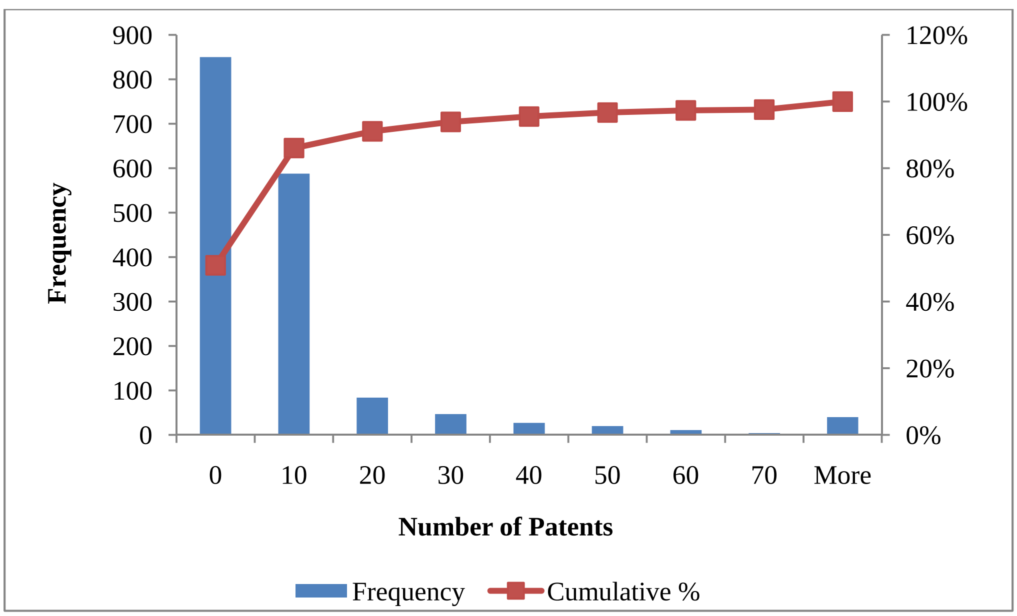

Figure 2. Histogram of Patents Produced by Firms in the Sample (1,671 Firms).

Table 1. Descriptive Statistics of the Variables Used in the Empirical Model.

Variable / Statistic	# of Observations[1]	Mean	Median	Mode	Std Deviation
Patents	1671	7.978	0.000	0.000	33.189
Size	1506	3.704	3.000	1.000	2.940
Age	1449	6.838	5.000	2.000	5.364
SBIR	1671	1.881	0.648	0.094	6.968
VC	545				
Scope	764	0.570	0.600	0.000	0.225
SBIRFirms	1671	52.211	28.000	145.000	53.990
PatentFirms	1671	29.906	14.000	0.000	33.229
FirmScope	1387	0.567	0.575	0.595	0.079
LQE	1167	3.653	4.141	6.094	2.717
UnivNIH	1619	845.209	670.639	1279.807	756.826

[1] In the case of the VC variable, the figure measures the number of firms that have received venture capital funds.

Patents	Total number of patents awarded to firm from 1983 to 2006
Size	Variable that is increasing with the number of the employees at the origin LSF (the variable values followed the following codification: 1 = 1 to 4 employees; 2 = 5 to 9 employees; 3 = 10 to 14 employees; 4 = 15 to 19 employees; 5 = 20 to 24 employees; 6 = 25 to 49 employees; 7 = 50 to 74 employees; 8 = 75 to 99 employees; 9 = 100 to 149 employees; 10 = 150 to 249 employees; 11 = 250 to 500 employees)
Age	Number of years from 1983 to 2006 that the firm was active as an independent firm and before potential mergers, acquisitions, and bankruptcies
SBIR	Total amount from Phase 1 and Phase 2 grants raised by the firm from 1983 to 2006 ($2.006 B.)
VC	Variable that takes the value of 1 (and 0 otherwise) for firms that received funds from venture capital firms from 1983 to 2006
Scope	Index that is increasing with the number of different patent classes cited by the firm's patents
SBIRFirms	Number of life science firms located within a 20-mile radius from the origin firm that have received at least one SBIR grant from 1983 to 2006
PatentFirms	Number of life science firms in a 20-mile radius that have received at least one SBIR grant and have been granted at least one patent from 1983 to 2006
FirmScope	Average Scope index for firms located within twenty miles from the firm
LQE	Location quotient for each firm's Metropolitan Statistical Area
UnivNIH	Total amount of dollars awarded from the National Institutes of Health (NIH) from 1983 to 2006 to universities located in the same Metropolitan Statistical Area with the firm ($2.006 B.)

