ABSTRACT

Title of Dissertation:	DO INDUSTRIAL CLUSTERS ENCOURAGE ESTABLISHMENT INNOVATION?
	Li Fang, Doctor of Philosophy, 2018
Dissertation directed by:	Professor Gerrit Knaap, Urban and Regional Planning and Design

Industrial clusters are geographical concentrations of related industries. They foster innovation, job creation and business formation. Previous studies find that firms in clusters on average are more innovative than firms outside. They interpret this as evidence that clusters encourage firms to innovate. This interpretation is misleading because two different mechanisms can lead to the same result. On the one hand, firms in clusters improve innovativeness through knowledge spillovers and network building. On the other, less innovative firms are forced out of clusters by tough competition. Most studies fail to differentiate these two mechanisms. I separate these mechanisms and examine their variations across industries and establishments. I also search for the optimal spatial scale of industrial clusters to maximize their effect on innovation.

In this dissertation, I match establishment data with patent data for the state of Maryland from 2004 to 2013. I improve the methodology of quantifying the causal

relationship between clusters and innovation, and apply this method to employment centers. Employment centers on average encourage establishments to file for 8% to 11% more patents. This effect is maximized within a one- to two-mile radius region. I also compare how much clusters encourage innovation across different industries, and find significant heterogeneity. In Metalworking Technology, the effect of clusters peaks at a three-mile radius region and increases patent applications by 18%. In contrast, in Business Services, the effect is essentially zero, even when it is maximized in a one-mile radius region. These differences can be explained by industrial characteristics, such as the different level of reliance on tacit knowledge. Finally, I examine how industrial clusters shape the originality of small versus large establishments. I find that small in-cluster establishments improve innovation numerically more than large establishments, but their differences are statistically insignificant.

This dissertation can provide guidance to the design of industrial policies. It helps to more precisely evaluate the benefit of cluster policies. Policymakers can also implement cluster policies targeting at the most beneficiary industries and the optimal spatial scales.

DO INDUSTRIAL CLUSTERS ENCOURAGE ESTABLISHMENT INNOVATION?

by

Li Fang

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2018

Advisory Committee:
Professor Gerrit Knaap, Chair
Professor Marie Howland
Professor Casey Dawkins
Professor Peter Reuter
Dr. Sevgi Erdogan

© Copyright by
Li Fang
2018

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 The Concept of Industrial Clusters

Industrial clusters are geographical concentrations of related industries (Porter, 2000). This concept includes two layers: the industrial layer and the geographical layer. The industrial layer states that clusters are formed by a group of related industries. The relationships between industries can be captured in different ways. For example, input-output relationship, which means outputs of an industry become the inputs of another, and co-location relationship, which means some industries tend to always locate together (Porter, 2003). The geographical layer states that clusters are geographical units within which a group of industries is highly concentrated. Empirically, most studies used administrative units to define clusters in geography, but a few treated space continuously (Wallsten, 2001; De Silva and McComb, 2012; Duranton and Overman, 2005; Kerr and Komiers, 2015).

Clusters are an important feature of economic landscape in any city, region, or nation (Sforzi, 1990; Enright, 1993; Breschi, 1995). For example, one of the most famous industrial clusters in the United States is the Silicon Valley. It employs 10 percent information technology workers and obtains 12 percent of the patents in the United State as of 2013[1]. Other top industrial clusters, such as the Route 128 and the Research

[1] http://www.bls.gov/oes/2013/may/naics2_51.htm
http://www.siliconvalleyindex.org/index.php/economy/innovation-and-entrepeneurship

Triangle Park, are also leading hubs for economic development and innovation[2]. Theoretical and empirical works have shown that isolated firms are not preforming as well as those in the clusters (Baptista and Swann, 1996; Fabiani and Pellegrini, 1998) and regions with clusters may see their economies grow faster than others (DRI/McGraw-Hill, 1995; Glaeser, 2000b).

Clusters happen in different ways, both incidentally and strategically. For example, the carpet cluster in Dalton, Georgia, which produces more than 85% of the carpets sold in the United States, was kick-started by one stitching genius Catherine Evans Whitener. Whitener, at the age of 15, crafted a bedspread for her brother's wedding, which caused a national sensation. This sensation caused her to receive orders more than she could fill, so she taught her neighbors the technique; the cottage industry was born. The industry quickly spread and became a local brand. On the other hand, the research triangle, one of the largest technology centers in the nation, was created strategically by the close collaboration between governments, businesses and universities. A non-profit organization, the Research Triangle Regional Partnership was created and has been leveraging on local resources, favorable policies and local collaborations to promote the cluster to the nation and the world. While we have not much control over incidental kick-start of some clusters, supporting existing clusters and strategically grow clusters align with a region's specialty are doable.

---

[2] http://engineered.typepad.com/thoughts_on_business_engi/2012/01/innovation-bostons-route-128-vs-san-franciscos-silicon-valley.html
http://www.ncbi.nlm.nih.gov/books/NBK158811/

Recognizing the role of industrial clusters in promoting local economic development, policymakers have started to see clusters as a new framework to organize economic activities. Boston launched a $1 billion initiative in 2006 to grow a life science cluster [3]. The U.S. Department of Commerce allocated $50 million in 2009 to assist regional cluster initiatives [4]. In 2010, the U.S. Department of Energy awarded $129 million to the Energy Regional Innovation Cluster in Philadelphia[5]. The Maryland Technology Development Corporation invested up to $225,000 to help startups in the Maryland cybersecurity cluster. Thus, it is both academically and practically important to quantify the benefits of industrial clusters on economic development, and particularly on innovation.

1.2 Research Questions

This study answers a general question: Do industrial clusters encourage innovation? And if so, by how much and at which geographical scale? In addition, this study also examines the heterogeneous effect of industrial clusters across industries and establishments, and explains which locational, industrial and establishment characteristics account for the heterogeneity.

Academically, answering these questions can help us better understand the impact of industrial clusters, and the mechanisms through which industrial clusters boost innovation. Practically, answering these questions can provide guidance for the

[3] https://www.clustermapping.us/sites/default/files/files/resource/Clusters_and_Competitiveness-_A_New_Federal_Role_for_Stimulating_Regional_Economies__Full_Report_.pdf
[4] https://www.ncbi.nlm.nih.gov/books/NBK115046/
[5] http://scienceprogress.org/2010/08/a-win-for-regional-innovation/

evaluation and design of industrial policies. We can determine whether a cluster policy is worthwhile, and implement the policy targeting at the optimal geographical scale and the industry that benefits the most from it.

1.3 Motivation

Clusters are important drivers of economic development and innovative activities (Feldman and Audretsch, 1999; Porter, 2003; Feser, Renski, and Goldstein, 2008; Glaeser and Kerr, 2009; Delgado, Porter, and Stern, 2010, 2014). As mentioned above, policymakers invested heavily in clusters to encourage innovation. Among other positive economic outcomes clusters help to achieve, innovation is a major outcome that policy makers and researchers care about, given its importance to long-term economic success (Romer, 1986; Grossman and Helpman, 1993). Though many cluster-oriented projects and policies have been initiated, policymakers till now have not been equipped with sufficient knowledge to make the most efficient policy choices. Four issues remain unresolved.

First, do industrial clusters actually encourage innovation? And if so, by how much? Although a large number of studies have been carried out on this topic, most studies overstate the causal effect of clusters on innovation by confusing learning and selection (Aharonson, Baum, and Feldman, 2004; Baptista and Swann, 1998; Delgado, Porter, and Stern, 2014). While learning reflects the desirable policy outcome that clusters help establishments improve, selection reflects that tough competition in clusters forces out the least innovative establishments and is not necessarily desirable. A few studies that controlled for selection focused on geographical regions outside of the United States

(Vásquez-Urriago, Barge-Gil, Rico, et al., 2014; Falck, Heblich, and Kipar, 2010; Chen, 2011). A more accurate estimation of how much industrial clusters encourage innovation is needed for the United States to make sure that cluster policies are a worthy use of public funds.

Second, what is the optimal geographical scale to form industrial clusters? No consensus to date has been achieved in terms of the optimal geographical scope of clusters. Due to the lack of detailed geographical information, most studies use administrative boundaries to define clusters. Clusters thus are forced to be no smaller than a county (Chrisinger, Fowler, and Kleit, 2015; Feser, Renski, and Goldstein, 2008), and sometimes as large as a metropolitan area or a state (Ellison, Glaeser, and Kerr, 2010; Combes, Duranton, Gobillon, et al., 2012). In practice, policymakers usually target at large regions to grow clusters, such as a whole state or a multi-county region[6]. These large geographical regions, unfortunately, may be well above the optimal scale of clusters. A few recent studies, using detailed geographical data, found that the effects of clusters on innovation and firm mortality are largely concentrated within a one-mile radius region (Wallsten, 2001; De Silva and McComb, 2012). Such studies are still rare and we need a systematic comparison of how much clusters encourage innovation across different scales. Such a comparison helps to determine the appropriate geographical unit to implement cluster policies.

[6] For example, the Maryland TEDCO supported establishments all over the state, without a particular focus on a specific geographical unit. The US Department of Commerce issued one million dollars to help grow technology clusters in Lower Grande Valley Region, Texas—a four-county region.

Third, which industries should we target to grow clusters? Though Porter (2000) suggested that we should pay attention to all industries, in practice, with limited funding, policymakers need to focus resources by targeting at specific industries[7]. How then to choose becomes an important question. Supporting a cluster in a specific industry incurs the opportunity cost of not supporting other industries. Previous studies provide limited guidance on how to prioritize. Most studies do not pay attention to the heterogeneous effect of clusters across industries. A few studies estimate the heterogeneous effect at a highly aggregated level, which is not very helpful for policymaking. For example, studies comparing how much clusters encourage innovation in manufacturing versus service sector provide no guidance for industry targeting within the manufacturing sector. Since the manufacturing sector is large, policymakers often have to target more detailed industries within this sector. Hence, we need a systematic comparison across industries at a less aggregated level.

Fourth, do the impacts of clusters land evenly across establishments? If not, what types of establishments benefit more? Establishments may not benefit uniformly in clusters. For example, if selection exists in clusters, some establishments may go bankrupt. These establishments bear a disproportionate share of the cost of cluster-oriented policies. Fang (2015) found that small in-cluster firms benefit more in terms of innovation than large

[7] Industry targeting is suggested to be a potentially useful tool for economic development by some researchers (among others, see Voytek and Ledebur, 1991; Shields, Barkley, and Emery, 2009; Barkley and Henry, 2005), but seen unfair and unhelpful by others (among others, see Copaken, 1982; Chiang, 1993; Feldman and Francis, 2004). This dissertation does not intend to dismiss the whole line of studies raising concerns over industry targeting, but simply to provide a method that helps with efficient targeting if a community or a local authority decides that they would like to target at certain industries in pursue of economic development. Moreover, although theoretically targeting at a certain industry is unfair, practically, with limited funds, focusing resources is hardly avoidable. If all industries are paid with equal attention, the consequence will be a dilution of resources (Peck and McGuinness, 2003; Shields, Barkley, and Emery, 2009).

ones, by summarizing results from previous studies. But to date, such studies are still rare, and we don't fully understand who are the winners and losers in clusters. For policymakers, it's important to know whether small establishments can benefit from clusters as they are the biggest contributors to innovation and job creation (Rothwell and Zegveld, 1982; Wagner, 2004).

1.4 Contribution

This dissertation makes four contributions. First, it delivers a more precise estimate of causal effect of industrial clusters on innovation by differentiating selection and learning. Second, it identifies the optimal size of clusters by comparing the magnitude of learning across different geographical scopes. It also allows this optimal size to vary across industries. Third, it quantifies the heterogeneous effect of clusters on innovation across a wide range of industries, and explains these heterogeneities with industrial characteristics. Fourth, it compares how much industrial clusters encourage originality in small versus large establishments.

Based on the results of this dissertation, more informed cluster policies can be designed to encourage innovation. We can be sure that cluster policies are heading towards the right direction if learning effects are economically large. We can also implement cluster policies at the optimal geographical scale. In addition, public funds can be distributed more efficiently by properly prioritizing among industries and targeting at industries that benefit the most from cluster. Further, once the heterogeneous effect of clusters on small and large establishments is identified and evaluated, we can help establishments of different sizes with different strategies.

Chapter 2: Agglomeration and Innovation: Selection or Learning?

2.1 Introduction

Over the past decades, researchers have devoted considerable attention to quantify the relationship between agglomeration and innovation. Agglomerations are dense employment or establishment centers. In theory, agglomerations enable firms to share tacit knowledge (Gertler, 2003), build personal networks (Bell, 2005), cultivate an innovative atmosphere (Saxenian, 1994), and save costs (Helsley and Strange, 2002). These expand the capacities of firms to make breakthroughs (Duranton and Puga, 2004). Prior studies also empirically established a positive correlation between agglomeration and innovation (Carlino and Kerr, 2014; Packalen and Bhattacharya, 2015).

However, the estimated effect of agglomerations on innovation usually is not causal and suffers from the selection bias: less innovative firms are more likely to be forced out of agglomerations. This occurs because (a) competition in the input or output market is tougher in agglomerations (Baldwin and Okubo, 2006), and (b) high-skilled labor in major employment centers doesn't match the need of less innovative firms (Combes, Duranton, Gobillon, 2008; Behrens, Duranton, and Rober-Nicoud, 2014).

A few prior studies dealt with the selection bias in the agglomeration-productivity relationship. For example, Combes, Duranton, Gobillon et al. (2012) employed a continuous quantile estimator and found no evidence that market competition forces out the least productive firms in denser employment areas in France. Arimoto, Nakajima, and Okazaki (2014) applied similar method in Japanese silk-reeling industry and identified the selection effect. Similar methods can be applied to study the agglomeration-innovation relationship.

This chapter takes on this job. Using the theoretical predictions in Combes, Duranton, Gobillon et al. (2012) and a quantile estimator, I place bounds on the learning effect (establishments improve innovation in agglomerations) and the selection effect (less innovative establishments are less likely to survive in agglomerations) at different percentiles of establishment innovation. I adopt a unique population-wide establishment-level[8] dataset from Quarterly Census of Employment and Wages for the state of Maryland, 2004-2013, and the patent application data from United State Patent and Trademark Office. I find that a one-mile radius area with above-median employment concentration increases citation-weighted patent applications by 7.8% to 11.4% for an average establishment that applied for patents during the study period. I also find a sizable selection effect, with non-innovators 2.5% less likely to survive in employment centers. These results are qualitatively robust across time periods and alternative specifications and measurements.

[8] An establishment is a plant. For a large firm with multiple plants in Maryland, each plant counts as an establishment.

The contribution of this chapter is three-fold. First, it separates selection from learning, and more precisely estimates how much agglomerations encourage establishment innovation. Second, this chapter speaks to the geographical dimension of agglomerations. To date, most studies define agglomerations by predetermined administrative boundaries (Ellison, Glaeser, and Kerr, 2010; Combes, Duranton, Gobillon et al., 2012), with no evidence to support that these are the appropriate geographical scopes at which agglomerations prevail. This chapter uses a unique, detailed plant-level dataset and locates every plant on map. It is thus able to measure the exact proximity between establishments and compare the effect of agglomerations across a wide range of geographical scopes. By doing so, it shows that a one to two-mile radius area in Maryland with above-median employment maximizes the effect of agglomerations on innovation. This can be a size for policymakers to target. Third, this chapter also contributes to the Marshall-Jacobs-Porter debate (among others, see Marshall, 1890; Jacobs, 1969, 1984; Porter, 1990; Glaeser, Kallal, Scheinkman et al., 1991). It finds that diversity, especially related diversity, is the most important contributor in the effect of agglomerations on innovation. It also allows the effects of localization, diversity and competition to vary across geographical scales, similar to Andersson, Larsson, and Lundblad (2015), and finds that while localization and competition prevails only within one-mile-radius areas, diversity expands beyond ten miles in radius.

2.2 Literature

The theory of agglomerations dates back to Marshall (1890). He emphasized input sharing, labor market pooling and knowledge spillovers as main benefits of agglomerations. Since then, agglomeration has been widely recognized as one of the main drivers of regional economic growth (Glaeser, 2000), and its benefits expand to diversified business environment (Jacobs, 1969, 1984), demand linkages (Krugman, 1991), localized competition (Porter, 1998), and learning and innovative atmosphere (Henderson, 1974; Gertler, 2003). As beneficiaries, co-located establishments tend to be more productive, more innovative and more competitive.

Empirical studies found that co-located establishments on average innovate more than isolated establishments. For example, Aharonson, Beam and Feldman (2004) found that biotechnology firms in Canada are eight times more innovative in clusters. Similarly, Baptista and Swann (1998) reported that UK manufacturing firms are considerably more innovative when locating in clusters. Negative and insignificant effects have also been identified (Wang, Lin, and Li, 2010; Ferrand, Kelton, Chen et al., 2009), but less frequently. These studies, however, potentially overestimated how much firms improve innovativeness in clusters or agglomerations (the learning effect), as mentioned above, because of the selection bias.

A few researchers attempted to tease out selection bias using instrumental variables and quasi-experimental design, but none has convincingly eliminated the bias. For example, Vásquez-Urriago, Barge-Gil, Rico, et al. (2014) applied the propensity score matching and the instrumental variable approaches, and used the number of companies in a Spanish science and technology park as a share of total companies in a region as the instrument. They found that Spanish science and technology parks increase product innovation by 9.75 percent, but the instrument does not appear to satisfy the exclusion restriction. Falck, Heblich, and Kipar (2010) employed a triple-difference design and found that cluster-oriented policies in Germany increase the likelihood of innovation by 4.6 to 5.7 percent.

But the cluster policies are likely to be endogenous and have favored more innovative regions in the first place.

While the learning effect in agglomerations is surely important to know, the effect of selection may also be of interest. It sheds light on establishment dynamics, and job creation and destruction. However, the selection effect has rarely been empirically estimated, except a few studies measuring selection with firm survival and location choices (Staber, 2001; Rosenthal and Strange, 2003; De Silva and McComb, 2012).

The Marshall-Jacobs-Porter debate also touches on the effects of learning and selection. Marshall (1890) suggested that localization, defined as the concentration of own-industry employment or establishments, is the source of agglomeration benefits. Jacobs (1969, 1984), on the contrary, endorsed the diversity of industries as the major contributor to productivity and innovativeness in agglomerations. Later, Porter (1990) pointed out that the intense competition (i.e., selection) in clusters motivates firms to “keep up with the Joneses”. To date, empirical studies have found evidence for both the localization and the diversity arguments (Rosenthal and Strange, 2004), but the Porter hypothesis has been less intensively studied. This chapter contributes to this line of research by quantifying the effect of competition in uplifting innovation, and compares the role of competition, localization and diversity.

2.3 Data and Method

2.3.1 Data

This dissertation studies the state of Maryland over the period 2004 to 2013. Maryland is a state with strong clusters in information technology, pharmaceutical and education industries, active innovation practices, and diligent economic development endeavors. It ranked third among U.S. states in terms of innovation[9]. Crucial state players, such as the governor, the Department of Business & Economic Development, universities, research institutions and local independent organizations[10] have closely collaborated to strengthen clusters and spur innovation. These characteristics make Maryland an ideal setting to study the relationship between clusters, agglomerations and innovation.

I measure agglomerations with the concentration of employment within a local region (the size varies), and measure innovation with the number of citation-weighted patent applications. This ignores unpatented innovation, but thus far, the only measurements for innovation that are comparable overtime and across geographical spaces are patents and US Small Business Innovation Research Program Awards, while the latter only applies to small businesses[11]. Thus, I use the patent data from the United States Patent and

---

[9] https://www.fastcompany.com/3007772/united-states-innovation-ranking-states-and-district-innovation

[10] http://commerce.maryland.gov/Documents/ResearchDocument/CybermarylandReport.pdf
http://tedco.md/about-us/who-we-are/

[11] Number of patent applications weighted by citation is my measurement for innovation in this disseration. In the literature, there are several different ways to measure innovation: 1) Patents (Beaudry, 2001; Beaudry and Breschi, 2003; Aharonson, Beam and Feldman, 2004). 2) New product announcements (Acs and Audretsch, 1988). 3) R&D expenditure or employee (Baten, Spadavecchia, Streb et al., 2007; Smith, Broberg and Overgaard, 2002). 4) Stock market values of inventive output (Pakes, 1985). 5) Small business innovation research award (Gilbert, McDougall, and Audretsch, 2008). Each of these measurements has its advantages and disadvantages, but 2) is not available over time, 3) is not available in my establishment dataset, 4) only applies to public traded companies, and 5) only applies to small and medium-sized companies. Thus, I use 1) as my measurement.

Trademark Office (USPTO) and apply different weighting schemes to overcome some of its problems. The establishment data are obtained from the restricted version of the Quarterly Census of Employment and Wages (QCEW).

The USPTO documents all patent applications and includes information about the date of application, the technology class, names of the inventor and assignee, and their locations (country, state and city). I tackle with several concerns about this data. First, a patent filed by an establishment sometimes is assigned to the firm headquarter (Blind and Grupp, 1999; Teichert, 2013). This distorts the spatial distribution of innovation. Following Criscuolo and Verspagen (2008) and Deyle and Grupp (2005), I address this concern by assigning an application to the address of the inventor rather than that of the assignee, whenever such information is available, as these studies pointed out that inventors are more likely to be local branches that initiated the inventions. Second, patents differ in quality (Harhoff, Narin, Scherer et al., 1999). This concern is mitigated by weighting patent applications with subsequent frequency of citations[12], following Harhoff, Narin, Scherer et al. (1999) and Hall, Jaffe, and Trajtenberg (2000). I also experiment with six alternative weighting schemes in the robustness tests. Third, the exact date of patent filing is uninformative, as a patent can be filed up to 12 months after the invention is first introduced to the market or otherwise disclosed[13]. Thus, I dismiss the exact date and measure patent filing annually. The total number of patent applications

[12] Subsequent citation frequency is calculated by the total number of subsequent citations plus one (to avoid zero weight) divided by patent age. Number of citations (till the end of March 2016) is collected from the USPTO. This weight is also alternatively calculated by excluding self-citations and relative to each technology class and application year. These weights are standardized to mean one.
[13] https://www.uspto.gov/web/offices/pac/mpep/s2133.html

traced to Maryland establishments during the study period is 10,355. Co-inventors are assumed to share patents equally.

The QCEW publishes quarterly counts of establishment, employment and wages reported by employers, which cover 98 percent of jobs in the United States. The publicly available data are aggregated to the level of counties, while the restricted version of the data contain micro-level information of every plant. I have the luxury of the restricted data which disclose the following information for every establishment: name, address, age, size, wages and six-digit North American Industry Classification System (NAICS) code. These data are quite unique and advantageous: they allow a continuous and flexible treatment of space. With the exact address of every establishment, I can map and measure the proximity between establishments and employment. Thus, the measured agglomerations can be independent of jurisdictional divisions (Duranton and Overman, 2005) and overcome a major problem in prior studies with aggregated data. These data also permit a flexible change of the geographical scale at which agglomerations are defined and measured, and thus can potentially reveal the optimal geographical scope for the prevalence of agglomeration benefits. To my knowledge, prior studies seldom have access to such desirable data, and a few with similar data showed that the agglomeration effect largely concentrates within a one-mile radius or smaller local area that cannot be revealed with aggregated data (Duranton and Overman, 2005; Wallsten, 2001; De Silva and McComb, 2012; Carlino, Carr, Hunt et al., 2012; Arzaghi and Henderson, 2008). To my knowledge, there are few studies that have adopted the restricted QCEW data to conduct analysis related to urban agglomerations or industrial clusters. Examples are De Silva and McComb (2012) and Renski (2013). The former used the QCEW establishment data for the state of Texas to examine firm survival in clusters and the latter adopted the data for Maine to study labor mobility and knowledge spillover in clusters.

For consistency, the establishment data are also analyzed annually14. The total number of establishment-year observations is 1,503,114, and 94 percent are successfully mapped. Only mapped establishments are used in the analysis. The excluded establishments may cause a bias if they are industrially or geographically concentrated. To address this concern, I manually checked a subset of 900 randomly picked non-mapped establishments (approximately a one percent sample), and found them spatially and industrially dispersed.

I match the USPTO and QCEW datasets by establishment name and location at the city level. Since the patent application records are read from patent filing forms, the name of the same establishment may be formatted differently in each filing and there exists many misspellings. Thus, I first standardize the format of establishment name. For example, 180s, inc., 180S, INC., 180's, Inc. are all changed to 180s, inc. Second, I correct for obvious misspelling of establishment name. Then I group patent records and establishment records by year and city. Within each group, I match records by establishment name using the fuzzy string matching method. Last, after the computer match, I also manually check for the matched and unmatched patent records. This employs my discretion to avoid mismatch and form additional patent-establishment pairs with misspelling in either the patent or the establishment record. Eventually, I managed to match 93% of the patent filings to the establishment sample. In the absolute term, I

[14] I use the first quarter whenever possible, and use the second quarter in 2009, 2012 and 2013 when the first quarter data are unavailable to me.

link 10,355 patent records to 9,016 establishment-year observations. Note that this means the cast majority of the observations (99.4%) for patent applications is zero, and I can only identify the effect of agglomerations on innovation at the right tail of establishment innovation distribution with non-zero observations.