Table 1 presents the descriptive statistics of the variables used in the empirical model and Figure 2 is a histogram of the dependent variable. Most of the firms in the dataset were relatively young and small; on average they were in business for about 7 years, had somewhere between 10 and 19 employees and had sourced approximately $2 million from SBIR grants. However, the considerably larger standard deviation of the *SBIR* variable coupled with its relatively small modal value indicate that there was significant variability across SBIR winners with some firms being more successful in accumulating SBIR funds that others. With regard to the *SCOPE* index, for those firms that we could calculate the index (e.g., those firms with at least one patent), the average value indicates that firms in the dataset sourced external knowledge outside their core research focus. Nevertheless, the modal value of 0 suggests that the majority of firms used external knowledge sources that were largely familiar to them. Concerning the regional environment, universities in each firm's MSA had sourced on average more than $845 million from NIH funds and firms were located in regions with high life science activity as the location quotient suggests. Finally, on average each firm in the dataset had about 52 SBIR winners in 20 miles radius, 30 of which were also granted at least one patent from 1983 to 2006. The striking difference between the modal values of the *SBIRFirms* and *PatentFirms* variables (145 versus 0, respectively) emanates from the highly left skewed distribution of the dependent variable as shown in Figure 2. About half of the firms in the dataset were not granted a patent between 1983 and 2006 while 86% of all firms (1,438 of the 1,671 firms in the dataset) had less than 10 patents. On the contrary, a handful of firms had more than 70 patents over the time period covered in the dataset. Taken together, the highly left skewed distribution of the dependent variable and the presence of firms with well above average patent production reinforce the motivation of the present study in analyzing selected factors that can affect the patenting rate of a given firm.

CHAPTER FIVE

ESTIMATION RESULTS

Because the dependent variable of the empirical specification (total number of patents for a firm from 1986 to 2003) is a count variable we use a Poisson model to examine the relationship between patent production and firm-specific and regional characteristics. The fitted Poisson count data model is based on a logarithmic specification for the conditional mean, and the model parameters were estimated by maximum likelihood (ML). In the fitted model we relax the assumption of the traditional Poisson model under which for any Poisson process the mean of the outcome equals the variance because the assumption in question is not supported by our data. Given the fitted model, we computed the estimated marginal effect for each of the continuous explanatory variables in the model, ($E_x\left[\frac{\partial E(y|x)}{\partial x_j}\right] = \beta_j E[\exp(x'\beta)]$) (Winkelmann, 2008). This expression indicates the marginal effects are potentially different for each observation, and we compute the sample average of the estimated marginal effects for all observations and report these values in Table 2.[14] Due to observed evidence of heteroskedasticity of unknown form, the parameter standard errors were estimated with White's robust estimator.[15,16] Finally, because we use two variables to capture the presence of successful

[14] Footnote [a] in Table 2 provides details on the marginal effects of the dummy variables.

[15] The Poisson model with White's standard errors was estimated with the Generalized Estimating Equation (GEE) method, which is not an ML method. Accordingly, the goodness-of-fit statistic provided by the software used is the quasi-likelihood value (QIC), which is analogous to the AIC statistic for the ML estimators (Hardin & Hilbe, 2003).

[16] Because the *Scope* variable was computed only for firms that have been issued at least one patent, the sample used for the empirical specification included only such firms. To ensure the robustness of our results to potential selection bias problems associated with the issue at hand (i.e. the firms with patents had

firms in vicinity of the origin firms Table 2 has two models. In Model 1 we employ the *PatentFirms* variable and in Model 2 we replace the *PatentFirms* variable with the *SBIRFirms* variable.