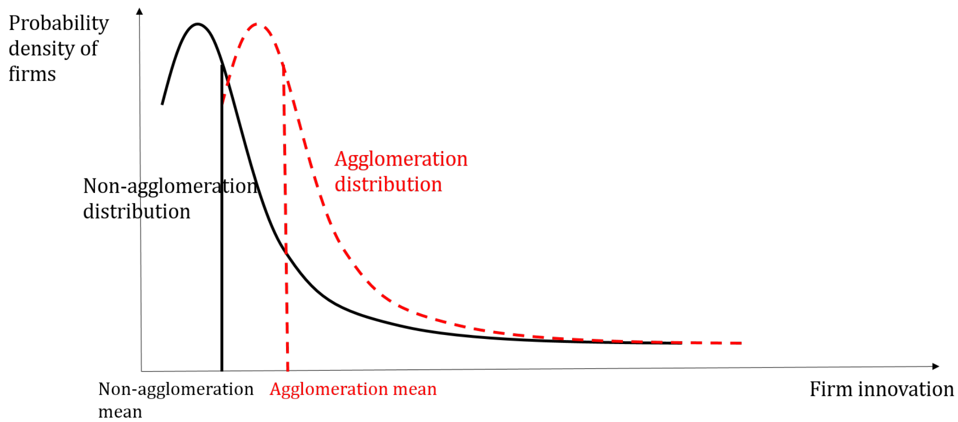
The rest 7% of patent records are excluded from the analysis. The exclusion of these patent filings is unlikely to bias the results due to the following reasons. First, the number of unmatched patent applications is small. A seven percent omission of the data shouldn't have a large impact on the results. Second, eight percent of the unmatched applications are invented by a person instead of an establishment. These applications are legitimately excludable from the analysis of establishment innovation. Third, the rest of the unmatched patent filings does not appear to be spatially concentrated (according to the city level address) or industrially concentrated (according to the technology class).

However, a measurement issue surfaces when the combination of establishment name and the city fails to identify a unique pair of establishment and patent filing. This happens if a patent is applied by a firm with multiple plants (and thus share the same name) located in the same city. In such cases, I assign the patent equally to these plants. Though obviously an imprecise measurement, this issue arises for only 31 firms, a fairly small proportion of the sample (less than 0.01 percent), and thereby unlikely to exhibit a quantitatively relevant impact on the results.

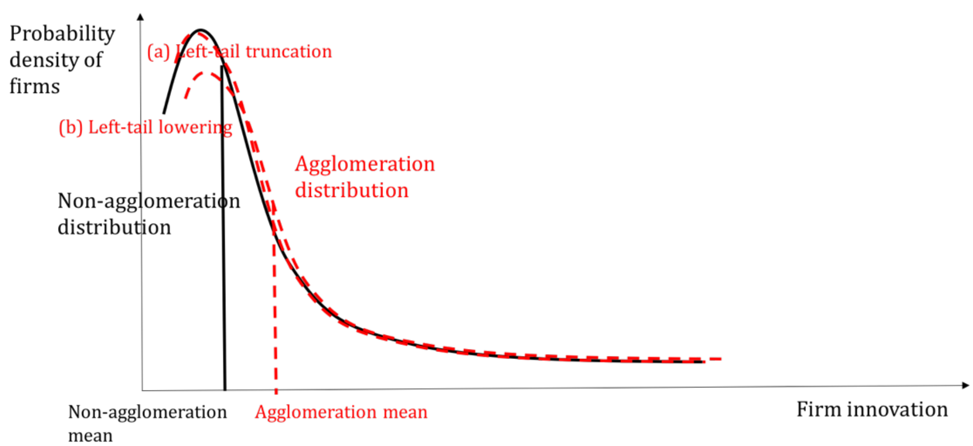
2.3.2 Method

This chapter measures agglomerations as local regions with above-median employment density, following Combes, Duranton, Gobillon et al. (2012). This implicitly assume that employment in all industries ubiquitously encourages innovation, which is later tested against the assumptions that only same-industry and related-industry employment encourage innovation, respectively. This chapter studies industries with at least 1% establishments filing for patents during the study period to ensure a minimum frequency of observations on patent filings. Different from the Combes paper and most prior studies, in this study, a local region is defined as a circular area around every establishment in the above industries rather than an administrative unit. The area is defined with a range of flexible radii: half-a-mile, one mile, two miles, five miles, and ten miles. This continuous and flexible treatment of space, as mentioned above, has two advantages: 1) it allows an establishment located at the border of an administrative unit to be affected by multiple neighboring units; and 2) it enables a comparison of the cluster or agglomeration effect across geographical scopes and overcomes the over-aggregation problem (Burger et al., 2008). The five radii applied in this chapter are chosen based on radii used in prior studies (De Silva and McComb, 2012; Wallsten, 2001; Arzaghi and Henderson, 2008), so that the results can be compared across studies. Though experimented with different radii, this study primarily focuses on the one-mile radius, within which signs of both learning and selection have been empirically detected (Wallsten, 2001; De Silva and McComb, 2012). If an establishment locates near the border of Maryland and a circle around it expands beyond the state boundary, that establishment will be excluded from the estimation. This is because I do not observe establishments in other states and therefore cannot count employment density in such a circle. This is the method adopted throughout this dissertation.

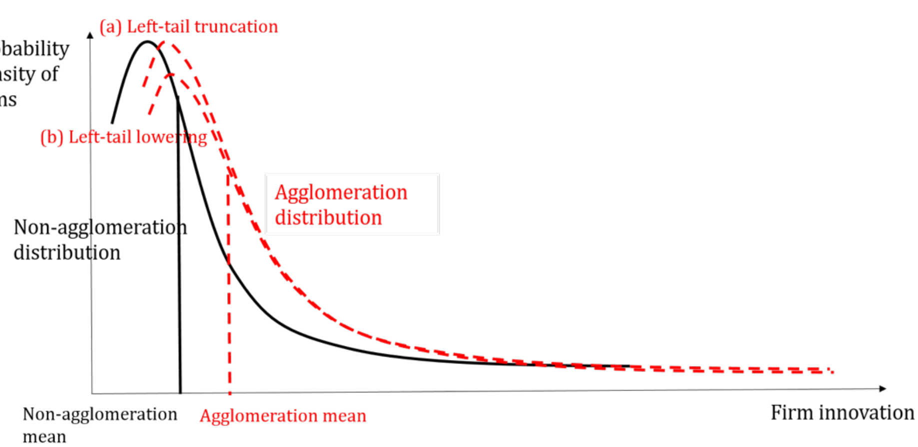
This chapter disentangles selection from learning. This is dealt with by estimating the distribution of establishment innovation, instead of the mean. According to Combes, Duranton, Gobillon et al. (2012), while the effects of learning and selection both upshift the average establishment innovation in agglomerations and therefore cannot be separated in an OLS regression, the distributional implications are different. The learning effect improves every establishment's innovation; thus, it right-shifts the distribution of establishment innovation, as shown in Figure 2.1, panel a. On the other hand, selection forces out the least innovation establishments in agglomerations, or lowers their chance of survival. As a result, it truncates or lowers the left tail of the innovation distribution, as shown in Figure 2.1, panel b. This two cases, as well as their combination as a third case (Figure 2.1, panel c), can be separated with an estimation of the establishment innovation distribution.



a. Learning: Establishment innovation shifts to the right in agglomerations.



b. Selection: The left tail of establishment innovation truncates or lowers in agglomerations.



c. Learning & selection: Establishment innovation both right-shifts and left-truncates in agglomerations.

Figure 2.1 Establishment innovation in and out of agglomerations under different mechanisms

I was going to stratify the sample into groups by quantiles and then run linear regressions on each quantile (group) of observations. However, doing so would lead to a sample selection bias since this procedure is equivalent to artificially select subsamples and discarding the rest of observations in each linear regression. This results in bias in the estimation. To deal with this problem while effectively estimate the distributional effect, economists have developed the quantile regression method (Koenker and Bassett, 1978). This method can be viewed as a natural way to extend the idea of estimating conditional mean in OLS regression to estimating conditional percentile. Unlike stratified OLS, quantile regression delivers consistent and asymptotic normal estimators.

In this case, a quantile regression estimates the difference in establishment innovation between agglomerations and non-agglomerations across a range of percentiles, contrary to the single mean estimate in an OLS regression. These quantile-specific estimates systematically compare the agglomeration with non-agglomeration distributions, and reveal selection at the left-tail and detect learning by the overall right-shift. This approach is implemented with the following specification.

$$logPat_{it} = \alpha_{1p} + \alpha_{2p} * Agg_{ijt} + \boldsymbol{\alpha_{3p} X_{ijt}} + \boldsymbol{\alpha_{4p} Z_{it}} + \boldsymbol{a_{tp}} + \varepsilon_{ijt} \quad (2.1)$$

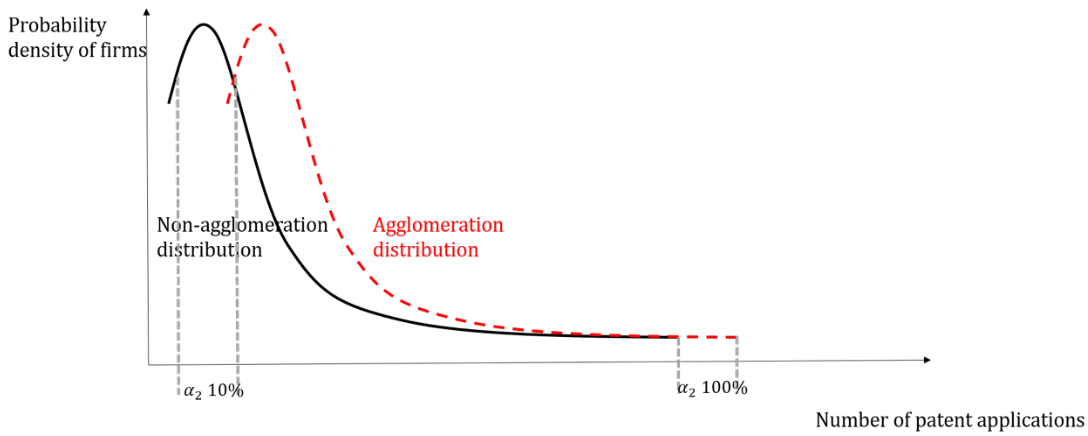
where Pat_{it} denotes one plus the number of per-year citation-weighted patents filed by establishment i year t. Agg_{ijt} is a dummy indicating whether establishment i locates in an agglomeration in year t, i.e., centering a region j with above-median employment density. Region j, as defined above, is a circular area around establishment i with a radius of half a mile, one mile, two miles, five miles, or ten miles. $\boldsymbol{X_{ijt}}$ denotes a set of locational characteristics. $\boldsymbol{Z_{it}}$ denotes a set of establishment characteristics. $\boldsymbol{a_t}$ denotes year fixed effects and ε_{ijt} denotes the random error.

For $\boldsymbol{Z_{it}}$, I include fixed effects indicating ownership types and six-digit NAICS industries. They are determinants of establishment innovation (Shefer and Frenkel, 2005; Hervas-Oliver, Sempere-Ripoll, Boronat-Moll, 2014). Other establishment characteristics correlated with establishment innovation such as establishment size (Simonen and McCann, 2008) and human capital (Winters, 2014), are also by-products of the agglomeration and innovation processes. They are excluded from equation (2.1) to maintain a causal interpretation of α_2, but included in an extended specification.

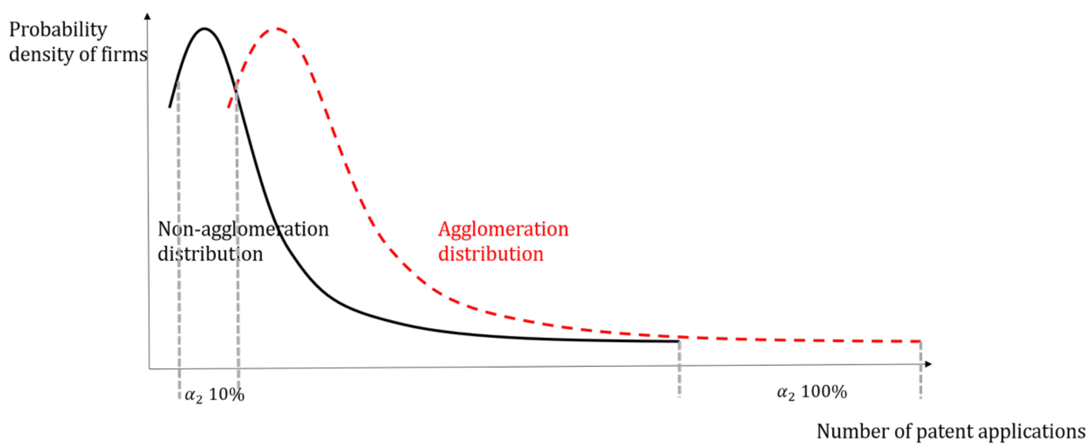
For $\boldsymbol{X_{ijt}}$, I include county fixed effects to control for the importance of political boundaries (Singh and Marx, 2013). In an extended specification, instead of using the aggregate Agg_{ijt} dummy to capture the agglomeration status, I break down the concept of agglomeration into the indices of localization, diversity and competition. This allows a comparison of which aspect of agglomeration accounts for more of its benefits on establishment innovation. Localization is measured by the concentration of employment/establishments in the same industry of establishment i. Diversity, on the other hand, measures the distribution of employment across industries. Competition measures the extent to which a location is dominated by a few large establishments. Localization, diversity and competition are found to greatly influence innovation in prior studies (Glaeser, Kallal, Scheinkman et al., 1991; Beaudry and Schiffauerova, 2009).

These indices are constructed as follows. The raw localization index is measured by the number of same-industry employment/establishments. A larger index indicates a greater same-industry concentration. The raw urbanization index is measured by an industry Gini index, capturing the inequality of employment distribution across six-digit NAICS industries. A Gini index closer to zero indicates greater industrial variety. The raw competition index is measured by the Herfindahl index, calculated by the sum of squared proportion of employment in three largest same-industry establishments, as a percent of total same industry employment in location j. A smaller Herfindahl index indicates tougher competition. These raw indices are standardized to ensure comparability of their regression coefficients. Alternative measurements of the diversity index are applied in the robustness tests.

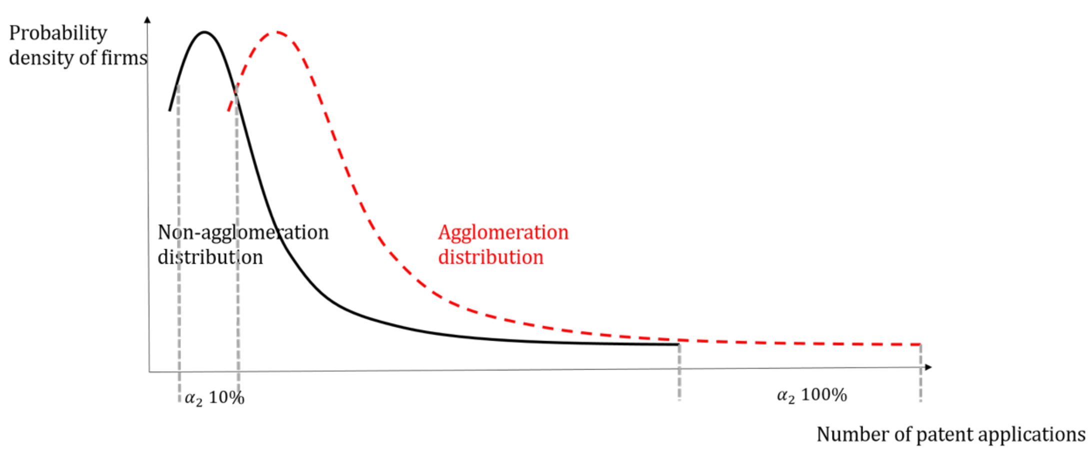
α_{2p} denotes coefficients estimated as percentile p. Variations of $\hat{\alpha}_2$ across quantiles separate selection and learning. If all establishments in agglomerations improve innovativeness by the same magnitude, $\hat{\alpha}_2$ stays a constant across percentiles (Figure 2.2, panel a). If more innovative establishments improve more, $\hat{\alpha}_2$ grows with quantiles (Figure 2.2, panel b). By contrast, with selection, $\hat{\alpha}_2$ is the largest at the lowest percentile and then weakly decreases by quantiles; it is zero at the 100 percentile (Figure 2.2, panel c).



a. All establishments learn the same: $\hat{\alpha}_2$ remains constant across percentiles



b. More innovative establishments learn more: $\hat{\alpha}_2$ increases with percentiles



c. Selection: $\hat{\alpha}_2$ decreases with percentiles

Figure 2.2 Identification of different mechanisms through the variations of $\hat{\alpha}_2$ across percentiles

Moreover, $\hat{\alpha}_2$ across different percentiles bound the magnitudes for learning and selection. $\hat{\alpha}_2$ has two components: the learning effect and a selection bias[15]. The selection bias, for the moment, is assumed to be caused solely by the fact that less innovative establishments are more likely to be forced out of agglomerations. The two components vary across percentiles with different patterns: The selection effect shrinks with the increase of quantiles, while the learning effect stays the same or escalates with percentiles. Thus, if $\hat{\alpha}_2$ at percentile p* is the smallest among all nonzero $\hat{\alpha}_2$, it bounds selection from below for percentiles p<p*, and from above for p>p*. Similarly, it bounds learning from above for p<p*, and from below for p>p*. A more detailed proof is provided in the appendix 1. The sharpness of these bounds depends on the magnitude of $\hat{\alpha}_2$ at p*; the smaller it is, the sharper the bounds are. Empirically, for the purpose of this chapter, these bounds are sharp enough to reveal the serious bias of an OLS estimator, but not to completely rank the magnitude of the learning effect across different geographical sizes.

Now, the bounded estimates for learning may still be contaminated by the self-selection bias. More innovative establishments may be systematically attracted by agglomerations (Baldwin and Okubo, 2006; Behrens, Duranton, and Rober-Nicoud, 2014), although they survive both agglomerations and non-agglomerations. This concern is especially relevant if $\hat{\alpha}_2$ increases with quantiles, which reveals a motive for self-selection: the highly creative establishments benefit the most from agglomerations.

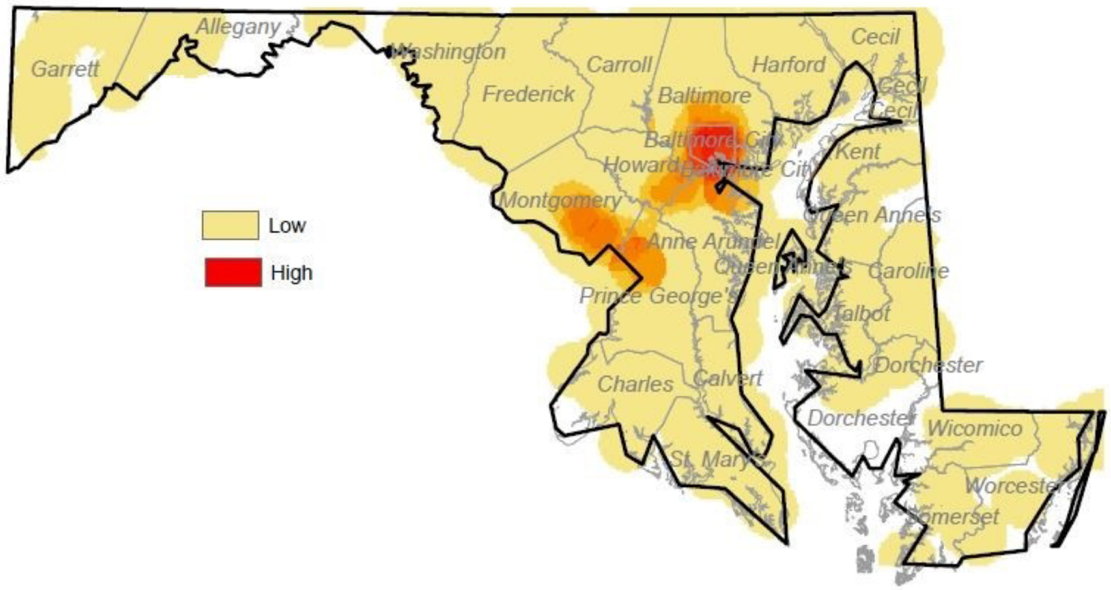
[15] See Angrist, Chernozhukov, and Fernández (2006).

The self-selection bias is eliminated by comparing the full sample result with the new establishment sample result. New establishments are establishments recorded in location j for the first year, which includes two types of establishments: the newly-borns and those that moved to location j for the first year. There are two types of agglomeration effects: static and dynamic effects (Neffke, Henning, Boschma et al., 2011). The static effect is immediately available for all establishments in agglomerations, including new establishments. These effects include sharing the local labor supply and lowered transportation cost. The dynamic effect is accumulative in nature. Being in the agglomeration itself does not suffice. A establishment needs to interact repeatedly with other establishments and workers to gain these benefits. These effects include knowledge spillover, network formation and trust building. Here I maintain the assumption that what matters for establishments to innovate in agglomerations, especially in local agglomerations as this chapter measures (one mile in radius), is the dynamic effect. Labor supply is shared in a region much larger than one mile in radius, and transportation costs (of products) play a relatively small role in establishment innovation. On the contrary, knowledge spillover and network building are essential in innovation, and these benefits all take time to reveal. As a result, they are almost absent in the new establishment sample since these establishments were at location j for only a short period of time. In contrast, innovative establishments that are self-selected into agglomeration locations would not be affected by the absence of these agglomeration effects. They can hit the ground running without assimilating these benefits. Thus, the estimates from the full sample include the learning effect and the self-selection bias, while those from the new establishment sample are mostly self-selection bias. I then can back out the learning effect by subtracting the latter from the former. Note that the static effect may still play some role, though relatively small, in fostering establishment innovation. These benefits can be captured by the new establishments. In that case, I'm backing out a conservative estimate for the learning effect.

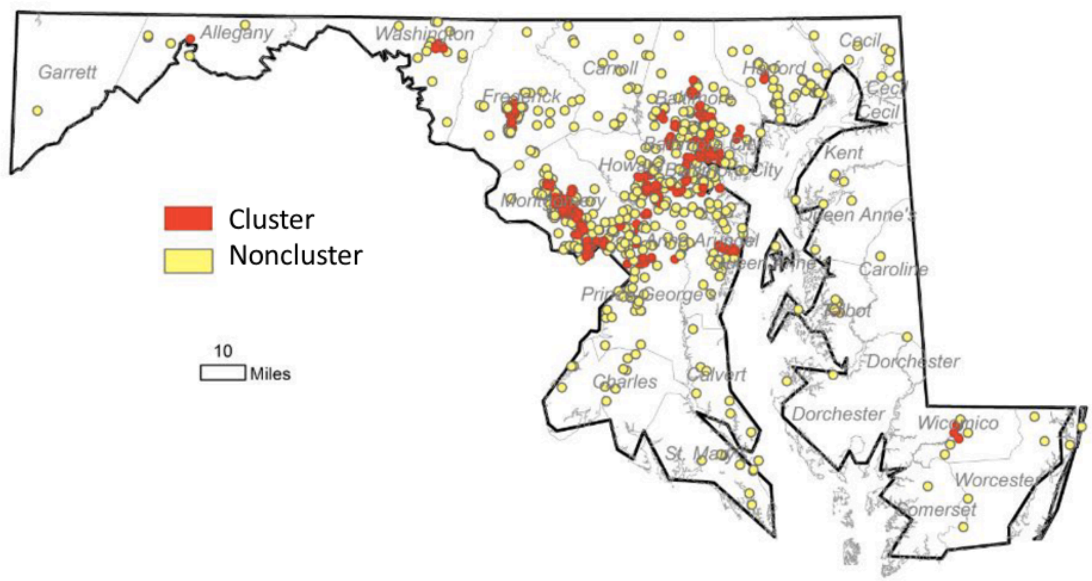
2.4 Results

2.4.1 Descriptive statistics

The employment density in Maryland is shown in Figure 2.3. Panel A maps the density continuously, and jobs are highly concentrated in the central areas. In comparison, panel B maps employment density discretely and in a "local" sense, with agglomerations and non-agglomerations defined as one-mile radius circles with above- and below-median employment density, respectively, around every establishment in industries with over 1% establishments filing for patents. Consistent with panel A, most agglomerations locate in central areas, but peripheral counties are not completely left out: Several agglomerations reside in their jurisdictions. The fact that most agglomerations locate in Montgomery, Howard and Baltimore Counties may bias the estimates if the county effects are not controlled for; thus, I control for county fixed effects in the regressions. Also, since these counties are populated with both agglomerations and non-agglomerations, the county-level characteristics should not be a major concern; they likely cancel out.



Panel A. Employment density



Panel B. Agglomerations and non-agglomerations

Figure 2.3 Maryland employment density, agglomerations and non-agglomerations in 2013

Table 2.1 contrasts establishments in and out of agglomerations to detect their differences in employment size, wages and innovation. Significant differences are found in all these aspects, consistent with previous studies (Baptista and Swann, 1998; Aharonson, Beam and Feldman, 2004). On average, establishments in agglomerations apply for 0.12

citation-weighted patents per year, compared to only 0.009 out of agglomerations. At the same time, establishments in agglomerations are twice as much likely to file for patents compared to establishments outside. Establishments in agglomerations are also triple the size of establishments outside and sustain a significant wage premium; these are either outcomes of the uplifted innovation or confounding factors that need to be controlled. Both possibilities are considered in the regression. These numerical summaries also reveal that over 90% of the establishments did not file for patents during the study period. Therefore, this chapter in fact focuses on a few right-tail percentiles of the distribution of establishment innovation where establishments did file for patents.

Table 2.1 Establishments in and out of agglomerations

Variable	Agglomeration	Non-agglomeration
Size	131.900***(761.515)	48.819(163.335)
Wage	1,451,011*** (8,821,925)	440,265 (3,101,964)
Probability of patent filing	0.070***(0.255)	0.032 (0.175)
Number of citation-weighted patent applications	0.116***(2.471)	0.009 (0.191)
Sample size	4692	6488

NOTE: *, ** and *** denote the difference between the two groups is significantly different at the 10%, 5% and 1% level, respectively, under a t-test allowing unequal variances. Standard deviation in parenthesis.

2.4.2 Main results

The effects of agglomerations on innovation at different percentiles are reported in Table 2.3, with descriptive statistics reported in Table 2.2. Panel A shows the baseline model with only the agglomeration dummy. All results start with the 97 percentile, because at percentiles below 97, establishments file no patents. Agglomerations exhibit significant positive effects on innovation at all percentiles, with the magnitudes first decreasing and then increasing. This indicates both selection and learning, consistent with Figure 2.4. At the 97 percentile, establishments in agglomerations file for 15.4% more citation-weighted patents than establishments outside, while this effect wanes to only 6.7% at the 97.5 percentile. Beyond the 97.5 percentile, the effect continuously increases with percentiles and finally scales to 91.6% at the 99.5 percentile. Therefore, the most innovative establishments, by locating in agglomerations, almost double their production of patents. These results remain qualitatively robust after adding year fixed effects in panel B, county fixed effects in panel C, and establishment characteristics in panel D, but the coefficients are generally smaller and the decrease of coefficients persists to higher percentiles (98 or 98.5). These indicate weaker learning effects, but not necessarily weaker selection effects; the magnitude of selection at each percentile declines, but it prevails at more percentiles.

Table 2.2 Descriptive statistics of key variables

Variable	Short name	Mean	Min	Max	Standard deviation	Sample size
Log (1+citation-weighted number of patent applications)	logPat	0.011	0	4.718	0.143	11,178
The agglomeration status	Agg	0.420	0	1	0.494	11,178
Log (1+establishment employment size)	LogSize	2.552	0	9.635	1.676	11,178
Log (1+wage per employee)	LogWage	4.005	0	13.383	4.496	10,256
Localization index	Loc	-3.84e-09	-0.318	3.373	1	11,178
Diversity index	Gini	-5.69e-09	-6.968	1.797	1	11,174
Competition index	Herf	-5.95e-09	-1.757	0.770	1	11,178

NOTE: Indices are standardized.

The goodness of fit is measured by both Pseudo R^2 and the sum of absolute weighted deviations. Pseudo R^2 is the analogy of R^2 in OLS regressions, but it may not have the range zero to one, and therefore does not have the nice interpretation as the percentage of variations explained by the model. Nonetheless, it still indicates how well the model fits the data. Note that in Table 2.3, at lower percentiles, the fit is not quite good. This is not unusual for disaggregated establishment data from various industries and locations, as important industrial and locational characteristics are controlled by fixed effects but not calculated in Pseudo R^2. An alternative measurement for goodness of fit in quantile regressions is the sum of absolute weighted deviations, which shows the remaining unexplained deviations of the model. The smaller the number, the better the fit. To put the numbers of the sum of absolute weighted deviations in Table 2.3 into perspective, the original sum of deviations of the data is in the neighborhood of 145. As a result, the percentage of deviations explained by the models varies approximately from 2% to 44%.

Table 2.3 The effect of agglomerations on patent filing at different percentiles

	LogPat					
	97	97.5	98	98.5	99	99.5
	Panel A: Baseline model					
Agg	0.154***	0.067***	0.114***	0.203***	0.560***	0.916***
	(0.028)	(0.020)	(0.028)	(0.053)	(0.104)	(0.188)
Pseudo R^2	0.0223	0.0209	0.0223	0.0329	0.0685	0.1405
Sum of absolute weighted deviations	143.790	140.425	134.316	125.589	111.173	82.130
Observations	11,178	11,178	11,178	11,178	11,178	11,178
	Panel B: Year fixed effects					
Agg	0.140***	0.101***	0.095***	0.181***	0.394***	0.885***
	(0.026)	(0.017)	(0.022)	(0.041)	(0.123)	(0.238)
Pseudo R^2	0.0148	0.0170	0.0162	0.0203	0.0427	0.1096
Sum of absolute weighted deviations	144.409	140.442	134.675	126.787	118.990	84.943
Observations	11,178	11,178	11,178	11,178	11,178	11,178
	Panel C: County fixed effects					
Agg	0.106***	0.073***	0.067***	0.073	0.208*	0.808***
	(0.026)	(0.018)	(0.014)	(0.036)	(0.117)	(0.193)
Pseudo R^2	0.0060	0.0096	0.0102	0.0084	0.0169	0.0610
Sum of absolute weighted deviations	144.273	140.083	134.016	126.778	115.402	88.278
Observations	11,178	11,178	11,178	11,178	11,178	11,178
	Panel D: Establishment characteristics					
Agg	0.090***	0.044	0.042	0.028	0.078	0.493**
	(0.024)	(0.028)	(0.010)	(0.024)	(0.092)	(0.201)
Pseudo R^2	0.0002	0.0021	0.0040	0.0025	0.0022	0.0123
Sum of absolute weighted deviations	143.334	139.087	132.667	125.294	114.936	90.752
Observations	11,178	11,178	11,178	11,178	11,178	11,178

NOTE: * p<0.1; ** p<0.05; *** p<0.01. Standard errors in parentheses.

Table 2.4 adds potentially endogenous establishment characteristics. The pattern of remains similar from 97.5 to 99 percentiles, but changes at the lowest (97) and the highest (99.5%) percentiles. The similarity in the middle percentiles confirms that regardless of whether to control for these characteristics, learning exists. But at the lowest percentile, the sign of selection disappears, and at the top percentile, learning diminishes. These are likely driven by two facts: 1) small in-agglomeration establishments are more likely to be forced out due to their lack of financial channels, and 2) large in-agglomeration establishments are more likely to file for an impressive number of patents. Thus, after controlling for the size of establishments, the coefficients on the agglomeration status become smaller. This may or may not imply the results of Table 2.3 to be spurious, depending on whether agglomerations attract larger establishments or establishments grow faster in agglomerations (potentially due to the improved innovativeness). By comparing the new establishments in and out of agglomerations conditional on counties, year and industries, I find that the size of new establishments in agglomerations are actually 4.7% smaller than those out of agglomerations; and the difference is not significantly different from zero. Thus, the results in Table 2.4 could not have been driven by the self-selection along establishment sizes. In contrast, given relatively smaller sizes to begin with, establishments grow 60% faster in agglomerations than outside of.

Table 2.4 The effect of agglomerations on patent filing with the control of additional establishment characteristics

	LogPat					
	97	97.5	98	98.5	99	99.5
Agg	0.022	0.031*	0.042*	0.050	0.061	0.089
	(0.019)	(0.018)	(0.025)	(0.047)	(0.099)	(0.177)
LogSize	0.033***	0.038***	0.045***	0.060***	0.101***	0.173***
	(0.006)	(0.006)	(0.001)	(0.015)	(0.033)	(0.058)
LogWage	-0.0003	0.0002	0.001	0.002	0.002	-0.0006
	(0.001)	(0.001)	(0.002)	(0.003)	(0.007)	(0.013)
Pseudo R^2	0.0434	0.0525	0.0592	0.0766	0.1076	0.1596
Sum of absolute weighted deviations	131.092	125.687	119.223	110.512	97.528	72.559
Observations	10,256	10,256	10,256	10,256	10,256	10,256

NOTE: * $p<0.1$; ** $p<0.05$; *** $p<0.01$. Standard errors in parentheses.

Table 2.5 compares the effects of localization, diversity and competition on innovation to address the Marshall-Jacobs-Porter debate. Panel A measures localization with same-industry employment and Panel B with the number of same-industry establishments. These two measurements turn out to be highly correlated with a correlation coefficient of 0.99. The results mildly support the Jacobs hypothesis, but are inconsistent with either the Marshall or the Porter hypotheses. A more diverse local economy (a smaller Gini index) increases innovation at most percentiles below 99, but hurts top innovators. This is consistent with the majority of prior studies in favor of the Jacob hypothesis (Beaudry and Schiffauerova, 2009), and the exception at the top percentile may be explained by top establishments being less reliant on the outside environment than their interior resources (Agrawal, Galasso, and Oettl, 2017). In contrast, localization demonstrates almost no effects on innovation at lower percentiles and negative effects at 98.5 percentiles and above.

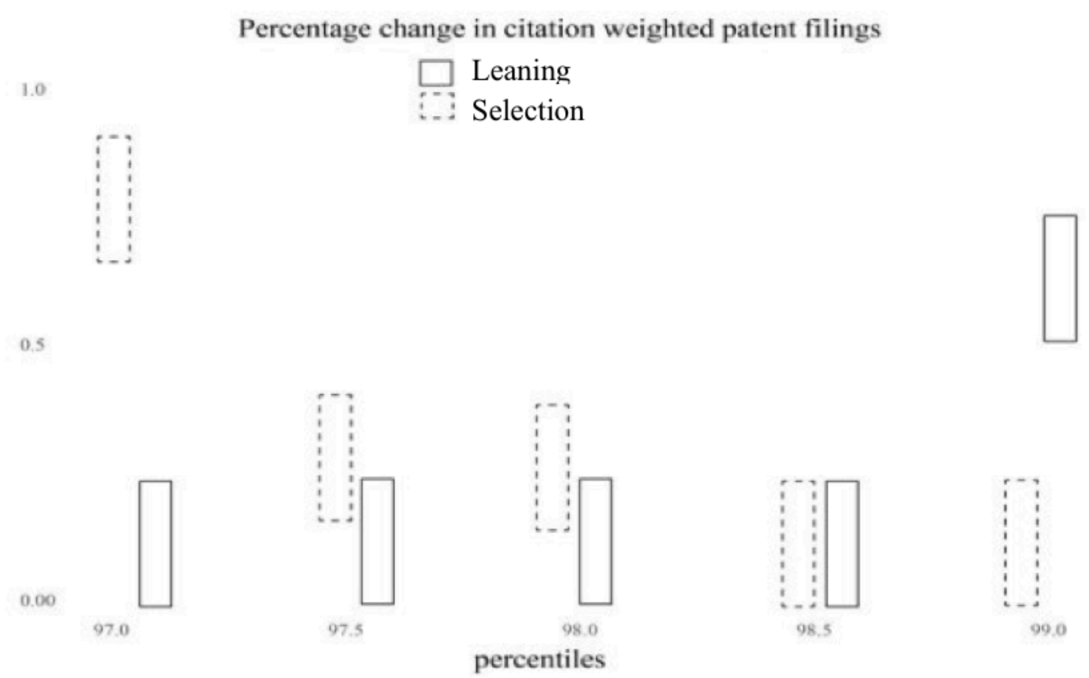
This echoes the findings of Feldman and Audretsch (1999) and Glaeser, Kallal, Scheinkman et al. (1991). Similarly, more intensive local competition (a smaller Herfindahl index) is also associated with fewer patent filings; this holds at all percentiles. This is in line with the predictions of the standard industrial organization theory (Dasgupta and Stiglitz, 1980; Aghion, and Howitt, 1992; Caballero and Jaffe, 1993; Dubey and Wu, 2002), consistent with some literature on the monopoly-innovation relationship (among others, see Swan, 1970; Reinganum, 1983; Geroski, 1994; Aghion, Bloom, Blundel et al., 2002), but clashes with the empirical findings of Feldman and Audretsch (1999) and Glaeser, Kallal, Scheinkman et al. (1991). Later in the robustness tests, I find that measuring diversity with the entropy index instead of the Gini index does not change the result, while the positive effect of diversity on innovation is completely driven by "related diversity" (Frenken, Van Oort, and Verburg, 2007) —diversity of related industries.

Table 2.5 Marshall v.s. Jacobs v.s. Porter

	LogPat					
	97	97.5	98	98.5	99	99.5
	Panel A: Localization measured by employment					
Loc	-4.34e-19	2.17e-19	-0.00009	-0.026	-0.069	-0.143
	(0.0002)	(0.0003)	(0.013)	(0.027)	(0.057)	(0.104)
Gini	-0.0005**	-0.0006*	-0.0002	0.0005	-0.004	0.124
	(0.0002)	(0.0003)	(0.014)	(0.029)	(0.059)	(0.109)
Herf	0.0006***	0.0009***	0.001	0.002	0.014	0.115
	(0.0002)	(0.0003)	(0.015)	(0.030)	(0.062)	(0.114)
Pseudo R^2	0.0006	0.0007	0.0006	0.0080	0.0132	0.0373
Sum of absolute weighted deviations	143.343	139.334	132.857	125.265	113.628	88.601
Observations	11,174	11,174	11,174	11,174	11,174	11,174
	Panel B: Localization measured by number of establishments					
Loc	9.75e-19	0	-0.00009	-0.026	-0.068	-0.145
	(0.0002)	(0.0003)	(0.013)	(0.028)	(0.057)	(0.104)
Gini	-0.0005**	-0.0006*	-0.0002	0.0004	-0.004	0.126
	(0.0002)	(0.0003)	(0.014)	(0.029)	(0.060)	(0.109)
Herf	0.0006***	0.0009***	0.001	0.002	0.014	0.116
	(0.0002)	(0.0003)	(0.015)	(0.030)	(0.063)	(0.114)
Pseudo R^2	0.0006	0.0007	0.0006	0.0082	0.0139	0.0400
Sum of absolute weighted deviations	143.342	139.334	132.858	125.258	113.562	88.360
Observations	11,174	11,174	11,174	11,174	11,174	11,174

NOTE: * p<0.1; ** p<0.05; *** p<0.01. Standard errors in parentheses.

Figure 2.4 places bounds on the learning effect and the selection effect. Using the results in panel D, Table 2.3, the full specification of equation (2.1), I find that learning at best encourages establishments at the 97 to 98.5 percentiles to apply for 2.8% additional citation-weighted patents per year. The effect almost triples at the 99 percentile, and at the 99.5 percentile, an establishment in agglomerations files for at least 46.5% more patents due to learning (results not shown in Figure 2.4). To put into perspectives, an average 99.5 percentile establishment in non-agglomerations files 3.33 citation-weighted patent applications per year, so a 46.5% increase implies 1.55 more patents; in monetary terms, this transforms to a $4.36-$5.04 million addition of firm value, using estimates of stock market return of patents in Hirschey and Richardson (2001). As Dechezleprêtre, Martin, and Mohnen (2016) suggested, the stock market return is likely a lower bound for the value of a patent, because it captures only the private value. To the extent that a patent spurs innovation in other organizations, the social value would exceed the private value. To generalize to all innovating establishments, learning lies between 7.8% and 11.4%, which is still considerable. At the same time, the magnitude of selection is much larger than that of learning at the 97 percentile, and of comparable size at the 97.5 to 98.5 percentiles. These imply a serious bias of an OLS estimator.



NOTE: The result corresponds to panel D, Table 2.3 (the full specification of equation (2.1)).

Figure 2.4 Bounds of learning and selection at different percentiles

To formally evaluate the bias, I run an OLS regression with equation (2.1) and the coefficient on the agglomeration status for all establishments, including those filing for no patent, is 1.41%. In contrast, the weighted averages of the upper and lower bounds imply that learning for all establishments lies between 0.26% and 0.34%. Thus, the bias of the OLS estimator is large, almost quintuples the true effect. Similarly, an OLS estimator on innovating establishments is 18.2%, which almost doubles the true effect between 7.8% and 11.4%. This is partly due to the problem of OLS with the piles of zero dependent variables, not completely driven by the selection effect. But a Tobit estimator that properly deals with the pile-up of zeros, still doubles the true effect.

These analyses show that the selection effect in Maryland agglomerations is sizable. However, remember that the agglomerations in focus are only one-mile in radius. Many sources for market selection are not present in such a small region: Most types of products are competed within a much larger market; labor flows within a much larger geographical region; raw materials usually are also not confined by such a small local region. One possible source for this detected selection effect is the competition for space itself. The benefit of agglomerations may have lured establishments to locate near to each other and bided up the price. Using the CoStar data, I confirm that the average rent per square feet per month for office, industrial and flexible spaces is \$20.59, \$6.03 and \$5.08, respectively, in agglomerations. And in non-agglomerations, it is \$23.07, \$4.14 and \$3.43, respectively. Though agglomerations have a lower office rent, the difference is not statistically significant from zero. In contrast, the higher rent for industrial and flexible spaces is significant at the 5% level under a t-test. This result, though still only suggestive, is consistent with the hypothesis that establishments bid up prices within a local region in order to co-locate with others. This puts pressure on the least innovative establishments and may have forced them out of the desirable locations.

Table 2.6 compares the effect of agglomerations across geographical sizes, using equation (2.1), and Table 2.7 presents the bounds. The results show that selection remains similar in magnitude in half-a-mile agglomerations, and wanes almost completely in larger ones. This reflects the fact that in larger agglomerations, most establishments are no longer direct rivalries for resources. This result is also consistent with De Silva and McComb (2012), who found that in one-mile-radius agglomerations, the share of own-industry employment increases firm mortality, while the opposite is true in larger agglomerations. At the same time, learning on innovating establishments in half-a-mile agglomerations ranges from 4.4% to 5%, in two-mile ones 10% to 10.6%, and it fades completely in beyond-five-mile agglomerations. In comparison, a one-mile-radius agglomeration exhibits an average effect of 7.8% to 11.4%; thus, while the bounded estimates allow us to conclude that a one-mile or two-mile agglomeration outperforms that of other sizes, they cannot rank between these two sizes. This finding is broadly consistent with Wallsten (2001), who also compared the effects across agglomeration sizes and found the largest effect prevails within less-than-one-mile-radius areas. It is also consistent with Rosenthal and Strange (2008a), who found that human capital spillovers decline sharply with distance and almost wane beyond five miles.

Table 2.6 The effect of geographical sizes

	LogPat					
	97	97.5	98	98.5	99	99.5
	Radius=0.5 mile					
Agg	0.022	0.091***	0.055*	0.046	0.006	0.270
	(0.023)	(0.033)	(0.029)	(0.046)	(0.087)	(0.235)
Pseudo R^2	0.0009	0.0045	0.0027	0.0005	0.0001	0.0890
Sum of absolute weighted deviations	106.725	105.662	104.185	100.094	88.360	66.614
Observations	11,170	11,170	11,170	11,170	11,170	11,170
	Radius=2 miles					
Agg	0.004***	0.004***	0.007	0.047	0.069	0.502**
	(0.0004)	(0.0009)	(0.021)	(0.044)	(0.090)	(0.225)
Pseudo R^2	0.0007	0.0005	0.0005	0.0005	0.0014	0.0209
Sum of absolute weighted deviations	122.700	121.981	121.199	119.089	111.570	89.077
Observations	11,170	11,170	11,170	11,170	11,170	11,170
	Radius=5 miles					
Agg	0.004***	0.003**	-0.0006	-0.062	-0.126	0.089
	(0.0006)	(0.001)	(0.023)	(0.043)	(0.085)	(0.240)
Pseudo R^2	0.0005	0.0001	0.0000	0.0038	0.0057	0.0003
Sum of absolute weighted deviations	126.591	125.865	125.038	122.111	114.041	91.907
Observations	11,170	11,170	11,170	11,170	11,170	11,170
	Radius=10 miles					
Agg	0.004***	0.001	-0.002	-0.058	-0.133*	-0.245
	(0.0008)	(0.001)	(0.023)	(0.042)	(0.074)	(0.265)
Pseudo R^2	0.0004	0.0000	0.0001	0.0044	0.0066	0.0052
Sum of absolute weighted deviations	121.061	120.361	119.554	117.029	109.768	88.998
Observations	11,170	11,170	11,170	11,170	11,170	11,170

NOTE: * p<0.1; ** p<0.05; *** p<0.01. Standard errors in parentheses.

These results indicate that what matters for innovation in agglomerations operate in close geographical proximity. One such candidate would be face-to-face communications to share ideas and possibly generate new ones, as face interactions are sensitive to distance due to people's time cost. Worker turnover may also help transmit knowledge, and workers are more likely to land a job in nearby establishments because of their fuller knowledge about these firms and lower search cost. Their knowledge of nearby firms may come from face-to-face communications with workers in those firms, or simply observing other workers in streets and restaurants.

Table 2.7 The bounds of the agglomeration effect and the natural selection bias

	Percentiles					
	97	97.5	98	98.5	99	99.5
	Radius=0.5 mile					
Learning_upper bound	0.006	0.006	0.006	0.006	0.006	0.270
Learning _lower bound	0	0	0	0	0	0.264
Selection_upper bound	0.022	0.091	0.055	0.046	0.006	0.006
Selection_lower bound	0.016	0.085	0.049	0.040	0	0
	Radius=1 mile					
Learning_upper bound	0.028	0.028	0.028	0.028	0.078	0.493
Learning_lower bound	0	0	0	0	0.050	0.465
Selection_upper bound	0.090	0.044	0.042	0.028	0.028	0.028
Selection_lower bound	0.062	0.016	0.014	0	0	0
	Radius=2 miles					
Learning_upper bound	0.004	0.004	0.007	0.047	0.069	0.502
Learning_lower bound	0	0	0.003	0.043	0.065	0.498
Selection_upper bound	0.004	0.004	0.004	0.004	0.004	0.004
Selection_lower bound	0	0	0	0	0	0
	Radius=5 miles					
Learning_upper bound	0	0	0	0	0	0.089
Learning_lower bound	0	0	0	0	0	0.089
Selection_upper bound	0.004	0.003	0	0	0	0
Selection_lower bound	0.004	0.003	0	0	0	0
	Radius=10 miles					
Learning_upper bound	0	0	0	0	0	0
Learning_lower bound	0	0	0	0	0	0
Selection_upper bound	0.004	0.001	0	0	0	0
Selection_lower bound	0.004	0.001	0	0	0	0

NOTE: Negative coefficients are regarded as zero, due to the assumption of non-negativity of learning and selection. See the appendix 1.

I also compare the effects of localization, diversity and competition across geographical scales (results not shown). Consistent with prior studies (Rosenthal and Strange, 2003; Arzaghi and Henderson, 2008), localization prevails only at local regions within half a mile in radius, competition prevails within one mile, while diversity expands beyond broader levels and stays strong in a ten-mile-radius agglomeration.

I also deal with the potential contamination of the self-selection effect. I run equation (2.1) with the new establishment sample and find self-selection negligibly small at all percentiles and in agglomerations of all sizes. They are no more than five percent of the coefficients in Table 2.3 and statistically indistinguishable from zero. Thus, the complication caused by self-selection is minimal and the bounded estimates are trustworthy.

2.5 Robustness Checks

The Combes, Duranton, Gobillon et al. (2012) method. I apply the Combes, Duranton, Gobillon et al. (2012) method to test whether its result agrees with the bounded estimates. The estimated share of left truncation is 0.9%, the average right-shift is 0.31%, and the right dilation is 311%. Note that the average agglomeration effect (right-shift) is in line with the bounded estimates: 0.26% to 0.34%. At the same time, the selection effect estimated by the Combes, Duranton, Gobillon et al. (2012) method is smaller than the lower bound of the bounded estimates (1.5%). This is likely due to the fact that the Combes, Duranton, Gobillon et al. (2012) method captures the left truncation but not the lowering of the left tail, while both are captured in the bounded estimates. Nevertheless, this exercise shows that learning estimated in both approaches are of comparable magnitudes.