Table 2. Marginal Effects[a] of Poisson Model with Mean Equal Variance Assumption Relaxed. The Dependent Variable is the Total Number of Patents Awarded to a Given Firm from 1983 to 2006.

	Model 1		**Model 2**	
Variables	**Estimates**	**Std Errors[b]**	**Estimates**	**Std Errors[b]**
Intercept	-0.459	0.940	-0.365	0.923
Age	0.458	0.014 ***	0.455	0.014 ***
Size	2.710	0.040 ***	2.733	0.040 ***
VC	3.186	0.172 ***	3.306	0.171 ***
SBIR	-0.014	0.001	-0.013	0.002
Scope	-6.524	0.356 *	-6.720	0.3556 *
ScopeFirms	9.162	1.606	8.175	1.585
PatentFirms	0.042	0.002 *		
SBIRFirms			0.020	0.002
Location Quotient	0.310	0.035	0.357	0.035
UnivNIH	-0.001	0.000	-0.001	0.000
Scale[c]	4.856		4.870	
GEE QICu	-2439.759		-2448.709	
Heteroskedasticity Test [d]	20.650 ***		20.8 00 ***	
Multicollinearity Condition #	40.018		40.103	
Number of Observations	500		500	

[a] The marginal effects for continuous variables are the average marginal effect for all observations. For the VC variable, the marginal effect is approximated as the change in the dependent variable resulting from going from the 0 to the 1 category.
[b] The standard errors correspond to the GEE estimates.
[c] The scale parameter comes from the model without robust standard errors.
[d] An LM test employing results from auxiliary regressions was used to test for heteroskedasticity.
*** .001 significance, ** .05 significance, * .10 significance
Note: The log link function was used for the Poisson model.

some unobserved characteristics that allowed them to enter the sample), we also run the Models in Table 2 excluding the *Scope* variable. The results of these models are presented in the Appendix table and are largely similar to the results presented in Table 2. Note that a standard technique to handle selection bias concerns is the two stage Heckman model. Unfortunately this model does not apply to models where the dependent variable is a count variable.

In line with theoretical expectations and empirical findings of previous studies, the empirical estimates presented in Table 2 indicate that larger and older firms are more prolific in producing patents. A one-unit increase in the size variable is associated with 2.7 more patents for a given firm. To put this figure in perspective recall that approximately 60% of the firms in the sample had 3 or less patents over time (see Figure 2). Therefore, larger firms appear to have an advantage in the rate they produce patents. As explained in the theoretical part, we expected the *Age* variable to capture residual effects of size that the *Size* variable does not capture. The magnitude and sign of the *Age* variable agree with our theoretical expectations. One extra year in business from 1983 to 2006 for a given firm was associated with close to half additional patent over time.

With regard to the variables that approximated the R&D intensity for a particular firm we find that venture capital backed firms were significantly better off in producing patents while the amount of funds from SBIR awards did not have strong explanatory power in the matter. In particular, firms that have received venture capital funds produced more than three additional patents from the rest of the firms. This estimates points towards beneficial effects that emanate from synergistic activities between private and federal sources of research funds. The finding that pertains to SBIR funds merits attention because at a first glance may signal an inefficiency of federal grants in promoting innovation. While the empirical estimates do not exclude such an interpretation there are at least two reasons to argue that this conclusion is premature. First, as explained previously SBIR grants are typically used to fund early stages of the research process while venture capital funds are used for later stages. But in order to reach later stages of research, the early stages need to occur first. Hence, the contribution of SBIR grants may rest mainly in setting up the necessary foundation for later research stages that are pursued by private markets and are heavily directed towards innovative outcomes such as patents and less towards developing the basic science. Second, SBIR funds have been found to offer a certification effect (Lerner, 1999) to the firms they secure them. It is then possible that venture capitalists invest more heavily on SBIR winners; because typically the amount invested by venture capitalists to a given firm greatly outweighs the amount from SBIR grants, the contribution of SBIR grants is maybe masked due to the difference in the invested amounts. For example, the majority of firms in our sample had sourced about 94,000 dollars from SBIR grants (see Table 1), which falls short of the millions of dollars typically invested by venture capital firms. Unfortunately, due to data limitations we are not able to test empirically the aforementioned arguments. As such, we do not provide a substantive interpretation for the statistically weak coefficient of the SBIR variable.