Alternative definition of agglomerations. I use the density of establishments instead of employment to define agglomerations. The results are qualitatively robust, while learning and selection are both weakened in magnitudes: a one-mile radius area with above-median establishment concentration increases citation-weighted patent applications by about 7.3% for an average establishment that files for patents during the study period. At the same time, non-innovators are 1.3% less likely to survive in agglomerations.

Alternative measures of diversity. Entropy, as the literature pointed out (Attaran and Zwick, 1987), can be another useful measure for industrial diversity. I therefore use it instead of the Gini index to test for robustness. The results barely change and the entropy-measured index are highly correlated with the Gini index with a coefficient of 0.82. Furthermore, Frenken, Van Oort, and Verburg (2007) suggested that diversity per se does not contribute to learning; it is related diversity, the diversity in the related industries that matters. To probe into that, I separate the diversity index into related and unrelated diversity. They are calculated with the Gini indices within a group of closely related industries16 and across such groups, respectively. As expected, I find the positive effects of diversity on innovation completely driven by related diversity, while unrelated diversity shows statistically zero or even negative effects.

Nonlinear specification. I also use polynomials of the localization, diversity and competition indices up to the power of three, to account for potential nonlinear relationships found in prior studies. Competition and localization exhibit inverse-U

[16] I apply the definition of Delgado, Porter and Stern (2014), who group industries by input-output, co-location and labor-pooling relationships using optimal clustering algorithm.

shaped effects on innovation, i.e., too much or too little are both harmful, while the effect of diversity remains linear.

Alternative weighting schemes of patent applications. In the main analysis, I use per year citation to weight patent applications. Alternatively, I apply six other weighting schemes. First, I assign zero weights on ungranted applications, i.e., I count patent grants instead of patent applications. This reduces the magnitudes of selection and learning, but does not change the conclusion. I also compare patent granting rates in and out of agglomerations, to see whether the higher rent in agglomerations prompts establishments to file for immature patents. As Farre-Mensa, Hegde, and Ljungqvist (2016) found, patents are important signals to potential investors and loaners; firms with a patent are more likely to get financed. Thus, the urge to finance may have prompted establishments in agglomerations to file for immature patents. I find the overall granting rate high, with more than 85% of patent filings eventually granted. This rate is 3.4% lower for establishments in agglomerations than those outside of, but the difference is not statistically significant at the 10% level. Therefore, the competition pressure in agglomerations may have hasted establishments to file for patents that are not ready, but the effect is insignificant. Third, I apply no weights to the count of patent filings. The selection effect for less innovative establishments rises by about 0.8 percent while learning falls by approximately 3.8 to 5.2 percent. This indicates that: 1) innovators are more likely to survive agglomerations than non-innovators even if they filed for patents and failed or their patents were never cited, which again can be explained by the signaling effect of patents; and 2) agglomerations lift the quality of patents at a similar rate as they speed up the applications. Fourth, I weight by number of citations relative to the average citation in the same technology class and of the same age. Compared to the main weighting scheme, this weight considers the difference in citations across technology class and the nonlinearity of the accumulation of citations over the life cycle of patents. However, subject to the relatively small number of patents in each technology-class-by-age group, this weight is also more sensitive to outliers. Nevertheless, I find the main results barely change either qualitatively or quantitatively with this weight. Fifth, I weight by total citations, and the results are still qualitatively robust, with learning boosting by about three to ten percent at the top three quantiles (98.5 to 99.5). Last, I exclude self-citations. This does not affect the main results except at the 99.5 percentile, where the coefficient decreases by 12.2%. It shows that only top innovators self-cite frequently.

The change over time. I also test the change of learning over time by dividing the ten-year study period into halves. The result, while qualitatively robust in both periods, is quantitatively enlarged in the second. Learning as well as selection grow over time in Maryland agglomerations. The potential erosion of learning in the modern age with the development of transportation and internet technologies, identified by some prior studies (Packalen and Bhattacharya, 2015), is not evident in this case. This can be explained by the nature of Maryland agglomerations. Maryland agglomerations are mostly IT, high-tech or education-related, and rely heavily on peer-to-peer interactions; this type of knowledge spillovers, as suggested by Glaeser (1998), are not losing their magic in the modern economy. On the contrary, they are on the rise.

The impediment effect of county boundaries. I further test the potential impediment effect of county boundaries on knowledge spillovers. I run the same regressions on agglomerations spanning across multiple counties, and the result remains fairly robust with minimal change of the coefficients. Thus, county boundaries don't appear to inhibit knowledge spillovers and this strengthens my case of treating space continuously instead of cutting them by administrative boundaries.

Permutation test. Last, I conduct a permutation test. I randomly assign locations to the agglomeration status at the actual probability for 10,000 times, run the same quantile regressions and summarize the results. The chance of a statistically significant coefficient at the 10% level at 97 to 99.5 percentiles is only 0.9% to 9.5%; the estimated coefficients nicely approximate a normal distribution centering zero.

2.6 Conclusion

This chapter brings to our attention that learning and selection co-exist in agglomerations. As a result, the effect of agglomerations on innovation estimated by ordinary least square regression is upwardly biased. The bias, using agglomerations in Maryland, 2004-2013, as an example, could be four times the size of the true learning effect.

A more precise estimate of learning is provided using quantile regression. An average establishment that files for at least one patent during the study period increases citation-weighted patent applications by 7.8% to 11.4% by locating in agglomerations, while at the same time, non-innovators are 2.5% less likely to survive in agglomerations. These effects are substantial.

This chapter informs policy makers and entrepreneurs making locational choices. The more precisely estimated learning effect can be applied to cost-benefit and cost-effectiveness analyses of industrial policies, while the different effects across the geographical scope can be used to determine the optimal size of agglomerations. For establishment managers making locational choices, the innovativeness of their establishments needs to be carefully evaluated. Highly innovative establishments benefit the most by locating in denser employment centers, while non-innovators can greatly lift their chance of survival by avoiding competitive business centers.

Chapter 3: The Heterogeneous Effect of Clusters across Industries

3.1 Introduction

In recent years, policymakers have invested millions of public funds to support industrial clusters in different regions. They are also targeting at specific industries and geographies. For instance, Boston launched a $1 billion initiative in 2006 to grow a life science cluster[17]. In 2010, the U.S. Department of Energy awarded $129 million to the Energy Regional Innovation Cluster at the 1,200-acre Navy Yard in Philadelphia[18]. These investments, however, may not be optimal: These targeted industries may not be the ones that benefit the most from clustering, and the targeted geographical regions may not be at the scale that the strongest effect of clusters prevails.

While optimizing the targeted industries and spatial scales can improve the efficiency of industrial policies, studies to date fail to provide such guidance. In terms of industry optimization, most studies aggregate industries at high levels of classification based on similarity of products (Feldman and Audretsch, 1999; Wallsten, 2001; Mariani, 2004), while others disaggregate, but only study a small subset of industries (Beaudry, 2001; Bloom, Schankerman, and Van Reenen, 2013). As a result, the spectrum of industrial clusters measured by previous studies does not

[17] https://www.clustermapping.us/sites/default/files/files/resource/Clusters_and_Competitiveness-_A_New_Federal_Role_for_Stimulating_Regional_Economies__Full_Report_.pdf

[18] http://scienceprogress.org/2010/08/a-win-for-regional-innovation/

capture the relationship among industries in production process (in the former case) and/or is incompetently measured (in the latter case).

In terms of spatial optimization, most studies define clusters by administrative boundaries due to the aggregated feature of business data (Combes, Duranton, Gobillon et al., 2012; Feser, Renski, and Goldstein, 2008; Ellison, Glaeser, and Kerr, 2010). These administrative boundaries are usually no smaller than a county. Thus, if the effect of clusters prevails at a smaller scale, prior studies could not have identified it. Moreover, administrative boundaries usually are not economically meaningful and do not make reasonable boundaries for clusters. A few recent studies adopt business data at the establishment level and find that the effect of clusters largely concentrates within a one-mile radius region (Wallsten, 2001; De Silva and McComb, 2012). These pieces of evidence reveal how aggregated data mask the true scale of clusters. But to date, studies using establishment-level data are still rare and none of them analyzes the spatial scope of clusters as industry-specific.

This chapter closes these two gaps. I estimate the effect of clusters on innovation across 34 industrial clusters defined by Delgado, Porter, and Stern (2014) according to input-output, co-location, and labor-sharing relationships between industries. A continuous quantile method following Combes, Duranton, Gobillon et al. (2012) and a comparison between the full sample and the new establishment sample are applied to eliminate the selection bias. Metalworking Technology cluster, consisting of 14 closely interrelated six-digit NAICS industries, encourages innovative by the greatest magnitude, followed by Food Processing and Manufacturing, and Automotive clusters. These three industrial clusters increase a establishment's patent filings by 17.6, 8.5, and 5.5 percent, respectively. I also optimize the spatial scale of each industrial cluster to maximize its effect on innovation. Metalworking Technology, Food Processing and Manufacturing, and Automotive clusters, for example, are associated with optimal sizes of three, one and two mile(s) in radius, respectively. Eventually, this chapter delivers a complete ranking of industrial clusters by the magnitude of their effects on innovation, and shows their optimal spatial scales. By consulting this list, policymakers can design efficient industrial policies. They can prioritize the industrial clusters on top of the list with cluster-oriented policies and generate the greatest spurt in innovation.

The contribution of this chapter is four-fold. It contributes to both studies on the heterogeneous effect of clusters across industries and studies on the spatial scope of clusters. It examines clusters in detailed industry classification based on meaningful industrial interconnectedness (34 groups of industries at the six-digit NAICS code level closely related by input-output, co-location and labor-sharing relationships) and, for the first time, allows the spatial scale of clusters to be industry-specific. In addition to identifying these heterogeneities, this chapter also explains the differences in the effect on innovation and spatial scales by a set of key industrial characteristics. The importance of tacit knowledge, education level of employees, and market size turn out to account for a great deal of both differences.

It also contributes to the literature that quantifies the cluster-innovation relationship, by eliminating two forms of selection bias: the selection bias caused by market competition (less innovative establishments are forced out of clusters), and the bias caused by self-selection (more innovative establishments are systematically attracted to clusters). The estimates in this chapter are 0.03% to 1.9% smaller than the (prior) OLS estimates, indicating upward bias in the latter.

This chapter also speaks to the literature on the sources of the benefits of clusters (among others, see Jaffe, Trajtenberg, and Henderson, 1993; Rosenthal and Strange, 2008b). Prior studies suggest that the prevalence of tacit knowledge, the intensity of co-patenting collaborations, the expansion of market size and the high skill level of employees (Elvery, 2010) may be the reasons that clusters encourage innovation. These effects are quantified in this chapter.

It also contributes to studies on industry targeting (among others, see Leatherman, Howard, and Kastens, 2002; Conroy, Deller, and Tsvetkova, 2016). This chapter provides a new strategy to help policymakers target industrial clusters with the greatest effect on innovation at their optimal spatial scales. This strategy can produce cost-effective industrial policies.

3.2 Literature

Researchers find that the effect of clusters on innovation vary across industries. For example, Wallsten (2001) finds that proximity to other innovators increases the chance of innovating in computers, bio-tech and life sciences, and electronics industries, but not in materials and energy sectors. But these studies suffer from two problems. First, in terms of industrial aggregation, they either aggregate industries at high levels based on product similarity (Mariani, 2004) or only study a small subset of detailed industries (Beaudry, 2001; Bloom, Schankerman, and Van Reenen, 2013). High-level industry aggregation based on product similarities masks the inherent unrelatedness in the production process, and therefore provides misleading policy applications by clustering unrelated industries together. At the same time, focusing on a small subset of industries also provides only limited guidance for industry targeting due to the narrow research scope. Second, the reasons behind the heterogeneous effect are not well understood. Researchers have put forth hypotheses that the reliance on tacit knowledge and market size may affect how much clusters uplift innovation (Carlino and Kerr, 2014; Delgado, Porter, and Stern, 2014), but empirical studies are rare. Identifying key industrial characteristics accounting for the heterogeneous effect will deepen our understanding of clusters. For example, if we find industries relying on tacit knowledge benefit more from clusters, then it implies that the spillover of uncodified knowledge plays a central role in clusters. Moreover, pinning down a set of key industrial characteristics also help to apply this study to general settings. For instance, in this chapter, I rank 34 industrial clusters defined based on NAICS by their effects on innovation. This ranking directly applies to policymaking in the United States but does not in other countries that classify industries in a different way. However, by knowing that clusters encourage innovation the most in industries relying on tacit knowledge for example, policymakers around the world can then identify and target such industries in their own jurisdictions.

The effect of clusters attenuates by distance. Most studies to date aggregate geographies by administrative units and thus cannot identify the attenuation effect precisely (Combes, Duranton, Gobillon et al., 2012; Feser, Renski, and Goldstein, 2008; Ellison, Glaeser, and Kerr, 2010). A few recent studies, using more detailed geographical data, find that the effect of clusters is quite local, largely concentrated within a one-mile radius region (Wallsten, 2001; Arzaghi and Henderson, 2008; De Silva and McComb, 2012). Moreover, the spatial scope at which the effect of clusters prevails should differ by industries, as industries have different levels of reliance on collaboration and knowledge spillovers. But to date, no studies allow the spatial scale of clusters to be industry-specific, except Rosenthal and Strange (2003), who find that own-industry employment encourages new employment only within a one-mile-radius ring for the apparel industry, while the radius expands to ten miles for the software and fabricated metal industries. While Rosenthal and Strange (2003) focus on the effect of clusters on employment and six two-digit SIC industries (corresponding to approximately three-digit NAICS industries), this chapter extends the analysis to the effect on innovation and more detailed industry classification (34 industrial clusters based on six-digit NAICS code).

3.3 Data and Method

3.3.1 Data

I adopt the same datasets as described in section 2.3.1. Aside from these two main datasets, I also use the American Community Survey and previous literature as complimentary data sources in this chapter.

3.3.2 Method

In this chapter, I extend my previous definition for agglomerations into clusters of different industries. I group industries by their input-output, co-location and labor-sharing relationships, following Delgado, Porter, and Stern (2014). They have classified a total of 67 groups of closely related industries based on 978 six-digit NAICS coded industries. In this chapter, I focus on 34 groups of industries in which at least ten establishments applied for patents during the study period 2004 to 2013; this is to ensure enough observations of patent applications in the estimation. In terms of geography, this chapter again defines a cluster as a region around every establishment with a flexible radius. The radius changes from one mile to twenty-five miles to search for the optimal scale at which clusters encourage patent filings by the greatest magnitude. This optimal radius is industry-specific. This definition triumphs those in most prior studies, which define clusters by administrative boundaries and maintain the spatial scale of clusters constant across industries. In this chapter, the spatial scale of clusters is empirically optimized rather than pre-determined. Putting together, if a region (with a changeable radius) around an establishment has an above-median employment density in related industries, it is defined as a cluster. I also shift the threshold of employment density to 1.2 times greater than average and change the definition to focus on establishment density, but the primary analysis looks at employment density above median.

I take care of two important forms of selection bias in estimating the effect of clusters on innovation: 1) less innovative establishments are forced out of clusters (market selection), and 2) more innovative establishments actively choose to locate in clusters (self-selection). I tackle with market selection with a continuous quantile estimator developed by Gobillon and Roux (2008) and Combes, Duranton, Gobillon et al. (2012). Compared to the discrete quantile regression adopted in chapter 2, this method is similar in logit but it only estimates three parameters. Since I'm comparing across a wide range of industries, reducing the number of parameters is essential. A functional form as follows is adopted to describe the relationship between distributions of establishment innovation in and out of clusters, which allows for the presence of both the true learning effect and market selection effect, and the cluster effect is allowed to increase with the innovativeness of establishments (i.e., more innovative establishments benefit more in clusters).

$$\varphi_{n,i}(u) = D_n \varphi_{n,j}(S_n + (1 - S_n)u) + A_n \text{ for } u \in \left[\max\left(0, \frac{-S_n}{1-S_n}\right), 1\right]$$

$\varphi_{n,i}(u)$ is the uth quantile of the distribution F_i of establishment innovation in clusters of industry n. $\varphi_{n,j}(u)$ is its counterpart in non-clusters. The former can be approximated by three changes in the latter: a right-shift (parameter A), a left-truncation (parameter S) and a dilation (parameter D). This is because clusters force out the least innovative establishments and thus lead to a left-truncation, improve the innovativeness of the remaining establishments and thus lead to a right-shift, and improve innovativeness the most for the best performers and thus lead to a dilation (Baldwin and Okubo, 2006; Behrens, Duranton, and Rober-Nicoud, 2014). A, if >0, is the average improvement of innovation for an in-cluster establishment compared to an out-of-cluster counterpart. D, if >1, stands for the magnitude by which more innovative establishments benefit more from clusters. S, if >0, stands for the magnitude of the market selection effect, or more specifically, the percentage of the least innovative establishments being forced out of clusters compared to that in non-clusters. A continuous quantile estimator developed by Gobillon and Roux (2010) and Combes, Duranton, Gobillon, et al. (2012) is applied to estimate A, D and S. The average learning effect (A) and the dilation (D) estimated from the above procedure are still contaminated by self-selection. I again contrast the full sample with the new establishment sample to deal with this issue.

I repeat the above procedure for each industrial cluster at different spatial scales to search for the optimal scale that maximizes the effect on innovation A. This is done by the following approach, under the assumption that learning is a smooth function of the spatial size of clusters. I let the computer randomly pick an integer x between one and twenty-five as the starting radius, and update it by one in both directions each time. The radius that returns the larger learning effect is used as the next starting point, until both x+n+1 and x+n-1 return a magnitude smaller than that of x+n. x+n then is a local optimal point. To obtain the global optimal point, I repeat the above steps ten times and compare the local optimal point(s). The one returning the largest effect is the global optimality. This is carried out for 34 industrial clusters, and their (optimized) spatial scales and effects on innovation are then recorded and compared. In the end, this delivers a ranking of industrial clusters by the strength of learning at their optimal spatial scales.

Last, I group industrial clusters by a series of industry-level characteristics and examine how these characteristics shape their spatial scales, effects on innovation and magnitudes of market selection. This is implemented with the main estimation procedure described in the above paragraphs combined with t-tests. The industry characteristics include: the importance of tacit knowledge, average number of patent collaborators, market size, patent intensity, and education level of employees. A dummy variable is constructed indicating whether tacit knowledge is mentioned in the literature as critical for an industry. Learning should be stronger if tacit knowledge is important, as clusters help establishments gain access to tacit, rather than codified knowledge. At the same time, the spatial scale of clusters should be smaller, as tacit knowledge, by definition, can only be transmitted through face-to-face interactions and is thus sensitive to distance. The literatures are compiled from Web of Science, Google Scholar, SSRN and NBER, with key words combinations of "tacit knowledge" or "uncodified knowledge" and the name of each industrial cluster or its sub-industries. Number of patent collaborators is constructed from the USPTO patent data and the QCEW data, and used to capture the importance of network in an industry. If network is important, clusters should show greater effect because they help establishments to build denser networks with neighbors. Market size is obtained from Berry and Garrison (1958) and measured by the minimum population an industry serves. A smaller market size is associated with a tougher local competition and therefore a stronger market selection effect and a smaller spatial scale; industries depending on a large market size may also be associated with a stronger learning effect, as clusters help establishments to form collective competitiveness and expand the market. A dummy variable indicating whether an industry is patent intensive, with patent intensity defined by the number of patents per employee, is obtained from Economics and Statistics Administration and USPTO (2012). High patent intensity implies that innovation is important for survival and therefore the market selection effect should be stronger. Education level of employees is measured by the percentage of workers that hold a bachelor's degree. This variable is obtained from the American Community Survey. A more educated workforce may contribute to both learning and market selection. On the one hand, highly educated employees may learn faster in clusters and improve innovation by a greater magnitude. On the other, less innovative establishments may have difficulty affording highly educated workers and thus more likely to be pushed out.

3.4 Results

3.4.1 Ranking of industrial clusters by their effects on innovation

The 34 industrial clusters are summarized in Table 3.1. Each consists a group of six-digit NAICS industries that are closely related in input-output, co-location and labor-pooling relationships. For example, the Metalworking Technology cluster, contains 14 six-digit industries including Machine Tool Manufacturing, Metal Heat Treating, and Industrial Mold Manufacturing. Figure 3.1 shows the distribution of the 34 clusters by the number of establishments applying for patents during the study period. In most clusters, fewer than 60 establishments applied, while three clusters, including Business Services, Education & Knowledge Creation and Distribution & Electronic Commerce, have over 100 establishments applied. These are both highly innovative and major clusters in Maryland. They account for 81% of establishments and 74% of jobs in the state. At the same time, an establishment in these three industrial clusters is four times more likely to apply for patents and produces ten times more citation-weighted patent filings than one in other industries; both differences are statistically significant at the 5% level under a t-test.

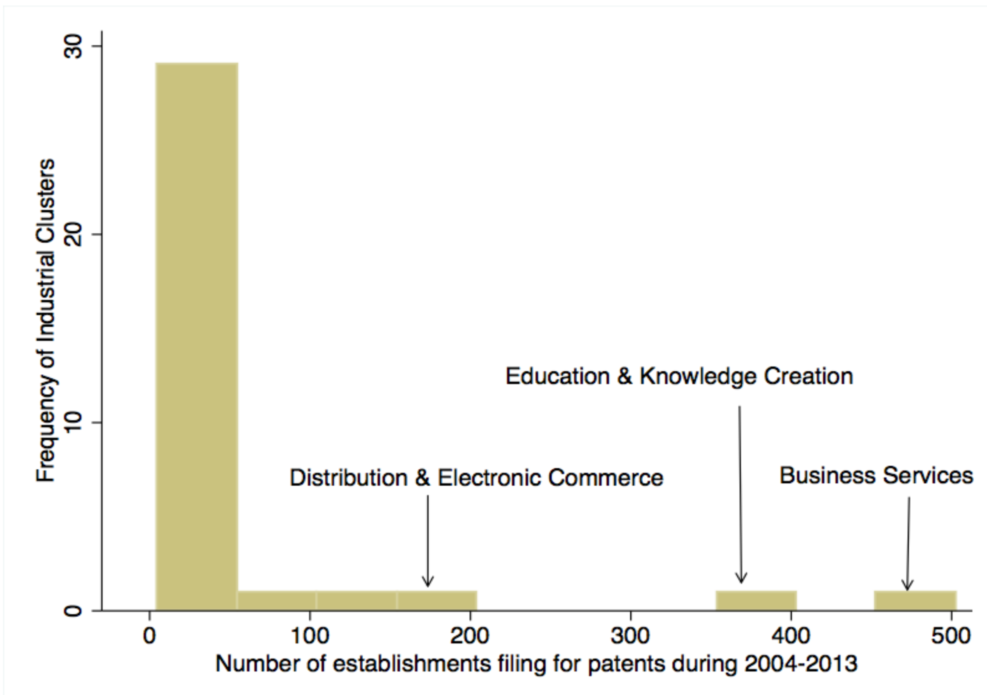
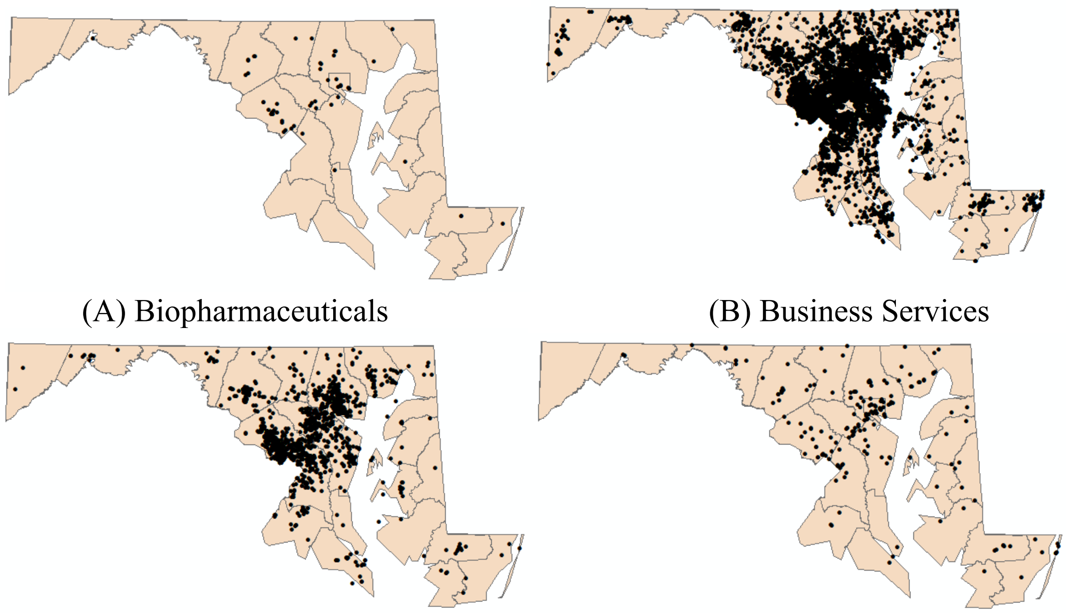


Figure 3.1 Number of establishments applying for patents during the study period (2004-2013) in the 34 industrial clusters

Geographically, the 34 industrial clusters do not concentrate at the same place. For example, as Figure 3.2 shows, establishments in Biopharmaceuticals are relatively few and spatially scattered. In contrast, there are a lot of establishments in Business Services; they are highly concentrated in the central Maryland—in Montgomery, Howard, Baltimore County and City, Anne Arundel and Prince George's. Establishments in Education and Knowledge Creation cluster are even more concentrated in Montgomery, Howard and Baltimore City. In Food Processing and

Manufacturing, on the contrary, establishments are mainly geographically dispersed with a slight concentration in Baltimore City. Employment distributions are similar in pattern. To conclude, different industrial clusters exhibit different geographical distributions; as a result, their effects can be spatially disentangled.