The results that pertain to the external scientific knowledge sources used by a firm are intriguing. An increase in the *Scope* index (i.e., an increase in the breadth of external scientific technology sources) brings about 6.5 less patents for a given firm. Alternatively, firms that employ knowledge close to their core science are considerable more prolific in producing patents. While our estimate is not conclusive, it does suggest that life science firms that follow a strategy that focuses on core competencies are the most successful in improving their patenting rate. Also note that the *Scope* index has the largest marginal effect among the variables with statistically strong coefficients. This finding highlights the economic significance of the type of external scientific knowledge that a particular firm employs. Contrary to the *Scope* index, the *ScopeFirms* variable did not have explanatory power which indicates that the type of external knowledge sources employed by nearby firms do not alter the patent production of a given firm.

Finally, the empirical estimates that refer to the regional environment suggest that what matters for the production of patents is collocation with the most successful firms and not the R&D intensity of close-by universities (*UnivNIH*) or the intensity of life science activity in the region (*Location Quotient*). In Model 1 the *PatentFirms* variable shows that one additional SBIR winner in close proximity to the origin firm that has been issued at least one patent is associated with 0.042 more patents for the origin firm. On first sight the 0.042 figure is small but if we assume a linear process and evaluate this figure at the average number of such firms in radius (see Table 1), then collocation with successful firms is associated with 1.25[17] more patents for the origin firm. The figure becomes more economically relevant because it illustrates the advantages that can be realized by firms located in the traditional life science clusters in the United States, where it is not uncommon to have more than 100 firms located within a twenty miles radius. To quantify this advantage, if we assume a linear process so that each nearby firm has the same and non-decreasing contribution to the patenting rate of the origin firm collocation with 100 successful firms would translate to 4.2 more patents for the origin firm. On the other hand, the statistically weak coefficient of the *SBIRFirms* coefficient in Model 2 indicates that firms can benefit from collocation with the most successful firms and not with collocation with less successful enterprises. Put alternatively, what matters for an increase in patents for a given firm is not the number of SBIR winners in proximity but the number of SBIR winners that have been issued patents.

[17] 29.906 (average number of *PatentFirms*)*0.042=1.25

CHAPTER SIX

CONCLUDING COMMENTS

The SBIR program is the largest innovation program in the United States and provides more than $2 billion per year to fund early stage research projects with innovative and commercial potential pursued by private firms that typically operate in high-technology industries such as life science. Two major goals of the program are to increase the innovative pace of funded firms and to boost the commercialization of research projects initially funded by the program. The contribution of the SBIR program with regard to commercialization efforts is well documented in the academic literature and exemplified by firms such as Apple Computer, Chiron, Compaq, and Intel that received early stage SBIR funds. On the other hand, whether the SBIR program promotes innovation is seen with skepticism from some scientists and it is a topic that surprisingly has received relatively little attention in the literature. In an effort to inform policy makers to increase the innovation of the program, this study focused on selected factors that have helped SBIR winners to produce more patents.

To complement standard determinants of patent production for a given firm, we are among the first studies to incorporate in the analysis the effects of the scope of scientific knowledge sources a firm is using. The empirical estimates suggest that the breadth of external knowledge employed by a particular firm is an important factor that needs to be accounted for. We find that firms that use external knowledge sources that are close to the core of the firms' knowledge produce significantly more patents that the rest of the firms. Further, we conclude that larger and venture capital-backed firms exhibit above average performance in producing patents. Finally, our empirical analysis suggests that firms located in close proximity to other innovative SBIR winners are also more prolific in producing patents. Note that collocation with SBIR winners *per se* (compared to collocation with SBIR winners with patents) appeared to have no effect on the patent production of a given firm.