(A) Biopharmaceuticals (B) Business Services

(C) Education & Knowledge Creation (D) Food Processing & Manufacturing

Figure 3.2 Spatial distributions of establishments in different industrial clusters

Table 3.1 ranks the 34 industrial clusters by how much they encourage innovation (parameter A), and reports their optimal geographical scale. The cluster that boosts innovation by the greatest magnitude is Metalworking Technology. When formed at a scale of three miles in radius, this cluster uplifts the number of citation-weighted patent applications by e0.162-1, or 17.6 percent. This is a sizable effect. The literature estimates the elasticity of establishment patenting to R&D expenditure close to 0.5 (Aghion, Van Reenen, and Zingales, 2013; Bloom, Schankerman, and Van Reenen, 2013). A 17.6% increase in patenting roughly equals a 35% increase in R&D expenditure. Agrawal, Galasso, and Oettl (2017) estimate that a 10% increase in a region's stock of highways increases the region's patenting by 1.7%. A 17.6% increase in patenting then is equivalent to doubling the stock of highways. For regions with a relatively strong Metalworking Technology industry, a cluster-oriented policy targeting at this industry would be cost-effective. Metalworking Technology covers a wide range of metalworking jobs that create parts for all kinds of machines, including engines, hand tools and accessories. The most competitive establishments in this industry is highly technology-based and innovative. For example, Black & Decker Corporation, an American metalworking manufacturer headquartered in Towson, Maryland, applied for 200 patents during the study period with a total citation of 85019. Studies also confirm that this industry relies heavily on tacit knowledge (Balconi, 2002), which may explain why clusters in this industry lead to a large improvement in patent filings. Food Processing and Manufacturing cluster and Automotive cluster rank the second and the third. When formed at a scale of one mile and two miles in radius, these two clusters increase patent filings by 8.5 and 5.5 percent, respectively. These effects although much smaller than that of the Metalworking Technology cluster, are still quite sizable compared to other policies or projects (Aghion, Van Reenen, and Zingales, 2013; Bloom, Schankerman, and Van Reenen, 2013; Agrawal, Galasso, and Oettl, 2017). Other industries on top of the list are Recreational and Small Electric Goods, Education and Knowledge Creation, and Upstream Metal Manufacturing. After the top ten industrial clusters on the list, the effect of a cluster on innovation becomes small (no more than 1%). Thus, cluster

[19] The main result is not completely driven by this particular establishment though; excluding it, the Metalworking Technology cluster still increases citation-weighted patent applications by 16.4% and ranks the first among all industrial clusters.

policies do not work for every industry; this signals the importance of appropriate industry targeting. Down at the bottom of the list are Communications Equipment and Services, Biopharmaceuticals, and Local Utilities, in clusters of which, the effect on innovation is essentially zero in magnitude. In Lighting and Electrical Equipment and Upstream Chemical Products, effects are negative by a magnitude of 1% and 5.7%, respectively. This implies that in these two industries, knowledge may become more rigid in dense employment centers than outside and actually inhibit innovation.

Table 3.1 Ranking of industrial clusters by effect on innovation (parameter A)

Rank	Cluster	Optimal Size (mile in radius)	Average effect on innovation, or learning (A)	Magnitude that more innovative establishment benefit more (D)	Market selection (S)	Pseudo R^2	Sampl Size
1	Metalworking Technology	3	0.162*** (0.003)	2.956*** (0.002)	0.050 (0.123)	0.516	1,983
2	Food Processing and Manufacturing	1	0.082*** (0.0001)	2.337*** (0.001)	0.024 (0.023)	0.485	3,537
3	Automotive	2	0.054*** (0.0003)	1.246*** (0.002)	0.046*** (0.002)	0.530	507
4	Recreational and Small Electric Goods	1	0.048*** (0.005)	1.250 (0.200)	0.043*** (0.012)	0.522	147,10
5	Education and Knowledge Creation	1	0.029*** (0.004)	1.004 (0.220)	0.034 (0.123)	0.022	20,62
6	Upstream Metal Manufacturing	1	0.028*** (0.0007)	0.856*** (0.003)	0.013 (0.231)	0.511	742
7	Information Technology and Analytical Instruments	1	0.025*** (0.005)	1.494* (0.212)	0.031*** (0.003)	0.252	4,180
8	Downstream Chemical Products	2	0.020*** (0.0007)	2.007*** (0.006)	0.015*** (0.004)	0.509	1,972
8	Distribution and Electronic Commerce	1	0.020*** (0.0007)	2.007*** (0.006)	0.004 (0.065)	0.509	67,48
10	Production Technology and Heavy Machinery	3	0.017* (0.009)	1.643*** (0.011)	0.025 (0.021)	0.441	4,186
11	Construction Products and Services	1	0.008*** (0.001)	0.559*** (0.089)	0.003 (0.211)	0.503	5,160
11	Downstream Metal Products	1	0.008 (0.011)	0.559*** (0.083)	0.008 (0.023)	0.503	1,778
13	Wood Products	1	0.007 (0.014)	0.559* (0.208)	0.007 (0.034)	0.503	139,17

14	Plastics	1	0.0065 (0)	1.646 (0)	0.013 (0)	0.470	1,193
15	Marketing, Design, and Publishing	1	0.00025*** (1.734e-26)	1.382 (0.328)	0.003 (0.382)	1.000	31,438
16	Aerospace Vehicles and Defense	2	0.00019*** (3.076e-26)	1.038 (0.160)	0.005 (0.015)	1.000	1,243
17	Local Personal Services (Non-Medical)	1	0.00018*** (1.329e-26)	1.739** (0.285)	0.000 (0.286)	1.000	61,590
18	Local Health Services	1	0.00015*** (2.025e-26)	1.260 (0.291)	0.000 (0.132)	1.000	126,52
19	Local Household Goods and Services	1	0.00004 (0.052)	1.022 (0.298)	0.000 (0.169)	0.679	23,846
19	Hospitality and Tourism	1	0.00004 (0.088)	0.987 (0.165)	-0.001 (0.434)	0.687	17,423
21	Financial Services	1	0.000033*** (1.908e-26)	1.050 (0.195)	0.002 (0.231)	1.000	28,135
22	Jewelry and Precious Metals	5	0.00002 (0.123)	1.043 (0.254)	-0.002 (0.257)	0.507	244
22	Insurance Services	1	0.00002 (0.044)	1.006 (0.365)	-0.001 (0.322)	0.758	9,008
24	Local Financial Services	1	7.515e-06 (0.002)	1.078 (109.25)	0.001 (0.259)	1.000	11,584
25	Business Services	1	1.744e-06*** (8.311e-26)	1.0008 (0.784)	0.006 (0.052)	1.000	113,88
26	Performing Arts	1	6.124e-08 (0.013)	1.001 (0.232)	0.002 (0.166)	0.575	3,609
27	Local Logistical Services	1	7.838e-09 (0.071)	1.003 (0.132)	0.000 (0.087)	0.640	20,101
28	Local Retailing of Clothing and General Merchandise	1	5.443e-09 (0.089)	0.999 (0.057)	-0.001 (0.063)	0.535	267,35
29	Medical Devices	1	0 (0)	1 (0)	0.003 (0)	/	126,39
30	Communications Equipment and Services	2	-2.877e-06*** (1.385e-26)	0.999*** (0.0008)	0.003 (0.072)	1.000	4,999
31	Biopharmaceuticals	1	-0.0005*** (7.455e-26)	0.813*** (0.018)	0.006 (0.132)	1.000	637
31	Local Utilities	1	-0.0005 (0.0007)	0.539 (46.360)	0.001 (0.131)	1.000	5,803
33	Lighting and Electrical Equipment	4	-0.011 (0.012)	0.788*** (0.007)	-0.007 (0.232)	0.313	1,755
34	Upstream Chemical Products	12	-0.059*** (0.002)	0.508*** (0.005)	-0.038*** (0.007)	0.418	1,094

Note: Bootstrapped standard errors are in parenthesis. *p<0.05, **p<0.01, ***p<0.005 We test for A=0, D=1 and S=0.

It is also evident from the list that the effects of most industrial clusters are locally concentrated. 24 out of the 34 industrial clusters, and six out of the top ten, exhibit their strongest effects within a one-mile radius region. Only one cluster, Upstream Chemical Products, has it effect prevail beyond five miles. This is consistent with previous studies that identify the sharp attenuation of the cluster effect over distance (Arzaghi and Henderson, 2008; Rosenthal and Strange, 2003). This also indicates that direct human interactions (face-to-face communications) and worker turnover likely play an essential role in fostering innovation in clusters; since the probability of a person running into another attenuates fast with the physical distance between them, the effect of clusters on innovation does not survive long distance.

The market selection effect (parameter S) is numerically unneglectable but statistically insignificant at the 5% level for most industries. But for a subset of industrial clusters, including Automotive, Recreational and Small Electric Good, Information and Technology and Analytical Instrument, and Upstream Chemical Products, market selection is statistically significant. In these four industries, clusters force out an additional 4.6, 4.3, 3.1 and 1.5 percent of the least innovative establishments compared to non-clusters, respectively. Downstream Chemical Products, on the contrary, has a negative market selection effect, with non-clusters forcing out 3.9 percent more less innovative establishments than clusters. This indicates that in Downstream Chemical Products, clusters are actually more tolerant of non-innovative establishments, likely through buffering them from bankruptcy with cost- and labor-sharing.

One thing worthy of special notice is that the common wisdom for industry targeting turns out to be misleading. Practitioners tend to target highly innovative industries for cluster development. Table 3.1 shows that while this strategy is right about some industries, such as the highly innovative Education and Knowledge Creation and Information Technology and Analytical Instruments industries, which do top this list, it is wrong about others. For example, the similarly highly innovative Biopharmaceuticals and Business Services industries, as well as the Medical Devices industry, do not appear to encourage patent filings at all, and certainly do not encourage more than the not-so-innovative Food Processing and Manufacturing and Woods Products industries. An industry being innovative in general does not guarantee that forming a cluster in this industry encourages more innovation; it could be that locations with either high or low employment density are equally highly innovative. These are just two separate issues and one does not infer the other.

I then examine whether the estimated average effect of clusters on innovation (A) and dilation (D) are still contaminated by self-selection bias with the new establishment sample. For most industrial clusters, self-selection is both small (less than 2% of the original A and S estimates) and statistically insignificant at the 5% level. There are a few exceptions. For Construction Products and Services (No. 11 in the ranking), Marketing, Design, and Publishing (No.15), Local Health Services (No.18) and Finance Services (No.21), self-selection is of comparable magnitude of the original estimate of parameter A and thus brings down the true effect of clusters to essentially zero. But all these self-selection coefficients are insignificant at the 5% level, except for the Finance Services cluster.

Finally, I formally evaluate the bias of an OLS estimator. I re-estimate the effect of the 34 clusters on innovation with OLS regressions. In 32 clusters, the OLS estimates are larger than the estimates for parameter A in Table 3.1. This indicates an upward bias of the OLS estimator due to selection bias (captured by parameter S). The upward bias varies from 0.3 to 1.9 percent across the 32 industrial clusters, which is not large but still unneglectable. This is consistent with the fact that in most of the industrial clusters, the parameter S is both small (<1%) and statistically insignificant, but in some industries, the selection bias is serious with S larger than 3% and statistically significantly different from zero at the 5% level.

3.4.2 Industrial characteristics that explain the difference

In this section, I explain the different optimal sizes of clusters and their heterogeneous effects on innovation using industrial characteristics. These industry-level characteristics include: the importance of tacit knowledge, average number of patent collaborators, market size, patent intensity, and education level of employees. The importance of tacit knowledge, as mentioned above, is measured by a dummy variable indicating whether tacit knowledge is mentioned in the literature as important for a certain industry. I expect the importance of tacit knowledge to be associated with a larger learning effect and a smaller spatial scale. The literatures are compiled from Web of Science, Google Scholar, SSRN and NBER, with key words combinations of "tacit knowledge" or "uncodified knowledge" and the name of each industrial cluster or its sub-industries. For example, for Construction Products and Services industry, I identify the paper "Tacit knowledge and organisational performance: construction industry perspective" by Pathirage, Amaratunga, and Haigh published in Journal of knowledge management in 2007. This paper states that tacit knowledge is critical in construction industry. As a result, the dummy variable is coded as one for this industry. For each group, I re-run the main estimation to obtain parameter A (learning) and S (market selection), and apply t-tests to compare the difference across groups in both parameters as well as the identified optimal sizes of clusters in Table 3.1. Table 3.2 shows the result. As expected, the spatial scale of clusters is 0.8 miles smaller in radius for industries with important tacit knowledge. At the same time, learning is also seven percent larger in these industries. Market selection is also much stronger; this is sensible as tacit knowledge makes location more critical, and the competition over location is thereby likely to be intensified. These differences are all statistically significant at the 5% level. Note that the results can be unreliable is only one study suggests that tacit knowledge is important in an industry. In my analysis, only two industries, Distribution and Electronic Commerce and Local Logistic Services, have no more than one study stating that they rely on tacit knowledge. Removing these two industries from the analysis does not change the result much. Cluster optimal size becomes slightly larger at 1.438 miles in radius, learning enlarges to 0.017 and selection enlarges to 0.012. Statistically significance does not change at all.

Table 3.2 Cluster size, learning and market selection by industrial characteristics

Industry Characteristics	Group	Cluster size (mile in radius)	Learning (A)	Market selection (S)	Sample size
Tacit Knowledge	Important	1.389*** (0.007)	0.016*** (0.00009)	0.011*** (0.0004)	580,550
	Unimportant	2.188*** (0.002)	0.009*** (0.00007)	0.005*** (0.00005)	676,045
Number of patent collaborators	Above median	1.75*** (0.005)	0.013*** (0.00008)	0.006*** (0.00004)	750,245
	Below median	1.8*** (0.004)	0.012*** (0.00007)	0.013*** (0.00006)	506,350
Market Size	Above median	2*** (0.005)	0.017*** (0.00009)	0.007*** (0.00005)	683,304
	Below median	1.333*** (0.003)	0.007*** (0.00005)	0.009*** (0.00004)	573,291
Patent Intensive	Yes	2.667*** (0.007)	0.014*** (0.0001)	0.0083*** (0.00007)	454,949
	No	1.053*** (0.0005)	0.012*** (0.00005)	0.0078*** (0.00003)	801,646
Share of workers holding a bachelor's degree	Above median	1.308*** (0.003)	0.014*** (0.00007)	0.010*** (0.00004)	472,067
	Below median	2.048*** (0.005)	0.009*** (0.0004)	0.007*** (0.00005)	784,528

Note: Bootstrapped standard errors are in parenthesis. *p<0.05, **p<0.01, ***p<0.005

Average number of patent collaborators in an industry, defined as how many establishments on average collaborate on a patent application, captures the importance of network for innovation in an industry. If networking is important, clusters should exhibit a greater learning effect as they help to build networks by concentrating establishments and workers in geographical proximity; at the same time, the spatial scales of clusters are expected to be smaller as networking requires proximity. Table 3.2 shows that as expected, industries with above median number of patent collaborators form clusters at a slightly smaller scale (0.05 miles smaller in radius) and learning is also slightly stronger (by 0.1%). These differences are statistically significant at the 5% level but numerically small.

Market size is measured by the minimum population an industry serves and constructed from Berry and Garrison (1958)[20]. These two authors quantify market sizes for 52 industries, most of which are service-based. I match these 52 industries as closely as possible to my 34 industrial clusters. For example, the Local Health Services cluster is linked to Hospital and Clinics, which has a market size of 1159 persons. In this way, I manage to match 24 out of the 34 clusters to the Berry and Garrison (1958) study, and use these 24 clusters to carry out the analysis. A smaller market size is expected to be associated with a smaller spatial scale for clusters and a tougher local competition (i.e., a stronger market selection effect). Table 3.2 shows that indeed, a smaller market size results in a spatial scale that is 0.7 miles smaller in radius and a one percent weaker learning effect. A smaller market size is also associated with a market selection effect 0.2 percent stronger in magnitude.

Patent intensity, defined as the number of patents per employee in an industry, is measured with a dummy variable indicating whether an industry is patent intensive, constructed from Economics and Statistics Administration and USPTO (2012). Innovation is important for survival in a patent intensive industry, and therefore in such an industry, the market selection effect should be stronger. Table 3.2 shows that patent intensive industries exhibit a marginally stronger selection effect (by 0.05%) and a slightly smaller effect of learning (by 0.2%); this is consistent with the findings from Table 3.1 that an industry being highly innovative and patent intensive does not necessarily infer that clusters in this industry encourage innovation by a large

[20] These estimates obviously are a bit out of date and may not accurately measure market size today. Nevertheless, Berry and Garrison (1958) provide the most complete list of industries with market sizes, and since I'm grouping industries by above/below median market size, the exact measurement is not that important.

magnitude. Table 3.2 also shows that patent intensive industries form clusters 1.7 miles larger in radius. While this difference is both large and statistically significant at the 5% level, there isn't a clear-cut theoretical explanation of why patent intensive industries would form larger clusters. It is likely a coincidence due to the fact that patent intensive industries are also mostly traded industries while patent unintensive industries are mostly local service industries.

Last, education level of employees in an industry is measured by the percentage of workers with a bachelor's degree, obtained from the American Community Survey. This variable may contribute to both the effect on innovation and the market selection effect as educated workers may learn faster and therefore benefit more from clusters, while at the same time, educated workers are costly and may become unaffordable for non-innovative establishments. Table 3.2 confirms both. Industrial clusters with more educated workers exhibit a 0.5 percent larger learning effect and a 0.3 percent stronger market selection effect.

3.5 Robustness Checks

Alternative definitions for clusters. I change the threshold of employment density for cluster definition from above median to 1.2 times above average. I find that in 32 out of the 34 industrial clusters, the change of the parameters is numerically small and does not shift their relative ranking. Two industries, Upstream Metal Manufacturing, and Distribution and Electronic Commerce, increase their effects on innovation under this new definition from 2.8% to 6.1%, and 2% to 4.3%, respectively. This also upshifts their rankings from No. 6 and No. 8 to No. 3 and No. 6.

I also use establishment density instead of employment density above median as an alternative definition for clusters. The assumption here is that knowledge spillovers and rivalries happen across establishments, not within establishments, as workers within an establishment hold more homogenous knowledge. This assumption is reasonable if most establishments are small and workers play similar roles in these establishments. The assumption implicitly adopted in the main definition is that workers are the carriers of knowledge, and thus either within- or across-establishment interactions can yield new knowledge. This assumption holds if establishments are large or workers in establishments, even small ones, play different roles and therefore hold different knowledges. Both assumption may be plausible and there is no reason to believe one more than the other. Given the plausibility of both assumptions, I use employment density as the main definition due to the nature of the data. The QCEW data are a count of employment and covers 98% of jobs in the United States, but the data do not include self-employed establishments and thus undercount the number of establishments. As a result, it is more accurate to use this dataset to measurement employment than establishments. Nevertheless, when defining clusters by above-median establishment density, I find the optimal sizes of all clusters completely unchanged, and the effects on innovation almost unchanged with differences smaller than 0.01% in 19 out of the 34 industrial clusters. The exceptions are as follows. For Communications Equipment and Services cluster, while the change of the parameter A is statistically significant at the 5% level, the magnitude is very small (1.00e-03) and numerically unimportant. Four industries experience an increase in their effects on innovation: Aerospace Vehicles and Defense cluster increases its effect from 0.2

to 0.7 percent, and its rank climbs from No.16 to No.13. Biopharmaceuticals cluster, the effect of which increases from -0.5 to 1.4 percent, now ranks No.11 instead of No.31. Lighting and Electrical Equipment cluster increases its effect on innovation from -1.1 to 3 percent, with its rank rocketing from No.33 to No.5. Finally, Upstream Chemical Products cluster increases its effect from -5.9 to one percent, with its rank soaring from No.34 to No.12. Seven other industrial clusters significantly drop their effects on innovation. For example, the effects of Downstream Chemical Products, Downstream Metal Products, Plastics, and Production Technology and Heavy Machinery clusters all become negative, ranging from -0.2 to -0.9 percent. The top industrial cluster, Metalworking Technology cluster, also scales back its effect to 7.9 percent. But it still tops the ranking as the second top industry, Food Processing and Manufacturing, also scales back its effect from 8.5 to 3.1 percent and drops from rank No.2 to No.5. Finally, Automotive cluster also shows a smaller effect on innovation in this new ranking, 2.1 instead of 5.5 percent, which drags its rank down from No.3 to No.8.

Alternative weights of patent applications. In the main analysis, I use per year citation to weight patent applications. I apply five other weights to check for robustness. First, I assign zero weights to ungranted applications, i.e., I count patent grants instead of patent applications. This weight reduces the effect of clusters on innovation for all industries, but does not change the ranking. Second, I apply no weights to patent filings. This again scales back the effect on innovation by almost a half for the top ten industrial clusters. Third, I weight patents by the number of citations relative to their age and average citation of patents in the same technology class. I find the main results barely change. Fourth, I weight by total citations, and the results are still qualitatively robust with the ranking almost unchanged. But the effect on innovation boosts by one to three percent for the top six industries. Last, I exclude self-citations. This barely affects the result, although it decreases the effect on innovation in some industries by a percentage point smaller than 0.1.

3.6 Cost-Effectiveness Analysis

While Table 3.1 compares the benefits of cluster policies on innovation, this section further considers the costs. A Metalworking Technology cluster, as shown in Table 3.1, increases citation-weighted patent applications by 17.6 percent and a Food Processing and Manufacturing cluster by 8.5 percent. While the benefits are different, the costs of developing a cluster in these two industries are also distinct. The data show that to develop an average non-cluster in Metalworking Technology into an average cluster requires the addition of 94 jobs, while in Food Processing and Manufacturing, 819. Thus, after adjusting the benefits by the number of jobs that need to be added, the difference between these two industries enlarges. Table 3.3 carries out a complete re-ranking of the 34 industrial clusters based on their effects on innovation adjusted by number of jobs added, and this ranking is very different from that in Table 3.1. We still have Metalworking Technology topping this list with 1.9‰ increase of citation-weighted patent applications per job added, but the second and the third change to Downstream Metal Products and Distribution and Electronic Commerce.

Table 3.3 Ranking of industrial clusters by effect on innovation, adjusted by number of jobs needed to be added to develop a cluster

Rank	Cluster	Effect on innovation per job added (‰)
1	Metalworking Technology	1.872
2	Downstream Metal Products	1.004
3	Distribution and Electronic Commerce	0.789
4	Information Technology and Analytical Instruments	0.524
5	Automotive	0.496
6	Construction Products and Services	0.275
7	Education and Knowledge Creation	0.259
8	Wood Products	0.190
9	Plastics	0.165
10	Food Processing and Manufacturing	0.104
11	Upstream Metal Manufacturing	0.097
12	Production Technology and Heavy Machinery	0.085
13	Downstream Chemical Products	0.068
14	Recreational and Small Electric Goods	0.017
15	Local Personal Services (Non-Medical)	0.013
16	Marketing, Design, and Publishing	0.010
17	Local Health Services	0.004
18	Local Household Goods and Services	0.003
19	Financial Services	0.0012
20	Hospitality and Tourism	0.00115
21	Insurance Services	0.00038
22	Local Financial Services	0.00035
23	Jewelry and Precious Metals	0.0002
24	Aerospace Vehicles and Defense	6.933e-05
25	Business Services	4.780e-05
26	Performing Arts	3.934e-06
27	Local Retailing of Clothing and General Merchandise	3.601e-07
28	Local Logistical Services	2.191e-07
29	Medical Devices	0
30	Communications Equipment and Services	-9.785e-06
31	Biopharmaceuticals	-0.002
32	Local Utilities	-0.057
33	Lighting and Electrical Equipment	-0.074
34	Upstream Chemical Products	-0.023

Table 3.3 still has not considered that adding a job costs differently across industries. This is taken care of in Table 3.4, which divides the magnitudes in Table 3.3 by average annual wages in these industries in the unit of $1,000,000. Average wage is

used to measure the cost of adding a job. Under the assumption of perfect market competition, the social cost of a job equals the wage and the productivity of the worker. Thus, if cluster policies generate jobs in non-clusters to create greater employment density, the social cost of such policies equals the number of jobs added multiplied by average wage. The wage data are compiled from the Occupational Employment Statistics, Bureau of Labor Statistics[21]. They are averaged across all occupations in every four-digit NAICS code industry. I match them to the industrial clusters using the first four digit of each six-digit sub-industry in a cluster, and then average over all sub-industries in a cluster. This type of aggregation clearly creates measurement errors, but it is still the best I can do to measure the costs. However, it is reassuring that the measurement errors are unlikely to drive the results, as the exercise in Table 3.4 only slightly changes the ranking in Table 3.3. For example, No. 4 Information Technology and Analytical Instruments and No. 5 Automotive in Table 3.3 flip their rankings in Table 3.4. While Information Technology encourages more innovation for every job added, jobs in this industry are much more expensive as well. The ranking in Table 3.4 can be interpreted as a cost-effectiveness analysis. It shows that given the same social cost, which industrial clusters generate more citation-weighted patent applications. Policymakers, by consulting to this ranking, can prioritize certain industries over others to maximize the gain in innovation for every penny they spend.

[21] https://www.bls.gov/oes/data.htm

Table 3.4 Ranking of industrial clusters by effect on innovation adjusted by costs (number of jobs that a cluster exceeds a non-cluster multiplied by average wages)

Rank	Cluster	Effect on innovation adjusted by costs (‰ per $1,000,000 costs)
1	Metalworking Technology	36.773
2	Downstream Metal Products	19.952
3	Distribution and Electronic Commerce	11.745
4	Automotive	11.237
5	Information Technology and Analytical Instruments	5.632
6	Wood Products	5.163
7	Construction Products and Services	4.840
8	Education and Knowledge Creation	4.782
9	Plastics	3.842
10	Food Processing and Manufacturing	2.835
11	Upstream Metal Manufacturing	2.005
12	Production Technology and Heavy Machinery	1.391
13	Downstream Chemical Products	1.076
14	Local Personal Services (Non-Medical)	0.421
15	Recreational and Small Electric Goods	0.341
16	Local Health Services	0.126
17	Marketing, Design, and Publishing	0.123
18	Local Household Goods and Services	0.078
19	Hospitality and Tourism	0.037
20	Financial Services	0.011
21	Insurance Services	0.0055
22	Jewelry and Precious Metals	0.0053
23	Local Financial Services	0.005
24	Business Services	0.001
25	Aerospace Vehicles and Defense	0.0009
26	Performing Arts	7.924e-05
27	Local Retailing of Clothing and General Merchandise	1.121e-05
28	Local Logistical Services	4.615e-06
29	Medical Devices	0
30	Communications Equipment and Services	-0.0001
31	Biopharmaceuticals	-0.024
32	Local Utilities	-0.758
33	Lighting and Electrical Equipment	-1.477
34	Upstream Chemical Products	-3.449

3.7 Conclusion

While previous students find that 1) clusters in different industries exhibit heterogeneous effects on innovation, and 2) the distance of knowledge spillovers differs across industries, a complete comparison of industries in these two aspects is still missing. This chapter ranks 34 industrial clusters by their effects on innovation and identifies the optimal geographical scale for each cluster in which it improves innovation by the greatest magnitude.

Using the establishment data and the patent application data for the state of Maryland, 2004-2013, this chapter finds the following. First, the effect of most industrial clusters on innovation is locally concentrated. 24 out of the 34 industrial clusters studied in this chapter exhibit the strongest effects on patent applications in a one-mile radius region. Second, the variation in the effects on innovation is striking across industries. The top cluster, Metalworking Technology, when formed at its optimal scale of three miles in radius, uplifts the number of citation-weighted patent applications by 17.6 percent. Food Processing and Manufacturing cluster and Automotive cluster rank the second and third. When formed at one-mile and two-mile radius optimal scales, respectively, they increase patent filings by 8.5 and 5.5 percent. These effects are substantial, equivalent to a 11% to 35% increase in R&D expenditure, or a 32% to 100% increase in highway stock (Aghion, Van Reenen, and Zingales, 2013; Bloom, Schankerman, and Van Reenen, 2013; Agrawal, Galasso, and Oettl, 2017). However, after the top ten industrial clusters, the rest 24 clusters exhibit only a minimal effect on innovation no more than one percent. Thus, cluster policies do not encourage innovation for every industry and policymakers should make a careful pick. This chapter ranks the 34 industrial clusters by their effects on innovation and takes the costs into consideration. This ranking provides a direct guide for policymakers to target industries for cluster policies.

This chapter also examines industrial characteristics that explain the heterogeneous effects on innovation and the various spatial scales of clusters. It finds that the importance of tacit knowledge and a small market size accounts for a more local scale of knowledge spillovers, while a higher patent intensity leads to a larger scale. The importance of tacit knowledge is also associated with a stronger effect on innovation, as is a larger market size and a higher education level of employees.

This chapter conveys three messages to practitioners. First, for policymakers, it's important to recognize that the effects of industrial clusters on innovation prevail at different, but in general small, geographical scales. Cluster policies can be designed in line with these scales to maximize their outcomes. Second, it is wise to target certain industries for cluster policies, as most industrial clusters do not exhibit a large positive impact on innovation. Table 3.4 in this chapter provides a reasonable cost-effectiveness analysis that policymakers can directly consult. While this ranking may not be directly applicable to states other than Maryland and the industry classification does not apply to countries other than the United States, the same methodology can be applied to analyze similar issues in other regions. Moreover, the analysis on how industrial characteristics shape the geographical scales of clusters and their effects on innovation is generalizable to regions that classify industries with a different system.

Policymakers in other regions can simply target cluster policies towards industries for which tacit knowledge is important, workers are highly educated, and market size is large. Third, this chapter also provides guidance to establishment managers in terms of location choices. It helps predict whether and by how much their establishments would benefit from a cluster location. Establishment mangers can weigh the benefits against the costs of such a location to make a knowledgeable location choice.

Chapter 4: Do Clusters Benefit Small Establishments More Than Large Ones?

4.1 Introduction

The effect of clusters on innovation lands unequally on small and large establishments. Most prior studies concluded that clusters benefit small firms more than large ones. For example, Baten, Spadavecchia, Streb et al. (2007) found that inter-industry externalities are important for small firm innovation, but not for that of large firms. To be specific, if employment in other industries increases by one percent, the number of important patents per worker in the cluster increases by 5.229. Beugelsdijk (2007) found that if regional R&D intensity (R&D expenditure divided by gross regional product) increases by one percent, the share of new products in total sales in small firms increases by 7.11 percent. For large firms, the increase is only 4.48 percent and statistically insignificant. Fang (2015), in a meta-analysis that summarizes all empirical work since 1980, concluded that the positive effect of clusters on innovation lands exclusively on small firms, while large firms benefit little.

Researchers are particularly interested in small firms due to their importance to the economy. Small firms with fewer than a hundred employees account for the vast majority of the establishment population in the United States (Ayyagari, Demirguc-Kunt, and Maksimovic, 2011; Neumark, Wall, and Zhang, 2011). Small firms may require different levels of regional support compared to large firms. As contextualized by Baten, Spadavecchia, Streb et al. (2007), small firm innovation is more related to the wide base of external knowledge environment, from which innovative ideas might diffuse. Thus, supporting a clustering environment could be more beneficial for small establishments. However, existing studies that compare the effect of clusters on small versus large firms, as mentioned above, are still rare and suffer from endogeneity problem. Unobserved variables, such as the abilities and charisma of firm founder, may have simultaneously affected firm size and innovation. This problem would lead to a biased estimate.

This chapter, using the size and usage types of land parcels in 1973 as instruments, solves the endogeneity problem. I find that the conclusion of previous studies holds numerically but not statistically. Industrial clusters within one mile in radius increase citation-weighted patent filings by 2.3 percent in small establishments, but only by 0.5 percent in large establishments. However, the numerical difference between small and large establishments is statistically insignificant. This chapter also studies the attenuation of knowledge spillover over distance. Consistent with previous studies (Wallsten, 2001; Arzaghi and Henderson, 2008; Rosenthal and Strange, 2003), it finds that the effect of clusters on innovation diminishes with distance and wanes almost completely beyond five miles.

The contribution of this chapter is three-fold. First, this chapter introduces new instrumental variables to relieve the endogeneity issue. Second, it adds to the literature on the determinants of small establishment innovation (Acs and Audretsch, 1988; Van Dijk, Den Hertog, Menkveld et al., 1997). It finds that industrial clusters play an important role in boosting innovation in small establishments. Third, by looking directly into the sample attrition issue, it also provides an additional piece of evidence concerning how much clusters affect establishment mortality and relocation. This adds to the literature such as De Silva and McComb (2012) and Nathan and Vandore (2014).

4.2 Literature

Compared to large firms, small establishments have both advantages and disadvantages in gaining innovativeness from clusters. Nieto and Santamaría (2010) pointed out that the strengths of large firms reside in their abundant resources, whereas those of small businesses lie in their entrepreneurship, flexibility, and rapid response to market conditions (Lewin and Massini, 2003; Schumpeter, 1942). As clusters provide common resources at a low cost, they may help small establishments overcome their disadvantages. However, compared to large businesses, small establishments also lack financial channels and are less likely to be adequately equipped (Cohen and Klepper, 1992). They also face limitations on internal human capital and innovation capabilities (Hewitt-Dundas, 2006; Rogers, 2004). They thus may be at a disadvantageous position to improve innovation in clusters. As Giuliani (2005) pointed out, firms that are "cognitively" distant from innovators are less likely to benefit from the knowledge pool.

Empirically, there is no clear-cut relationship between firm size and innovation per se (Rothwell and Dodgson, 1994). But some evidence suggests the determinants of innovation to differ in small and large firms (Acs and Audretsch, 1988; Van Dijk, Den Hertog, Menkveld et al., 1997). Prior research also suggests that clusters have different effects on innovation in small and large firms (Freel, 2003; Baten, Spadavecchia, Streb et al., 2007). While small businesses have limited internal resources, clusters may provide them more external resources and collaboration opportunities to make up (Nieto and Santamaría, 2010). On the contrary, large firms rely more on internal resources and benefit less from clusters (Agrawal, Galasso and Oettl, 2016). Similarly, Baten, Spadavecchia, Streb et al. (2007) suggested that small firm innovation depends more on the external knowledge environment. These arguments are confirmed by Acs, Audretsch and Feldman (1994) showing that university R&D plays a more decisive role in innovation for small firms, while corporate R&D plays a relatively more important role in large firm innovations. Similarly, Beugelsdijk (2007) found that if regional R&D intensity increases by one percent the share of new products in total sales in small firms increases by 7.11 percent. For large firms, the increase is only 4.48 percent and statistically insignificant. Wallsten (2001) found that the number of firms within one-tenth mile in distance predicts whether a firm wins the Small Business Innovation Research (SBIR) awards. One additional firm within that close proximity, a small firm becomes 1.4 percent more likely to win a SBIR award. Baten, Spadavecchia, Streb et al. (2007) found if employment in other industries increases by one percent, the number of important patents per worker in the cluster increases by 5.229. In addition, if employment in the same industry and innovative firms increases by one percent, the number of important patents per worker in the cluster increases by 0.003; the effect is enlarged to 0.005 for small businesses but also becomes statistically insignificant.

They also found that employment in same industry and non-innovative firms hurts innovation in large firms, but encourages innovation in small ones, though insignificantly so. Fang (2015), in a meta-analysis that summarizes all empirical work since 1980, concluded that the positive effect of clusters on innovation lands exclusively on small firms, while large firms benefit little. Opposite conclusions also exist, but are relatively rare. Baptista and Swann (1998), using a dataset containing mostly large firms, found that own industry employment is significantly associated with firm innovation.

A major problem in previous studies is that they fail to address the endogeneity between establishment size and innovation. Establishment size affects how much an establishment benefit from clusters, but a more innovative establishment also grow faster. This leads to an overstatement of clusters' effect on large establishments and an understatement of that on small establishments, as the most successful small establishments quickly join the pool of large establishments. This problem is mitigated using the initial size of establishments at birth instead of their current size, but since most business data do not trace back to the date of establishment birth, previous studies remain muted on this issue. The overall bias on how much clusters improve small establishment innovation is ambiguous.

This chapter improves the estimates in prior studies with an instrumental variable approach. Using the 1973 land parcel size and use types as instruments, this chapter addresses the endogeneity issues. It also directly examines the effect of clusters on establishment bankruptcy and relocation to address the sample attrition issue. Thus, this chapter delivers a more precise estimate on how much clusters help small establishments innovate.

4.3 Data and Method

4.3.1 Data

I adopt the same datasets as described in section 2.3.1. I also obtain the digital maps of land parcels in 1973 for the state of Maryland from the Maryland Department of Planning[22]. These maps contain information on parcel size and parcel use types. I overlap these maps with the establishment maps and match each establishment with the parcel it resides.

4.3.2 Method

Clusters are measured the same way as in section 3.3.2, and I change the radius of clusters from one to, two, five and ten miles. This chapter adopts a linear regression following prior studies (Wallsten, 2001; Baten, Spadavecchia, Streb et al., 2007).

$$Pat_{int} = \alpha_1 Cluster_{int} + \alpha_2 Cluster_{int} * Large_{int} + \alpha_3 Large_{in} + \alpha_4 Payroll_{int} + \boldsymbol{a_i} + \boldsymbol{a_t} + \varepsilon_{int}$$

Pat_{int} denotes innovation of establishment i in industry n year t, measured by number of per-year citation-weighted patent applications. Since y contains lots of zeros, I adopt a Tobit regression to deal with this feature. $Cluster_{int}$ is a dummy variable indicating whether establishment i locates in a cluster in industry n year t.

[22] http://planning.maryland.gov/OurProducts/downloadFiles.shtml

$Large_{int}$ is a dummy variable indicating whether establishment i is a large establishment. This is measured either in year t or at the year of establishment i's birth. The definition for small and large establishments varies by each six-digit NAICS code industries, according to the U. S. Small Business Administration[23]. A plant of a larger establishment is different from a self-standing establishment, and a dummy variable should be added to account for their difference. Nonetheless, with a panel data structure, such difference will be factored out by establishment fixed effects.

Aside from these key variables, I have also controlled for establishment average payroll in dollars ($Payroll_{int}$) and establishment and year fixed effects ($\boldsymbol{a_i}$ and $\boldsymbol{a_t}$). These fixed effects take care of establishment level time-invariance characteristics and common trend. Other variables adopted in previous studies (Baten, Spadavecchia, Streb et al., 2007), such as R&D investment and export intensity, are unavailable in the QCEW dataset. As a result, they are included in the error term ε_{int}. The error term thus may be correlated with establishment size.

Aside from the omitted establishment-level observable characteristics, establishment size may be correlated with unobservable characteristics as well. These issues are addressed by instrumental variables. This chapter employs the 1973 parcel size and usage types as instruments. Parcel size affects establishment size. A small parcel restricts an establishment's ability to physically expand in place, as neighboring parcels may or may not be available. Even if they are available, it may be costly to

[23] https://www.sba.gov/contracting/getting-started-contractor/make-sure-you-meet-sba-size-standards/table-small-business-size-standards

consolidate them. Of course, a establishment can still expand vertically, but that also imposes a higher cost. Land use types[24] also affect establishment size; they restrict a establishment's ability to physical expand due to the cost of converting land use. All else equal, a establishment on industrial or commercial land is more likely to expand than a establishment on the edge of farms.

Using current parcel size and usage types as instruments, however, violates the exclusion restriction. Parcel size and usage are determined partly the political process of urban planning, and politics can be affected by current residing establishments. Therefore, I trace parcel size and usage types back to 1973, when the majority of current establishments did not even exist[25]. Past parcel size and usage clearly affect current size and uses. Also, they are determined by past political negotiations independent from current residing establishments. This satisfies the exclusion restriction. These instrumental variables are constructed by overlaying the 1973 land parcel map with the location of establishments and clusters. I obtain the exact parcel where an establishment resides.

In addition to the instrumental variable approach, I also measure establishment size at their birth. The initial size at an establishment's birth, compared to its current size, suffers less from the endogeneity issue, since the initial size cannot be a result of how much an establishment improves in clusters. To make sure I correctly measure establishment initial size, in that particular specification I limit the sample to a subset

[24] Land use types classify whether a land is low-, medium- or high-density residential, commercial, industrial, institutional, open urban land, deciduous forest, water, etc. A full list can be found in Appendix Table A1.
[25] The average longevity of a firm in the United State is only ten years; this is also consistent with the data used in this chapter, in which only 3% of establishments exist throughout the ten-year study period.

of establishments born during the study period. It's worth noting that the initial size specification may still suffer from omitted variable biases, including the omission of observable establishment characteristics and unobservable characteristics such as the inherent innovative capability of an establishment. A more capable establishment may start at a larger size than one less capable. Thus, I still apply the instrumental variables to this specification.

4.4 Results

4.4.1 Descriptive statistics

Small establishments dominate in Maryland. Among the total of 916,916 establishment-year observations, 734,222 (80%) can be classified as small or large according to the standards published by the U. S. Small Business Administration. Among these classified establishments, 733,742 (99.93%) are small establishments and only 480 (0.07%) are large ones. This is not much different from the US on average, where 99.7% of establishments are small[26].

Only a small portion of establishments are innovative, and large establishments are more innovative than small ones. Among the 723,222 classified establishments, only 0.13% applied for patents during the study period; they on average applied for 3.8 patents over ten years. Large establishments are more innovative than small ones: 2.7% of them applied for patents during the study period, compared to only 0.12% among small establishments; the difference is significant at the 5% level under a t-

[26] https://www.sba.gov/sites/default/files/FAQ_Sept_2012.pdf

test. This signals the need to encourage innovation in general, and especially in small establishments.

Establishment turnover is frequent. 24% of these 723,222 establishments only appear once in the data: 18% only survive for one year and the other 6% only exist in 2004. Those that only exist for 2004 may only survive for one year, or last for several years but went bankrupt or relocated outside of Maryland in 2004. 52% of the establishments appear for no more than three years. Only about 3% endured the entire study period. This is consistent with the estimates of average establishment longevity in previous studies[27]. It implies the seriousness of the sample attrition issue.

More than half of the establishment population is born during the study period. Among the 723,222 establishment-year observations, there are 231,874 unique establishments. 115,962 (56%) are born during 2005 to 2013. There are likely also establishments born in 2004, but I cannot separate them from those born before 2004. This large set of new establishments provides an opportunity to test the relationship between establishment initial size and innovation. This helps to mitigate the concern that establishments that benefit the most from clusters grow in size.

Table 4.1 compares the characteristics of establishments in and out of clusters. It demonstrates the significant benefits of clusters in fostering innovation, and sheds light on the severity of the sample attrition issue. The first three rows compare the total number of establishments in and out of clusters. 388,500 (42.4%) establishments locate in clusters, while the rest resides in non-clusters. A slightly greater share (44%)

[27] http://fortune.com/2015/04/02/this-is-how-long-your-business-will-last-according-to-science/

of small establishments and a smaller share of large establishments (40.8%) locate in clusters. This may be driven by small establishments benefiting more than large ones from clusters. The following six rows focus on establishment innovation. In general, 0.16% of establishments in clusters apply for patents, and they on average apply for 0.006 citation-weighted patents; out of clusters, only 0.09% establishments apply and they apply for only 0.003 citation-weighted patents on average. Both differences are statistically significant at the 5% level under a t-test. Similar patterns hold for small establishments. For large establishments, while the numeral difference enlarges, it also becomes statistically insignificant. This indicates that the impact of clusters on large establishments is more volatile: A few large establishments greatly uplift their innovation in clusters, but this is not the general pattern. The bottom seven rows examine establishment survival, birth and growth. As expected, establishments in clusters sustain a significantly lower survival rate. This is due to the tougher competitions in clusters for customers, labor and space. This pattern again holds for small establishments as well as small and patenting establishments, but becomes insignificant for large establishments. The bottom three rows show that while clusters significantly deter establishment birth, startups in clusters grow twice as fast as those outside, and their patent filings also increase at a 0.6% faster rate.

Table 4.1 Establishment characteristics in and out of clusters

Characteristics/Group	Cluster	Noncluster
Number of establishments	388,500	528,416
Number of small establishments	323,080	410,662
Number of large establishments	196	284
Percentage of establishments applying for patents	0.16%***	0.09%
Percentage of small establishments applying for patents	0.16%***	0.09%
Percentage of large establishments applying for patents	3.57%	2.11%
Average number of citation-weighted patent applications	0.006**	0.003
Average number of citation-weighted patent applications in small establishments	0.005**	0.003
Average number of citation-weighted patent applications in large establishments	2.964	0.332
Establishment survival rate	83.01%***	83.80%
Small establishment survival rate	83.47%***	84.01%
Small and patenting establishment survival rate	92.55%**	88.86%
Large establishment survival rate	88.07%	89.73%
Share of new establishments per year	25.67%***	26.88%
The yearly growth rate of new establishments	16.62%***	8.84%
The yearly growth rate in citation-weighted patent applications of new establishments	38.78%**	38.19%

NOTE: * p<0.1; ** p<0.05; *** p<0.01. Robust standard errors in parentheses.

4.4.2 Baseline results

Table 4.3 shows the baseline result, with establishments measured at current size. Table 4.2 presents descriptive statistics. The first column includes only the dummy variable for cluster status. An establishment in clusters increase citation-weighted patent applications by 62.2%; this improvement is both numerically large and statistically significant at the 1% level. The second column adds the dummy variable for large establishments as well as the interaction between large establishments and cluster status. While clusters encourage innovation by 56% in small establishments, the effect is 90.5% for large establishments. The additional benefit for large establishments is statistically insignificant though. This finding is in contradiction with previous studies, most of which found greater benefits of clusters on small establishments. This contradiction may be caused by the endogeneity issue that small establishments that benefit most soon grow into large ones.

Indeed, the sign flips after adding establishment fixed effects in columns (4) to (6). Column (2) also shows that large establishments file for four times as many patents as small ones. The third column includes additional control variables, including payroll and establishment ownership type. The result remains qualitatively similar, but the effect of clusters on small establishments halves. Columns (4) and (5) add establishment and year fixed effects, respectively. Exploiting the panel feature of the data allows me to track individual establishments over time. This mitigates the problem of sample selection. The results are in line with previous studies. The benefits of clusters on innovation, an improvement of 4.6% to 6.5%, are completely captured by small establishments. Large establishments actually do worse in clusters, probably due to the leakage of their knowledge to local competitors. Moreover, the effect of clusters on innovation is smaller in these columns compared to without fixed effects. This indicates serious upward bias from the sample selection issue in the cross-sectional model. Note that columns (4) and (5) exclude payroll from control variables because the large volume of missing data, but the result remains qualitatively similar after adding it in column (6).

Table 4.2 Descriptive statistics

Variables	Mean	Standard Deviation	Min	Max	Observations
y (Log Citation-weighted patent applications)	0.001	0.043	0	5.716	916,916
Cluster (yes=1, no=0)	0.424	0.494	0	1	916,916
Large (establishment current size, yes=1, no=0)	0.0007	0.026	0	1	734,222
Inilarge (establishment initial size, year=1, no=0)	0.0005	0.023	0	1	349,686
Log Payroll ($)	10.209	1.708	0	19.747	309,770
Private (ownership type, private=1, other=0)	0.995	0.068	0	1	916,916
Land size 1973 (acre)	8129.304	24750.31	1.600	322010	458,507

Table 4.3 The effect of clusters on innovation in small versus large establishments, with establishment size measured by current size

	Log Pat					
	(1)	(2)	(3)	(4)	(5)	(6)
Cluster	0.622***	0.560***	0.263*	0.046***	0.065***	0.075***
	(0.068)	(0.073)	(0.157)	(0.007)	(0.006)	(0.011)
Cluster*Large		0.345	1.687	-0.276***	-1.240***	-0.656***
		(0.887)	(1.594)	(0.005)	(0.200)	(0.221)
Large		4.007***	-1.180	1.452	1.196	1.051
		(0.628)	(1.256)	(1.764)	(1.343)	(1.346)
Log Payroll			0.972***			0.416***
			(0.067)			(0.035)
Private			1.779			
			(1.343)			
Establishment fixed effects	N	N	N	Y	Y	Y
Year fixed effects	N	N	N	N	Y	Y
R^2	0.005	0.008	0.166	0.005	0.004	0.157
Observations	916,916	734,222	237,931	734,222	734,222	237,931

NOTE: * p<0.1; ** p<0.05; *** p<0.01. Robust standard errors in parentheses.