We believe our results can initiate further research on issues related to the SBIR program and more generally to the relationship between private and federal funds, as well as on issues that pertain to spatial effects among firms. First, due to data limitations, we were not able to identify why, as our empirical estimates suggest, the involvement of venture capital firms and not the amount raised from SBIR grants for a given firm alter the patent production of the firm. We provide relevant arguments on the issue but further research can provide a finer analysis. Second, a shortcoming of our study is that we could not assess the exact mechanisms that allow only SBIR winners with patents to affect the patenting trends of a given firm. Further research can unravel the ways that allow only the most successful firms to affect the performance of nearby firms.

BIBLIOGRAPHY

Acs, Z.J., Anselin, L., & Varga, A. (2002). Patents and innovation counts as measures of regional production of new knowledge. *Research Policy*, *31*(7), 1069-1085.

Adams, J.D., & Jaffe, A.B. (1996). Bounding the effects of R&D: An investigation using matched establishment-firm data. *RAND Journal of Economics*, *27*(4), 700-721.

Alcacer, J, & Gittelman, M. (2006). Patent citations as a measure of knowledge flows: The influence of examiner citations. *The Review of Economics and Statistics*, *88*(4), 774-779.

Aldrich, H.E., & Wiedenmayer, G. (1993). From traits to rates: An ecological perspective on organizational foundings. *Advances in Entrepreneurship, Firm emergence, and Growth*, *1*, 145-195.

Anselin, L., Varga, A., & Acs, Z.J. (1997). Local geographic spillovers between university research and high technology innovations. *Journal of Urban Economics*, *42*(3), 422.

Audretsch, D.B., & Feldman, M.P. (1996). R&D spillovers and the geography of innovation and production, *American Economic Review*, *86*(3), 630-40.

Audretsch, D.B., Weigand, J., & Weigand, C. (2002). The impact of the SBIR on creating entrepreneurial behavior. *Economic Development Quarterly*, *16*(1), 32.

Audretsch, D.B. (2003). Standing on the shoulders of midgets: The U.S. Small Business Innovation Research program (SBIR). *Small Business Economics*, *20*(2), 129.

Audretsch, D.B., Link, A.N., & Scott, J.T. (2002). Public/private technology partnerships: Evaluating SBIR-supported research. *Research Policy*, *31*(1), 145.

Autant-Bernard, C. (2001). The geography of knowledge spillovers and technological proximity. *Economics of Innovation and New Technology*, *10*(4), 237-254.

Basu, P. (2004). Startups hope federal study fixes SBIR flaws. *Nature Biotechnology*, *22*(6), 644-644.

Beaudry, C., & Breschi, S. (2003). Are firms in clusters really more innovative? *Economics of Innovation and New Technology*, *12*(4), 325-342.

Bessen, J. (2008). The value of US patents by owner and patent characteristics. *Research Policy*, *37*(5), 932-945.

Blundell, R., Griffith, R., &Windmeijer, F. (2002). Individual effects and dynamics in count data models. *Journal of Econometrics*, *108*(1), 113-131.

Boix, R., & Galletto, V. (2009). Innovation and industrial districts: A first approach to the measurement and determinants of the I-district effect. *Regional Studies*, *43*(9), 1117-1133.

Boschma, R.A. (2005). Proximity and innovation: A critical assessment. *Regional Studies*, *39*(1), 61-74.

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

Breschi, S., & Lissoni, F. (2001). Localised knowledge spillovers vs. innovative milieux: Knowledge "tacitness" reconsidered. *Papers in Regional Science*, *80*(3), 255-273.

Carpenter, R.E., & Petersen, B.C. (2002). Capital market imperfections, high tech investment, and new equity financing. *The Economic Journal*, *112*(477), F54-F72.

Coenen, L., Moodysson, J., & Asheim, B.T. (2004). Nodes, networks and proximities: On the knowledge dynamics of the Medicon Valley biotech cluster. *European Planning Studies*, *12*(7), 1003-1018.