While fixed effects mitigate sample selection, if the benefits an establishment receive grow over time, then the results once again will be biased. For example, if an establishment grows from a small establishment into a large one, and at the same time, the benefits it obtains from clusters also increase, then a spurious relationship between establishment size and cluster benefit would still be present, even after controlling for establishment fixed effects. Since establishments improve innovation through face-to-face communications and network buildings, it is likely that they reap more benefits as they age and develop more sophisticated relationships with other establishments in clusters. To eliminate this type of bias, Table 4.4 measures establishment size at birth. The results are qualitatively consistent with the last three columns of Table 4.3, but the disadvantages of large establishments become insignificant. In the fixed effects model, columns (3) and (4), clusters are not significantly improving establishment innovation.

Table 4.4 The effect of clusters on innovation in small versus large establishments, with establishment size measured at birth

	Log Pat			
	(1)	(2)	(3)	(4)
Cluster	0.607***	0.171***	0.064	0.417
	(0.095)	(0.058)	(0.104)	(0.558)
Cluster*Large	-1.060	-0.126		
	(1.279)	(0.447)		
Large	4.224***	12.487***		
	(0.861)	(0.447)		
Log Payroll		0.750***		
		(0.005)		
Ownership		-0.303***		
		(0.070)		
Establishment fixed effects	N	N	Y	Y
Year fixed effects	N	N	N	Y
R^2	0.009	0.154	0.0001	0.002
Observations	349,686	69,465	349,686	349,686

NOTE: * p<0.1; ** p<0.05; *** p<0.01. Robust standard errors in parentheses.

4.4.3 Instrumental variable approach

While using establishment size at birth mitigates the sample selection issue, it does not address the endogeneity problem that the initial size of establishments may be a function of other establishment unobservable characteristics. This problem needs to be handled through an instrumental variable approach. As mentioned above, this chapter uses the 1973 parcel size and use types as instruments.

Table 4.5 shows the results for the first stage regression between the instruments and establishment current size. When entering parcel size and usage separately in columns (1) and (2), both statistically significantly affect establishment size at the 5% level.

Residing on a larger parcel indicates a greater possibility of being a large establishment. However, the partial F statistics is smaller than 10, indicating parcel size as a weak instrument. This is sensible as the average parcel size is quite large (see Table 4.2). Therefore, most parcels do not restrict establishments' abilities to expand. Only those small parcels, which constitutes a limited proportion of the overall land parcel samples, may affect establishment expansion. Columns (3) to (5) add both instruments and different combinations of control variables. While land use type remains significant, parcel size does not in columns (3) and (4). Partial F for all instrumental variables is greater than 10. Table 4.6 presents the first stage for establishment initial size. The results are essentially the same, except that parcel size does not exhibit a significant effect in columns (1) and (5). Thus, parcel size is not a particularly good instrument for establishment initial size.

Table 4.5 First stage for establishment current size

	Large				
	(1)	(2)	(3)	(4)	(5)
Land size 1973	8.01e-13*		5.78e-13	5.85e-13	6.76e-12***
	(3.53e-13)		(4.02e-13)	(3.83e-13)	(1.50e-12)
Land use type 1973 dummy variables	N	Y***	Y***	Y***	Y***
Cluster				-0.00002	-0.001***
				(0.00009)	(0.0003)
Log Payroll					0.002***
					(0.0003)
Ownership				0.001***	0.006***
				(0.00007)	(0.001)
Partial F	4.54	20.60	18.21	16.20	28.44
R^2	0.000	0.001	0.001	0.001	0.009
Observations	450,538	419,088	418,812	419,088	51,702

NOTE: * p<0.1; ** p<0.05; *** p<0.01. Robust standard errors in parentheses.

Table 4.6 First stage for establishment initial size

	Large				
	(1)	(2)	(3)	(4)	(5)
Land size 1973	1.94e-13		9.72e-13	1.05e-12	1.08e-12
	(7.23e-13)		(7.79e-13)	(7.95e-13)	(3.65e-12)
Land use type 1973 dummy variables	N	Y***	Y***	Y***	Y***
Cluster				0.0001	0.0002
				(0.0001)	(0.0006)
Log Payroll					0.003***
					(0.0006)
Ownership				0.001***	0.005***
				(0.0001)	(0.001)
Partial F	0.120	27.59	25.93	23.01	8.91
R^2	0.000	0.002	0.002	0.002	0.011
Observations	250,335	231,433	231,433	231,433	14,985

NOTE: * p<0.1; ** p<0.05; *** p<0.01. Robust standard errors in parentheses.

Table A1 in the appendix 2 further verifies the credibility of land use types as instrumental variables. With few exceptions, use types easier to be converted to industrial and commercial use are associated with a larger establishment size. These include commercial, industrial, bare ground and open urban land. In contrast, land use types that are hard to be converted, such as residential, agricultural, forest and water are more likely associated with a smaller establishment size. However, institutional land, while physically easy to be converted, turns out to be less likely associated with a large establishment. Institutional land tends to remain in institutional use, even after decades. This tendency restricts a establishment's expansion in place. Overall, this exercise validates the intuition behind these instruments.

Tables 4.7 and 4.8 show the 2SLS results. Column (1) in Table 4.7 shows that clusters increase innovation by 92.8%, while columns (2) and (3) demonstrate that only small establishments benefit from clusters, by 42% to 99.2%. The effect in column (3) is statistically insignificant though. After adding establishment and year fixed effects, the effect drops to only 2.3% in columns (4) and 1.7% in column (5). The difference between large and small establishments also become statistically insignificant. Compared to the estimates in columns (4) and (5) in Table 4.3, the upward bias, 4.2% to 5.8%, exceeds the true effect itself. This bias likely also exists in prior studies. For example, Wallsten (2001) found that one additional establishment within a one-mile radius region, a small establishment becomes 0.15% more likely to win a SBIR award. An average cluster location in this chapter has 18.38 more establishments than a noncluster location; this implies a 2.8% increase in the probability of winning a SBIR award. Although the probability of winning a SBIR award and the percentage increase in patent applications isn't directly comparable, their means in Wallsten's and my datasets are in fact similar. Thus, my estimate is about 0.5% smaller than that of Wallsten's. Baten, Spadavecchia, Streb et al. (2007) found that if employment in the same industry and innovative establishments increases by 1%, the number of important patents per worker increases by 0.003. In my data, if a location turns from a noncluster to a cluster, the employment in same the industry and innovative establishments would increase by 1.29 times. Thus, the number of important patents per worker should increase by 0.387 with Baten, Spadavecchia, Streb et al. (2007)'s estimate. In comparison, with my own estimates, clusters increase patent filings by 2.3%, equivalent to only 0.013 per worker. Beugelsdijk (2007) found that if regional R&D intensity increases by 1%, the share of new products in total sales in small establishments increases by 7.11%. Again, although not directly comparable, the elasticity estimated in this chapter (0.97%) is much smaller. In Table 4.8, where establishment size is measured at birth, the effect on establishment innovation is only 1.3% in columns (3) and (4). In columns (1) and (2), the effect remains large, 19.9% to 62.3%. This effect is completely captured by small establishments while large establishments actually are less innovative when locating in clusters.

Table 4.7 The 2SLS result with establishment current size

	Log Pat				
	(1)	(2)	(3)	(4)	(5)
Cluster	0.928*** (0.060)	0.992*** (0.059)	0.420 (0.405)	0.023** (0.011)	0.017* (0.010)
Cluster*Large		-0.938*** (0.116)	-0.841** (0.423)	-0.018 (0.034)	-0.018 (0.017)
Large		0.793*** (0.732)	0.219 (0.280)	0.108* (0.056)	0.119 (0.362)
Log Payroll			0.925*** (0.171)		-0.00003 (0.0001)
Ownership			1.035 (1.680)		
Establishment fixed effects	N	N	N	Y	Y
Year fixed effects	N	N	N	Y	Y
Observations	272,445	248,401	32,337	268,401	32,337

NOTE: * p<0.1; ** p<0.05; *** p<0.01. Robust standard errors in parentheses.

Table 4.8 The 2SLS result with establishment initial size

	Log Pat			
	(1)	(2)	(3)	(4)
Cluster	0.623*** (0.061)	0.199*** (0.031)	0.013** (0.006)	0.013** (0.006)
Cluster*Large	-0.785*** (0.167)	-0.518*** (0.031)		
Large	0.675*** (0.097)	0.953*** (0.223)		
Log Payroll		0.676*** (0.166)		
Ownership		-1.935 (1.759)		
Establishment fixed effects	N	N	Y	Y
Year fixed effects	N	N	N	Y
Observations	154,906	13,616	154,906	154,906

NOTE: * p<0.1; ** p<0.05; *** p<0.01. Robust standard errors in parentheses.

4.4.4 Sample attrition

This part directly probes into how clusters affect establishment bankruptcy and re-location. In clusters, the survival rate for small establishments is 75.02%, while that for large establishments is 79.08%. Survival here means an establishment still operates and does not relocate. In comparison, in non-clusters, the rate of survival for small establishments is 79.89%, and that for large establishments is 83.10%. All these numerical differences are statistically significant at the 5% level. As expected, clusters are tougher to survival. Both small and large establishments lower their survival rates by about 4% in clusters, though large establishments have an overall greater chance of survival. This is consistent with the fact that large establishments have more abundant internal resources and better financial channels.

Table 4.9 regresses establishment survival status from year t to year t+1 on cluster status and its interaction with establishment size. Measured by current size, large establishments are more likely to survive both in and out of clusters by 2.4% to 2.6%, though not significantly so. In the meantime, establishments are in general 0.6% less likely to survive in clusters than out of in OLS regressions and 2.7% to 4.7% so in 2SLS regressions. This is consistent with the fact that clusters sustain tougher competition. These results are robust to alternative specifications such as probit, logit and survival models. These results show the severity of sample attrition: Up to 4.7% more establishments exit from clusters than non-clusters every year. These quitters likely benefit less from clusters. Thus, an OLS regression overstates the positive effect of clusters on innovation.

Table 4.9 The effect of clusters on the survival of small versus large establishments

	Survival			
	Establishment current size		Establishment initial size	
	(1)	(2)	(3)	(4)
	OLS	2SLS	OLS	2SLS
Cluster	-0.006***	-0.027***	-0.006***	-0.047***
	(0.0009)	(0.008)	(0.001)	(0.008)
Cluster*Large	0.017	0.067		
	(0.033)	(0.057)		
Large	0.024	0.026		
	(0.025)	(0.025)		
Establishment fixed effects	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y
Within R^2/Chi^2	0.585	0.053	0.624	0.619
Observations	734,222	268,401	349,686	154,906

NOTE: * p<0.1; ** p<0.05; *** p<0.01. Robust standard errors in parentheses.

4.4.5 Geographical scale of clusters

I change the spatial scale of clusters to gauge the decay of knowledge spillover across distance. As Table 4.10 shows, from one mile (Table 4.7, column (4)) to two miles in radius, the effect of cluster actually increases by 1.7% but insignificantly so. However, expanding to five miles in radius brings down the benefits to only 0.7%. This benefit is not significantly different from zero. Further expanding to ten miles in radius kills the benefits completely. In fact, at a scale of ten miles in radius, clusters even discourage innovation by 6.8%, though insignificantly so.

These findings are broadly consistent with Wallsten (2001), who found that the effect of clusters maximizes within areas smaller than one mile in radius. It also perfectly meshes with Rosenthal and Strange (2008a), who found that human capital spillovers decline sharply over distance and almost wane beyond five miles. All these findings imply that whatever encourages innovation in clusters is very sensitive to physical proximity. One such candidate is face-to-face interactions. As Allen and Cohen (1969) and Allen and Fustfeld (1975) showed, even within a single building, physical distance still significantly affects the frequency of communication between people. Thus, at a spatial scale within five miles in radius, distance also matters in alternating the frequency of worker and establishment interactions. These human interactions likely spark ideas and boost innovation. Other factors may also play a role, such as the quality of public spaces and get-together locations (e.g., coffee shops). These places may become the hotspots for people to meet up and chat. Worker turnover may also be a candidate. Through observing and talking to workers in nearby firms, workers may accumulate enough knowledge for nearby firms to help them make a decision of job change. As when they do change jobs, they bring the knowledge from their old employer with them to the new one.

Table 4.10 Attenuation of the cluster effect over distance

	Log Pat					
	2 miles in radius		5 miles		10 miles	
Cluster	0.040*	(0.023)	0.007	(0.007)	-0.068	(0.052)
Cluster*Large	-0.010	(0.110)	-0.025	(0.041)	-0.091	(0.174)
Large	0.033	(0.029)	0.060	(0.048)	0.241	(0.719)
Establishment fixed effects	Y		Y		Y	
Year fixed effects	Y		Y		Y	
Observations	268,401		268,401		268,401	

NOTE: * p<0.1; ** p<0.05; *** p<0.01. Robust standard errors in parentheses.

4.5 Robustness Checks

Alternative definitions for clusters. To test the sensitivity of the result to the specific cluster definition, I adopt two alternative definitions: 1) employment density in related industries 1.2 times above average, and 2) establishment density in related industries above median. 1.2 is a frequently used threshold to defined clusters (Lazzeretti, Boix, and Capone, 2008; Shields, Barkley, and Emery, 2009). Using this definition, I find a stronger effect on innovation in both small and large establishments. The qualitative result, that small establishments benefit more than large ones numerically but not statistically, remain robust. Using establishment density also changes the results little. There is an 82.4% overlap between clusters defined by employment and establishment density, and thus the results are not only qualitatively but also quantitatively similar.

Alternative weights for patent applications. Aside from per year citation used in the main analysis, I adopt five other weights to check for robustness. First, I count patent grants instead of patent applications. This weight dampens clusters' effect on establishment innovation by about 0.3%. Second, I apply no weights to patent filings. This again scales back the effect on establishment innovation, with large establishments hit especially hard. This indicates that large establishments in clusters produce more heavily cited patents. Third, I weight patents by the number of citations relative to their age and average citation of patents in the same technology class, and the result remains similar. Fourth, I weight by total citations. This results in a larger effect on establishment innovation, especially on that of large establishments. This again indicates that patents applied by large in-cluster establishments are more heavily cited. Last, I exclude self-citations and the results barely change.

4.6 Conclusion

Small establishments account for the majority of the establishment population in the United States. Thus, it is essential to understand what makes small establishments successful. This chapter studies how much industrial clusters boost innovation in small versus large establishments.

Although prior studies on this topic is abundant, this chapter contributes to the literature by adopting two instrumental variables, the size and usage of land parcels in 1973, to solves the endogeneity issue. It finds that industrial clusters at the spatial scale of one mile in radius increase patent filings by 2.3% in small establishments, but only by 0.5% in large establishments. However, the difference between small and large establishments is statistically insignificant. The 2.3% increase in innovation is equivalent to a 4.6% increase in R&D investment.

This chapter also examines how quickly the effect of clusters attenuates over distance. Like previous studies, it finds sharp attenuation with all benefits gone beyond five miles. At the same time, clusters sustain tougher competitions: Small establishments are 2.7% less likely to survival in clusters than they are out of clusters. In contrast, large establishments are 4% more likely to survival in clusters than they are out of clusters, but the difference between small and large establishments is also statistically insignificant.

In general, with an instrumental variable approach, this chapter finds the difference between small and large establishments only exists numerically. This is likely due to the volatility of clusters' effect on large establishments. Some large establishments may benefit a great deal from clusters, but others do not. The heterogeneity among large establishments is worthy of further examination in future project.

The prosperity of small establishments is important for people's livelihood and economic vibrancy. This chapter finds that industrial clusters boost innovation in small establishments by a significant magnitude. Promoting business concentrations at a scale of one to two miles in radius makes a promising option for practitioners to encourage innovative activities and boost local economic development.

Chapter 5: Conclusion

5.1 Summary of Findings

This dissertation matches the establishment data with the patent data for the state of Maryland from 2004 to 2013, and quantifies how much industrial clusters encourage establishments to file for patents. I adopt novel statistical methods, including discrete and continuous quantile regressions, to separate two mechanisms in clusters: selection and learning. While learning improves establishments' innovativeness, selection forces out the least innovative establishments. Both increase the average innovativeness in clusters and therefore cannot be disentangled by an ordinary least square regression. The quantile methods, by estimating the distribution of establishment innovation, can separate learning from selection.

I find that by locating in one-mile-radius locations with employment density above median, establishments increase citation-weighted patent filings by 8 to 11 percent. This effect remains at a similar magnitude in locations of two miles in radius, but declines at a larger geographical scale and disappears beyond ten miles in radius. At the same time, selection in dense employment centers at the scale of one mile in radius reduces the chance of survival for non-innovators by 2.5%.

Learning likely comes from face-to-face interactions. That can take the form of people chatting in coffee shops about new ideas, or simply observing and talking to each other and make more knowledgeable job changes between nearby establishments. Selection, on the other hand, is likely due to the high rent at desirable locations, which pushes out less innovative establishments.

I also extend the analysis to 34 groups of related industries. I find that out of the 34 industrial clusters, only 14 significantly improve establishment innovation. The cluster that encourages establishment innovation by the largest magnitude is Metalworking Technology. In a region of three miles in radius with employment density in Metalworking Technology industries above median, establishment patent applications increase by 18%. This is equivalent to a 36% increase in R&D investment. In contrast, in Business Services, the effect is essentially zero. I find that industrial characteristics, such as the different level of reliance on tacit knowledge and the education level of employees can explain some of the heterogeneity across industries.

Finally, I examine how much industrial clusters benefit small versus large establishments in terms of innovation. I use land parcel size and usage in 1973 to instrument establishment size, to mitigate the endogeneity issue between establishment size and innovation. I find that small in-cluster establishments improve innovation numerically more than large establishments, but their differences are statistically insignificant. An average small establishment improves by 2.3%. This is consistent with the theory that small establishments may rely more than external environment, while large establishments rely more on internal resources; but the difference in the Maryland case is not statistically significant. This also implies that large establishments may be more heterogeneous, with some benefiting a lot while others don't. This is an issue worthy of further study.

5.2 Generalizability of the Results

The results of this dissertation may not generalize outside of the specific geography and time frame, just like any empirical work based on data from a particular region and period. These results can be sensitive to the particular industrial structure in the state of Maryland and the specific stage for cluster development during 2004 to 2013, but some findings are more generalizable than others.

More specifically, the conclusion that learning and selection coexist in clusters is not unique to Maryland, though also may not hold universally. Signs of learning and selection in industrial clusters have been identified in previous studies (Wallsten, 2001; De Silva and McComb, 2012), and business centers often sustain higher rent, which is one of the main sources for selection. Arimoto, Nakajima, and Okazaki (2014) identified selection in Japanese silk industrial cluster, while Combes, Duranton, Gobillon et al. (2012) did not find selection for French employment centers. Thus, while the coexistence of selection and learning is not something unique to Maryland, there may be other areas with presence of only one of these two mechanisms. The finding that the impact of industrial clusters and business centers prevails within a quite local region has some potential for generalizability. This is a result that has been confirmed by studies with establishment level data such as

Wallsten (2001) and Rosenthal and Strange (2008a) for other regions and time frames.

In contrast, the ranking of industrial clusters by their impact on establishment innovation likely does not generalize. This ranking is sensitive to the particular context of Maryland, its industrial composition, development stage and economic policies. Therefore, the Maryland ranking does not provide policy implications for other regions, and may also need to be adjusted by current trends after 2013 for Maryland practitioners to apply. However, while the ranking does not generalize, the methods can be adopted by researchers and practitioners in other regions to evaluate their own clusters and employment centers and generate their own rankings. In addition, the analysis of how industrial characteristics shape the impacts of clusters on innovation helps with generalizability. While the particular ranking is context-dependent, this analysis is based on theory and backed by data. It provides a general instruction to policymakers about what type of industries can benefit more from clustering.

Finally, the finding that small establishments may benefit more than large ones should have some generalizability, as this is not an uncommon finding in previous studies. However, since most of these prior studies have endogeneity issues and my findings do not confirm a statistically significant difference between small and large establishments, this issue should be further explored. With more studies focusing on other regions and periods, we could eventually tell whether these results hold more generally.

5.3 Policy Implications

This dissertation is closely related to policies. First, the more precisely measured causal effect of industrial clusters (and urban agglomerations) on establishment innovation can produce more rigorous cost-benefit and cost-effectiveness analyses of cluster and urban policies. As I find that an ordinary least square regression which most previous studies adopted significantly overestimates how much clusters and employment centers encourage innovation, they could justify undesirable policies. This dissertation helps avoid that problem.

Second, while industrial clusters and urban agglomerations are widely supported by local governments and non-profit organizations as a way to boost innovation, the consensus has not been reached about the geographical unit at which we encourage business concentration. This dissertation, by empirically searching for the optimal geographical scale of business centers and industrial clusters to maximize their effect on innovation, concludes that the optimal geographical scale is quite local, in general only one to two miles in radius. Therefore, policymakers could consider encourage establishments to locate in close proximity. As mentioned above, this is a result that has some potential for generalization beyond Maryland and the specific study period, so this policy recommendation is not completely restricted within Maryland.

Third, with limited public funds, policymakers are frequently faced with the challenges of industry targeting. They often pick specific industries to implement cluster policies. This dissertation conducts a systematic comparison of how much clusters improve establishment innovation across industries, and therefore can provide some guidance to policymakers to properly prioritize among industries. Some industries benefit by a great deal through clustering while others don't. Keeping all else constant, targeting the former makes sure that we obtain a more impressive outcome in terms of innovation. Note that while the methods put forward by this dissertation to facilitate industrial targeting are generalizable, the specific ranking of industrial clusters is not.

Fourth, I find selection effect in Maryland clusters and business centers. As mentioned above, this effect may also exist in clusters of other regions, but may not be an issue for all clusters. The selection effect has both benefits and costs. On the one hand, tough competition motivates establishments to improve innovativeness and productivity, and redistributes social resources such as labor and physical capital from failing establishments to successful ones. One the other, establishments being forced out of the market lay off workers, and some workers may experience a prolonged period of unemployment. Local authorities need to be aware of this effect of industrial clusters, and provide training programs to help displaced labor force.

Finally, not all establishments benefit the same from clusters. I find that small establishments benefit numerically more than large establishments. This distributional effect of industrial clusters should be considered by practitioners when implementing cluster policies. However, since the difference is statistically insignificant, this distributional effect may not be as serious as many previous studies have suggested. The generalizability of this result needs future exploration.

5.4 Limitation and Future Directions

This dissertation has several limitations. First, patent applications are not a perfect measurement for establishment innovation. Many innovative activities are not patented. As a result, in general, using patent application to measure establishment innovation would lead to an underestimate of how much industrial clusters and business centers encourage establishment innovation, though the estimates are still comparable with most previous studies that also used patents as the measurement. This issue is of particular concern when comparing the effects of clusters across industries. Different industries have significantly different patenting rates; as a result, the comparison may not be fair. At the same time, since innovative activities happen continuously, but patenting only happens at certain points of the innovation cycle, using patents to measure innovation does not control for this cycling effect. Industry life cycles may also contaminate the results. Emerging, growing and mature industries are likely to have different patenting rates. This would be a concern for Chapter 3 and makes the results sensitive to the case of Maryland in the particular time period. In the future, if more comprehensive datasets on establishment innovation become available, I would love to apply alternative measurement to examine the robustness of these results.

Second, the specific process of knowledge spillover remains a black box. In this dissertation, while I quantify the magnitude for learning, I do not reveal how learning happens and which establishments exchange knowledge with which. In my future research plan, I will visualize the network of patent citation among establishments and identify how knowledge is transmitted from one establishment to another. Qualitative case studies can also help with revealing the underlying processes.

Appendix 1: Proof for the bounded estimates of selection and learning

Let $\hat{\alpha}_{2p}$ denotes $\hat{\alpha}_2$ at percentile p. According to Angrist, Chernozhukov, and Fernández (2006),

$$\hat{\alpha}_{2p} = \tilde{\alpha}_{2p,L} + \tilde{\alpha}_{2p,S} \tag{A.1}$$

where $E(\tilde{\alpha}_{2p,L}) = \alpha_{2p,L}$ and $E(\tilde{\alpha}_{2p,S}) = \alpha_{2p,S}$, with $\alpha_{2p,L}$ and $\alpha_{2p,S}$ denote learning and selection, respectively, and assume that

$$\tilde{\alpha}_{2p,L} \geq 0, \text{ and} \tag{A.2}$$

$$\tilde{\alpha}_{2p,S} \geq 0.\text{[28]} \tag{A.3}$$

While $\tilde{\alpha}_{2p,L}$ and $\tilde{\alpha}_{2p,S}$ are unobservable, $\hat{\alpha}_{2p}$ can be estimated from equation (2.1). Thus we try to bound $\tilde{\alpha}_{2p,L}$ and $\tilde{\alpha}_{2p,S}$ with $\hat{\alpha}_{2p}$.