Cohen, W.M. (1995). Empirical studies of innovative activity and performance. In P. Stoneman (Ed.), *Handbook of the Economics of Innovation and Technical Change*. Oxford: Blackwell.

Cohen, W.M., & Levin, R.C. (1989). Empirical studies of innovation and market structure. In R. Schmalensee & R. Willig (Eds.), *Handbook of industrial organization*. Amsterdam: Elsevier.

Cooke, P. (2002). *Knowledge economies. Clusters, learning and cooperative advantage*. New York: Routledge.

Cooper, R.S. (2003). Purpose and performance of the Small Business Innovation Research (SBIR) program. *Small Business Economics*, *20*(2), 137.

Dahlander, L., & McKelvey, M.. (2005). The occurrence and spatial distribution of collaboration: Biotech firms in Gothenburg, Sweden. *Technology Analysis & Strategic Management*, *17*(4), 409-431.

Delaney, E.J. (1993). Technology search and firm bounds in biotechnology: New firms as agents of change. *Growth and Change*, *24*, 206-206.

Eisenhardt, K., & Tabrizi, B.N. (1995). Accelerating adaptive processes: Product innovation in the global computer industry. *Administrative Science Quarterly*, *40*, (1).

Elston, J.A., & Audretsch, D.B. (2011). Financing the entrepreneurial decision: An empirical approach using experimental data on risk attitudes. *Small Business Economics*, 1-14.

Feldman, M.P. (1999). The new economics of innovation, spillovers and agglomeration: A review of empirical studies. *Economics of Innovation & New Technology*, *8*(1/2), 21.

Fontes, M. (2005). Distant networking: The knowledge acquisition strategies of 'out-cluster' biotechnology firms. *European Planning Studies*, *13*(6), 899-920.

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

Garnsey, E., Cannon-Brooks, A. (1993). The Cambridge phenomenon revisited: Aggregate change among Cambridge high technology firms since 1985. *Entrepreneurship and Regional Development*, *5*, 179-207.

Gittelman, M. (2007). Does geography matter for science-based firms? Epistemic communities and the geography of research and patenting in biotechnology. *Organization Science*, *18*(4), 18.

Griliches, Z. (1992). The search for R&D spillovers. *Scandinavian Journal of Economics*, *94*(Supplement), 24-97

Gurmu, S., & Pérez-Sebastián, F.. (2008). Patents, R&D and lag effects: Evidence from flexible methods for count panel data on manufacturing firms. *Empirical Economics*, *35*(3), 507-526.

Håkanson, L. (2005). Epistemic communities and cluster dynamics: On the role of knowledge in industrial districts. *Industry & Innovation*, *12*(4), 433-463.

Hall, B.H., & Ziedonis, R.H. (2001). The patent paradox revisited: An empirical study of patenting in the US semiconductor industry, 1979-1995. *RAND Journal of Economics*, *32*(1), 101-128.

Hardin, J.W., & Hilbe, J. (2003). *Generalized estimating equations*. Boca Raton, FL: CRC Press.

Hausman, J., Hall, B.H., & Griliches, Z. (1984). Econometric models for count data with an application to the patents-R & D relationship. *Econometrica: Journal of the Econometric Society*, *52*(4), 909-938.

Henderson, J.V. (1994). *Externalities and industrial development* (NBER Working Paper 4878). Cambridge, MA: National Bureau of Economic Research.

Huggins, R, & Johnston, A. (2010). Knowledge flow and inter-firm networks: The influence of network resources, spatial proximity and firm size. *Entrepreneurship & Regional Development*, *22*(5), 457-484.

Jacobs, J. (1969). The economy of cities. London: Penguin.

Jaffe, A.B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, *108*(3), 577-98.

Jaffe, A.B., Trajtenberg, M., Henderson, R., & Henderson, J.V. (2005). Geographic localization of knowledge spillovers as evidenced by patent citations. In *New economic geography*, International Library of Critical Writings in Economics, vol. 184. Northampton, MA: Elgar.

Katila, R. (2002). New product search over time: past ideas in their prime? *Academy of Management Journal*, *45*(5), 995-1010.

Katila, R., & Ahuja, G. (2002). Something old, something new: A longitudinal study of search behavior and new product introduction. *Academy of Management Journal*, *45*(6), 1183-1194.

Keller, W. (2002). Geographic localization of international technology diffusion. *American Economic Review*, *92*(1), 120-142.

Kogut, B., & Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science*, 383-397.

Laursen, K., & Salter, A. (2006). Open for innovation: The role of openness in explaining innovation performance among UK manufacturing firms. *Strategic Management Journal*, *27*(2), 131-150.

Lazerson, M. (1995). A new Phoenix? Modern putting-out in the modena knitwear industry. *Administrative Science Quarterly*, *40*, 34-59.

Leiponen, A. (2005). Skills and innovation. *International Journal of Industrial Organization*, *23*(5-6), 303-323.

Leiponen, A., & Helfat, C.E. (2010). Innovation objectives, knowledge sources, and the benefits of breadth. *Strategic Management Journal*, *31*(2), 224-236.

Lerner, J. (1999). The government as venture capitalist: The long-run impact of the SBIR program. *Journal of Business*, *72*(3), 285-318.

Levinthal, D.A., & March, J.G. (1993). The myopia of learning. *Strategic Management Journal*, *14*(S2), 95-112.

Liebeskind, J.P., Oliver, A.L., Zucker, L., & Brewer, M. (1996). Social networks, learning, and flexibility: Sourcing scientific knowledge in new biotechnology firms. *Organization Science*, 428-443.

Link, A.N., & Ruhm, C.J. (2009). Bringing science to market: Commercializing from NIH SBIR awards. *Economics of Innovation and New Technology*, *18*(4), 381-402.

Lunn, J. (1987). An empirical analysis of firm process and product patenting. *Applied Economics*, *19*(6), 743-751.

Magazzini, L., Pammolli, F., Riccaboni, M., & Rossi, M.A. (2009). Patent disclosure and R&D competition in pharmaceuticals. *Economics of Innovation and New Technology*, *18*(5), 467-486.

March, J.G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, *2*(1), 71-87.

McKelvey, M., Alm, H., & Riccaboni, M. (2003). Does co-location matter for formal knowledge collaboration in the Swedish biotechnology–pharmaceutical sector? *Research Policy*, *32*(3), 483-501.

Meyer, J.W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, *83*(2), 24.

Montalvo, J.G. (1997). GMM estimation of count-panel-data models with fixed effects and predetermined instruments. *Journal of Business & Economic Statistics*, *15*(1), 82-89.

Moodysson, J., & Jonsson, O. (2007). Knowledge collaboration and proximity: The spatial organization of biotech innovation projects. *European Urban and Regional Studies*, *14*(2), 17.

Oliver, A.L. (1994). *In between markets and hierarchies—Networking through the life cycle of new biotechnology firms* (Institute for Social Science Research, Working Paper Series 91). Los Angeles, CA: UCLA, Institute for Social Science Research.

Orlando, M.J. (2004). Measuring spillovers from industrial R&D: On the importance of geographic and technological proximity. *RAND Journal of Economics*, *35*(4), 777-786.

Owen-Smith, J, & Powell, W.W. (2004). Knowledge networks as channels and conduits: The effects of spillovers in the Boston biotechnology community. *Organization Science*, *15*(1), 5-21.

Paci, R., & Usai, S. (2000, August). Externalities, knowledge spillovers and the spatial distribution of innovation. Paper presented at the European Regional Science Association, Barcelona, Spain.

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

Peeters, C., & van Pottelsberghe de la Potterie, B. (2006). Innovation strategy and the patenting behavior of firms. *Journal of Evolutionary Economics*, *16*(1), 109-135.

Powell, W.W., Koput, K.W., Bowie, J.I., & Smith-Doerr, L. (2002). The spatial clustering of science and capital: Accounting for biotech firm-venture capital relationships. *Regional Studies*, *36*(3), 291-305.

Rosenthal, S.S., & Strange, W.C. (2003). Geography, industrial organization, and agglomeration. *Review of Economics and Statistics*, *85*(2), 377-393.

Saxenian, A. (1994). Regional advantage: Culture and competition in Silicon Valley and Route 128. Cambridge, MA: Harvard University Press.

Schumpeter, J.A. (1942). *Capitalism, socialism and democracy*. New York: Harper.

Shefer, D., & Frenkel, A. (1998). Local milieu and innovations: Some empirical results. *The Annals of Regional Science*, *32*(1), 185-200.

Sorenson, O., & Stuart, T.E. (2001). Syndication networks and the spatial distribution of venture capital investments. *American Journal of Sociology*, *106*(6), 1546-1588.

Teece, D.J. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy*, *15*(6), 285-305.

Teece, D.J. (2005). Technology and technology transfer: Mansfieldian inspirations and subsequent developments. *The Journal of Technology Transfer*, *30*(2/2), *17-33*.

Thornton, P.H. (1999). The sociology of entrepreneurship. *Annual Review of Sociology*, *25*(1), 19-46.

Trajtenberg, M., Henderson, R., & Jaffe, A. (1997). University versus corporate patents: A window on the basicness of invention. *Economics of Innovation and New Technology*, *5*(1), 19-50.

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

Varga, A. (2000). Local academic knowledge transfers and the concentration of economic activity. *Journal of Regional Science*, *40*(2), 289.

Wallsten, S.J. (2001). An empirical test of geographic knowledge spillovers using geographic information systems and firm-level data. *Regional Science and Urban Economics*, *31*(5), 571-599.

Wessner, W.C. (2002). Entrepreneurial finance and the new economy. *Venture Capital*, *4*(4), 349-355.

———. (2005). Driving innovations across the valley of death. *Research Technology Management*, *48*(1), 9-12.

———. (2009a). *An Assessment of the Small Business Innovation Research Program at the National Institutes of Health*. Washington, DC: The National Academies Press.

———. (2009b). *Venture Funding and the NIH SBIR program*. Washington, DC: The National Academies Press.

Winkelmann, R. (2008). *Econometric analysis of count data*. Berlin: Springer Verlag.

Zander, U., & Kogut, B. (1995). Knowledge and the speed of the transfer and imitation of organizational capabilities: An empirical test. *Organization Science*, *6*(1), 76-92.

Appendix Table. Marginal Effects[a] of Poisson Model with Mean Equal Variance Assumption Relaxed Excluding Scope Variable. The Dependent Variable is the Total Number of Patents Awarded to a Given Firm from 1983 to 2006.

Variables	Model 1		Model 2	
	Estimates	Std Errors[b]	Estimates	Std Errors[b]
Intercept	-1.638	0.960	-1.551	0.944
Age	0.421	0.014 ***	0.418	0.014 ***
Size	2.374	0.039 ***	2.405	0.039 ***
VC	2.644	0.191 ***	2.773	0.190 ***
SBIR	-0.011	0.001	-0.010	0.001
ScopeFirms	6.520	1.651	5.798	1.635
PatentFirms	0.041	0.002 **		
SBIRFirms			0.020	0.002
Location Quotient	0.245	0.035	0.288	0.035
UnivNIH	-0.001	0.000	-0.001	0.000
Scale[c]	3.943		3.958	
GEE QICu	-3588.170		-3502.573	
Heteroskedasticity Test [d]	42.090 **		36.392 ***	
Multicollinearity Condition Number	35.061		35.173	
Number of Observations	919		919	

[a] The marginal effects for continuous variables are the average marginal effect for all observations. For the VC variable the marginal effect is approximated as the change in the dependent variable resulting going from the 0 to the 1 category.

[b] The standard errors correspond to the GEE estimates.

[c] The scale parameter comes from the model without robust standard errors.

[d] An LM test employing results from an auxiliary regressions was used to test for heteroskedasticity.

*** .001 significance, ** .05 significance, * .10 significance

Note: The log link function was used for the Poisson model.