According to the prediction of the theory, we have

$$\tilde{\alpha}_{2p,L} \leq \tilde{\alpha}_{2p',L} \text{ and} \tag{A.4}$$

$$\tilde{\alpha}_{2p,S} \geq \tilde{\alpha}_{2p',S}, \text{ for } p < p' \tag{A.5}$$

Let $\hat{\alpha}_{2p*}$ denotes the minimum nonzero[29] $\hat{\alpha}_{2p}$ across all p's. By (A.1) and (A.3),

[28] This is not a restrictive assumption for this chapter per se, as I find almost none significantly negative effect of clusters on innovation at any percentile in any specification, but it does impose a restriction on data when applied elsewhere.

[29] $\hat{\alpha}_{2p}$ may be zero at the left-tail percentiles, as the least innovative establishments in or out of clusters may be non-innovators. These estimates are uninformative and therefore dismissed throughout this chapter.

$\hat{\alpha}_{2p*} = \tilde{\alpha}_{2p*,L} + \tilde{\alpha}_{2p*,S}$ (A.6)

$\tilde{\alpha}_{2p*,S} \geq 0$ (A.7)

By (A.6) and (A.7), $\tilde{\alpha}_{2p*,L} \leq \hat{\alpha}_{2p*}$ (A.8)

By (A.8) and (A.2), $0 \leq \tilde{\alpha}_{2p*,L} \leq \hat{\alpha}_{2p*}$ (A.9)

By (A.4) and (A.9), $\tilde{\alpha}_{2p,L} \leq \tilde{\alpha}_{2p*,L} \leq \hat{\alpha}_{2p*}$ for $p < p^*$ (A.10)

By (A.1) and (A.10), $\tilde{\alpha}_{2p,S} \geq \hat{\alpha}_{2p} - \hat{\alpha}_{2p*}$ for $p < p^*$ (A.11)

Also, by (A.1)-(A.3), (A.10) and (A.11), we obtain

$0 \leq \tilde{\alpha}_{2p,L} \leq \hat{\alpha}_{2p*}$, and (A.12)

$\hat{\alpha}_{2p} - \hat{\alpha}_{2p*} \leq \tilde{\alpha}_{2p,S} \leq \hat{\alpha}_{2p}$ for $p < p^*$

(A.13)

These two equations bound selection and learning at percentiles smaller than p*.

By (A.2), $\tilde{\alpha}_{2p*,L} \geq 0$ (A.14)

By (A.6) and (A.14), $\tilde{\alpha}_{2p*,s} \leq \hat{\alpha}_{2p*}$ (A.15)

By (A.5) and (A.15), $0 \leq \tilde{\alpha}_{2p*,s} \leq \hat{\alpha}_{2p*}$ (A.16)

By (A.5), $\tilde{\alpha}_{2p,S} \leq \tilde{\alpha}_{2p*,S} \leq \hat{\alpha}_{2p*}$ for $p > p^*$ (A.17)

By (A.1) and (A.17), $\tilde{\alpha}_{2p,L} \geq \hat{\alpha}_{2p} - \hat{\alpha}_{2p*}$ for $p > p^*$

(A.18)

Also, by (A.1)-(A.3), (A.17) and (A.18), we obtain

$0 \leq \tilde{\alpha}_{2p,S} \leq \hat{\alpha}_{2p*}$, and (A.19)

$\hat{\alpha}_{2p} - \hat{\alpha}_{2p*} \leq \tilde{\alpha}_{2p,L} \leq \hat{\alpha}_{2p}$ for $p > p^*$

(A.20)

These two equations bound selection and learning at percentiles greater than p*.

Summarizing (A.9), (A.12), (A.13), (A.16), (A.19) and (A.20), we obtain

$$\begin{cases} p = p^* \begin{cases} 0 \leq \tilde{\alpha}_{2p,L} \leq \hat{\alpha}_{2p*} \\ 0 \leq \tilde{\alpha}_{2p,s} \leq \hat{\alpha}_{2p*} \end{cases} \\ p < p^* \begin{cases} 0 \leq \tilde{\alpha}_{2p,L} \leq \hat{\alpha}_{2p*} \\ \hat{\alpha}_{2p} - \hat{\alpha}_{2p*} \leq \tilde{\alpha}_{2p,S} \leq \hat{\alpha}_{2p} \end{cases} \\ p > p^* \begin{cases} \hat{\alpha}_{2p} - \hat{\alpha}_{2p*} \leq \tilde{\alpha}_{2p,L} \leq \hat{\alpha}_{2p} \\ 0 \leq \tilde{\alpha}_{2p,S} \leq \hat{\alpha}_{2p*} \end{cases} \end{cases}$$

These inequalities finally bound selection and learning at every percentile with estimated coefficients.

Appendix 2: First stage regression result between land use type and establishment size

Table A1 Land use type and establishment size

	Establishment current size is large	Establishment initial size is large
	(1)	(2)
Easy to convert:		
Commercial	-0.0003	-0.00006
	(0.0003)	(0.0004)
Industrial	0.0004	0.004*
	(0.001)	(0.002)
Institutional	-0.001***	-0.002***
	(0.0008)	(0.0003)
Bare ground	0.001*	-0.007
	(00005)	(0.004)
Open urban land	0.001***	0.002***
	(0.00008)	(0.0003)
Hard to convert:		
Brush	-0.0004*	0.050
	(0.0002)	(0.028)
Deciduous forest	-0.001***	-0.002***
	(0.00007)	(0.0007)
Extractive	0.005***	-0.019
	(0.001)	(0.013)
High-density residential	-0.001***	-0.001***
	(0.00008)	(0.0003)
Low-density residential	-0.0005***	-0.005*
	(0.0001)	(0.003)
Medium-density residential	-0.001***	-0.001***
	(0.00008)	(0.0003)
Mixed forest	0.0008	-0.007*
	(0.004)	(0.004)
Pasture	-0.0002	-0.006
	(0.0002)	(0.004)
Water	-0.0009***	-0.002***
	(0.00008)	(0.0008)

NOTE: * $p<0.05$; ** $p<0.01$; *** $p<0.005$. Robust standard errors in parentheses.

Bibliography

Acs, Z. J., & Audretsch, D. B. (1988). Innovation in large and small firms: an empirical analysis. *The American Economic Review*, 78(4): 678-690.

Acs, Z. J., Audretsch, D. B., & Feldman, M. P. (1994). R & D spillovers and recipient firm size. *The Review of Economics and Statistics*, 76(2): 336-340.

Aghion, P., Bloom, N., Blundel, R., et al. (2002). Competition and innovation: An inverted U relationship. National Bureau of Economic Research Working Paper. http://www.nber.org/papers/w9269

Aghion, P. & Howitt P. (1992). A model of growth through creative destruction. *Econometrica*, 60(2): 323-351.

Aghion, P., Van Reenen, J., & Zingales, L. (2013). Innovation and institutional ownership. *American Economic Review*, 103 (1): 277–304.

Agrawal, A., Galasso, A., & Oettl, A. (2017). Roads and innovation. *Review of Economics and Statistics*, 99(3): 417-434.

Aharonson, B. S., Baum, J. A., & Feldman M.P. (2004). Industrial clustering and the returns to inventive activity: Canadian biotechnology firms, 1991-2000. Working Paper. http://www3.druid.dk/wp/20040003.pdf

Andersson, M., Larsson, J. P., & Lundblad J. (2015). The productive city needs both. Working Paper. http://www-sre.wu.ac.at/ersa/ersaconfs/ersa15/e150825aFinal00385.pdf

Angrist, J., Chernozhukov, V., & Fernández- Val, I. (2006). Quantile regression under misspecification, with an application to the US wage structure. *Econometrica*, 74(2): 539-563.

Arimoto, Y., Nakajima, K., & Okazaki, T. (2014). Sources of productivity improvement in industrial clusters: The case of the prewar Japanese silk-reeling industry. *Regional Science and Urban Economics*, 46: 27-41.

Arzaghi, M., & Henderson, J. V. (2008). Networking off Madison Avenue. *The Review of Economic Studies*, 75(4): 1011-1038.

Attaran, M., & Zwick, M. (1987). Entropy and other measures of industrial diversification. *Quarterly Journal of Business and Economics*, 26(4): 17-34.

Audretsch, D., & Feldman M. (1996). Knowledge spillovers and the geography of innovation and production. *American Economic Review*, 86 (3): 630–640.

Ayyagari, M., Demirguc-Kunt, A., & Maksimovic, V. (2011). Small vs. young firms across the world: contribution to employment, job creation, and growth. Working paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1807732

Babbage, C. (1832). *On the economy of machinery and manufactures*. London: Charles Knight.

Bacolod, M., Blum, B. S., & Strange, W. C. (2009). Skills in the city. *Journal of Urban Economics*, 65(2): 136-153.

Balconi, M. (2002). Tacitness, codification of technological knowledge and the organisation of industry. *Research Policy*, 31(3): 357-379.

Baldwin, R. E., & Okubo T. (2006). Heterogeneous firms, agglomeration and economic geography: Spatial selection and sorting. *Journal of Economic Geography*, 6(3): 323-346.

Baptista, R., & Swann P. (1998). Do firms in clusters innovate more? *Research Policy*, 27(5): 525-540

Barkley, D. L., &M. S. Henry (2005). Targeting industry clusters for regional economic development: An overview of the REDRL approach. Regional Economic Development Research Laboratory Working Paper. http://media.clemson.edu/public/extension/redrl_pubs/redrl_rpt15.pdf

Baten, J., A. Spadavecchia, J. Streb, et al. (2007). What made southwest German firms innovative around 1900? Assessing the importance of intra-and inter-industry externalities. *Oxford Economic Papers*, 59(suppl 1): i105-i126.

Beaudry, C. (2001). Entry, growth and patenting in industrial clusters: A study of the aerospace industry in the UK. *International Journal of the Economics of Business*, 8(3): 405-436.

Beaudry, C., & Schiffauerova A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy*, 38(2): 318-337.

Behrens, K., Duranton, G., & Rober-Nicoud, F. (2014). Productive cities: Sorting, selection, and agglomeration. *Journal of Political Economy*, 122(3): 507-553.

Bell, G. G. (2005). Research notes and commentaries: Clusters, networks, and firm innovativeness. *Strategic Management Journal*, 26(3): 287-295.

Berry, B. J., & Garrison, W. L. (1958). The functional bases of the central place hierarchy. *Economic Geography*, 34(2): 145-154.

Beugelsdijk, S. (2007). The regional environment and a firm's innovative performance: A plea for a multilevel interactionist approach. *Economic Geography*, 83(2): 181-199.

Blind, K., & Grupp H. (1999). Interdependencies between the science and technology infrastructure and innovation activities in German regions: empirical findings and policy consequences. *Research Policy*, 28(5): 451-468.

Bloom, N., Schankerman M., & Van Reenen J. (2013). Identifying technology spillovers and product market rivalry. *Econometrica*, 81(4): 1347-1393.

Breschi, S. (1995). Identifying regional patterns of innovation using patent data. Paper presented at the workshop on Regional Innovation Systems, Regional Networks and Regional Polocy, organized by the STEP group at Lysebu Conference Center (Oslo), Norway.

Burger, M. J., van Oort, F. G., & van der Knaap, B. (2008). A treatise on the geographical scale of agglomeration externalities and the modifiable areal unit

problem. SSRN Working Paper.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1304670

Caballero, R., & Jaffe, A. (1993). How high are the giants' shoulders? An empirical assessment of knowledge spillovers and creative destruction in a model of economic growth. *NBER Macroeconomic Annual*, 8: 15-74.

Carlino, G. A., Carr, J., Hunt, R. M., & Smith, T. E. (2012). The agglomeration of R & D labs, Federal Reserve Bank of Philadelphia Working Paper.
http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.844.9959&rep=rep1&type=pdf

Carlino, G., Chatterjee, S., & Hunt, R. (2007). Urban density and the rate of invention. *Journal of Urban Economics*, 61 (3): 389-419.

Carlino, G., & Kerr, W. R. (2014). Agglomeration and innovation. National Bureau of Economic Research Working Paper.
http://www.hbs.edu/faculty/Publication%20Files/15-007_e181fd00-4426-4db8-8f70-89b1b5054a8f.pdf

Chen, H. S. (2011). The Relationship between technology industrial cluster and innovation in Taiwan. *Asia Pacific Management Review*, 16(3): 277-288.

Chiang, J. T. (1993). From industry targeting to technology targeting: A policy paradigm shift in the 1980s. *Technology in Society*, 15(4): 341-357.

Chrisinger, C. K., C. S. Fowler, & R.G. Kleit. (2015). Industry Clusters and Employment Outcomes in Washington State. *Economic Development Quarterly*, 29(3): 199-210.

Cohen, W. M., & Klepper, S. (1992). The anatomy of industry R&D intensity distributions. *American Economic Review*, 82(4): 773–799.

Combes, P.P., Duranton G., & Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, 63(2): 723-742.

Combes, P.P., Duranton, G., Gobillon, L., et al. (2012). The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica*, 80(6): 2543-2594.

Conroy, T., Deller, S., & Tsvetkova, A. (2016). Regional business climate and interstate manufacturing relocation decisions. *Regional Science and Urban Economics*, 60 (9): 155-168.

Copaken, R. D. (1982). The houdaille petition: A new weapon against unfair industry targeting practices. *The George Washington Journal of International Law and Economics*, 17(2): 211-247.

Criscuolo, P., & Verspagen B. (2008). Does it matter where patent citations come from? Inventor vs. examiner citations in European patents. *Research Policy*, 37(10): 1892-1908.

Dahlman, C.J. (1979). The problem of externality. *Journal of Law and Economics*, 22 (1): 141-162.

Dasgupta, P., & Stiglitz, J. (1980). Industrial structure and the nature of innovative activity. *The Economic Journal*, 90 (358): 266-293.

Dechezleprêtre, A., Martin, R., & Mohnen, M. (2014). Knowledge spillovers from clean and dirty technologies. Working Paper. http://cep.lse.ac.uk/pubs/download/dp1300.pdf

De Silva, D. G., & McComb, R. P. (2012). Geographic concentration and high tech firm survival. *Regional Science and Urban Economics*, 42(4): 691-701.

Delgado, M., M.E. Porter, & S. Stern. (2010). Clusters and entrepreneurship. *Journal of Economic Geography*, 10 (4): 495–518.

Delgado, M., Porter, M.E., & Stern, S. (2014). Clusters, convergence, and economic performance. *Research Policy*, 43(10): 1785-1799.

Deyle, H.G., & Grupp, H. (2005). Commuters and the regional assignment of innovative activities: A methodological patent study of German districts. *Research Policy*, 34(2): 221-234.

Dubey, P., & Wu, C. W. (2002). When less competition induces more product innovation. *Economics Letters*, 74(3): 309-312.

Duranton, G., & Overman, H. G. (2005). Testing for localization using micro-geographic data. *The Review of Economic Studies*, 72(4): 1077-1106.

Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In Henderson J. V. & Thisse, J. F. (Eds.), *Handbook of Regional and Urban Economics*, vol. 4, in: Cities and Geography, Elsevier, New York, pp. 2063-2117.

DRI/McGraw-Hill. (1995). *America's clusters*. DRI/McGraw-Hill, Lexington MA.

Economics and Statistics Administration and USPTO. (2012). Intellectual property and the U.S. economy: Industries in focus. https://www.uspto.gov/sites/default/files/news/publications/IP_Report_March_2012.pdf

Ellison, G., Glaeser, E., & Kerr, W.R. (2010). What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review*, 100(3): 1195–1213.

Elvery, J. A. (2010). City size and skill intensity. *Regional Science and Urban Economics*, 40(6): 367-379.

Enright, M.J. (1993). The geographic scope of competitive advantage. Division of Research, Harvard Business School.

Fabiani, S. & Pellegrini, G. (1998). Un'analisi quantitativa delle imprese nei distretti industriali italiani: Redditività, produttività e costo del lavoro. *L'industria*, 19 (4): 811-832.

Falck, O., Heblich, S., & Kipar, S. (2010). Industrial innovation: Direct evidence from a cluster-oriented policy. *Regional Science and Urban Economics*, 40(6): 574-582.

Fang, L. (2015). Do clusters encourage innovation? A Meta-analysis. *Journal of Planning Literature*, 30(3): 239-260.

Farre-Mensa, J., Hegde, D., & Ljungqvist, A. (2016). The bright side of patents. National Bureau of Economic Research Working Paper. https://www.uspto.gov/sites/default/files/documents/Patents%20030216%20USPTO%20Cover.pdf

Feldman, M. P. (1994). *The geography of innovation*. Dordrecht, the Netherlands: Kluwer Academic.

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

Feldman, M. P. & Francis, J. L. (2004). Homegrown solutions: Fostering cluster formation. *Economic Development Quarterly*, 18(2): 127-137.

Ferrand, Y., Kelton, C. M., Chen, K., et al. (2009). Biotechnology in Cincinnati clustering or colocation? *Economic Development Quarterly*, 23(2): 127-140.

Feser, E. J., & Luger, M. I. (2003). Cluster analysis as a mode of inquiry: Its use in science and technology policymaking in North Carolina. *European Planning Studies*, 11 (1): 11–24.

Feser, E.J., H. Renski, & H. Goldstein (2008). Clusters and economic development outcomes: An analysis of the link between clustering and industry growth. *Economic Development Quarterly*, 22 (4): 324-344.

Freel, M. S. (2003). Sectoral patterns of small firm innovation, networking and proximity. *Research Policy*, 32(5): 751-770.

Frenken, K., Van Oort, F., Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional Studies*, 41(5): 685-697.

Fritsch, M., & Slavtchev, V. (2010). How does industry specialization affect the efficiency of regional innovation systems? *The Annals of Regional Science*, 45(1): 87-108.

Geroski, P. (1994). *Market structure, corporate performance, and innovative activity*. Oxford University Press, New York.

Gertler, M. S. (2003). Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there). *Journal of Economic Geography*, 3(1): 75-99.

Giuliani, E. (2005). Cluster absorptive capacity: why do some clusters forge ahead and others lag behind? *European Urban and Regional Studies*, 12(3): 269-288.

Glaeser, E. L. (1998). Are cities dying? *The Journal of Economic Perspectives*, 12(2): 139-160.

Glaeser, E. L. (Eds.) (2010a). *Agglomeration economics*. University of Chicago Press, Chicago.

Glaeser, E. L. (2000b). The new economics of urban and regional growth. In Clark, G. L., Feldman, M. P., & Gertler, M. S. (Eds.), *The Oxford Handbook of Economic Geography*, pp. 83-98.

Glaeser, E. L., Kallal, H. D., Scheinkman, G. A., et al. (1991). Growth in cities. *Journal of Political Economics*, 100(6): 1126-1152.

Glaeser, E.L., & Kerr, W.R. (2009). Local industrial conditions and entrepreneurship: How much of the spatial distribution can we explain? *Journal of Economics and Management Strategy*, 18 (3): 623–663.

Glaeser, E. L., Rosenthal, S. S., & Strange, W. C. (2010). Urban economics and entrepreneurship. *Journal of Urban Economics*, 67(1): 1-14.

Gobillon, L., & Roux, S. (2008). Quantile-based inference of parametric transformations between two distributions. Processed, crest-insee. http://jms.insee.fr/files/documents/2009/80_2-JMS2009_S10-3_GOBILLON-ACTE.PDF

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

Grossman, G. M., & Helpman, E. (1993). *Innovation and growth in the global economy*. Cambridge: MIT press.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2000). Market value and patent citations: A first look. National Bureau of Economic Research Working Paper. http://www.nber.org/papers/w7741

Harhoff, D., Narin, F., Scherer F. M., et al. (1999). Citation frequency and the value of patented inventions. *Review of Economics and Statistics*, 81(3): 511-515.

Harrison, B., Kelley, M. R., & Gant, J. (1996). Innovative firm behavior and local milieu: Exploring the intersection of agglomeration, firm effects, and technological change. *Economic Geography*, 72 (3): 233–258.

Helsley, R. W., & Strange, W. C. (2002). Innovation and input sharing. *Journal of Urban Economics*, 51(1): 25-45.

Henderson, J. V. (1974). The sizes and types of cities. *The American Economic Review*, 64(4): 640-656.

Henderson, J. V. (1986). Efficiency of resource usage and city size. *Journal of Urban Economics*, 19 (1): 47–70.

Hervas-Oliver, J.L., Sempere-Ripoll, F., & Boronat-Moll, C. (2014). Process innovation strategy in SMEs, organizational innovation and performance: a misleading debate? *Small Business Economics*, 43(4): 873-886.

Hewitt-Dundas, N. (2006). Resource and capability constraints to innovation in small and large plants. *Small Business Economics*, 26(3): 257–277.

Hirschey, M., & Richardson V. J. (2001). Valuation effects of patent quality: A comparison for Japanese and US firms. *Pacific-Basin Finance Journal*, 9(1): 65-82.

Jacobs, J. (1969). *The economy of cities*. Vintage, New York.

Jacobs, J. (1984). *Cities and the wealth of nations.* Random House, New York.

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): 577-598.

Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. *Econometrica: Journal of the Econometric Society*, 46(1): 33-50.

Kramarz, F. (2000). Inter-industry and firm-size wage differentials: New evidence from linked employer-employee data, Working Paper. http://www.crest.fr/ckfinder/userfiles/files/pageperso/kramarz/regular-version-inter-industry-20000717.pdf

Krugman, P. R. (1991). *Geography and trade*. MIT press, Cambridge.

Lazzeretti, L., Boix, R., & Capone, F. (2008). Do creative industries cluster? Mapping creative local production systems in Italy and Spain. *Industry and innovation*, 15(5): 549-567.

Leatherman, J. C., Howard, D. J., & Kastens, T. L. (2002) Improved prospects for rural development: An industrial targeting system for the Great Plains. *Review of Agricultural Economics*, 24(1): 59-77.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*, 76(3): 1071-1102.

Lewin, A. Y., & Massini, S. (2003). Knowledge creation and organizational capabilities of innovating and imitating firms. In Tsoukas, H. and Mylonopoulos, N. (Eds.) *Organizations as Knowledge Systems*. New York: Palgrave, pp. 209–237.

Mariani, M. (2004). What determines technological hits?: Geography versus firm competencies. *Research Policy*, 33(10): 1565-1582.

Marshall, A. (1890). *Principles of political economy*. Maxmillan, New York.

Maskell, P. (2001). Towards a knowledge-based theory of the geographical cluster. *Industrial and Corporate Change*, 10(4): 921-943.

Nathan, M. & Vandore, E. (2014). Here be startups: Exploring London's 'tech city' digital cluster. *Environment and Planning A*, 46(10): 2283-2299.

Neffke, F., Henning, M., Boschma, R., Lundquist, K. J., & Olander, L. O. (2011). The dynamics of agglomeration externalities along the life cycle of industries. *Regional Studies*, 45(1): 49-65.

Neumark, D., Wall, B., & Zhang, J. (2011). Do small businesses create more jobs? New evidence for the United States from the National Establishment Time Series. *The Review of Economics and Statistics*, 93(1): 16-29.

Nieto, M. J., & Santamaría, L. (2010). Technological collaboration: Bridging the innovation gap between small and large firms. *Journal of Small Business Management*, 48(1): 44-69.

Packalen, M., & Bhattacharya, J. (2015). Cities and ideas. National Bureau of Economic Research Working Paper. http://www.nber.org/papers/w20921

Pathirage, C. P., Amaratunga, D. G., & Haigh, R. P. (2007). Tacit knowledge and organizational performance: construction industry perspective. *Journal of Knowledge Management*, 11(1): 115-126.

Peck, F. & McGuinness, D. (2003). Regional development agencies and cluster strategies: Engaging the knowledge-base in the North of England. *Local Economy*, 18(1): 49-62.

Renski, H. (2013). Using matched employee-employer data to measure labour mobility and knowledge flows in supply-chain and labour-based industry clusters. *Regional Science Policy & Practice*, 5(1), 25-43.

Porter, M. E. (1990). The competitive advantage of nations. *Harvard Business Review*, 68(2): 73-93.

Porter, M. E. (1998). Clusters and the new economics of competition. *Harvard Business Review*, 76(6): 77–90.

Porter, M. E. (2000). Location, competition, and economic development: Local clusters in a global economy. *Economic Development Quarterly*, 14(1): 15-34.

Porter, M.E. (2003). The economic performance of regions. *Regional Studies*. 37 (6-7): 549-578.

Reinganum, J. F. (1983). Uncertain innovation and the persistence of monopoly. *The American Economic Review*, 73(4): 741-748.

Rogers, M. (2004). Networks, firm size and innovation. *Small Business Economics*, 22(2): 141–153.

Romer, P. M. (1986). Increasing returns and long-run growth. *The Journal of Political Economy*, 94(5): 1002-1037.

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

Rosenthal, S.S., & Strange, W.C. (2004). Evidence on the nature and sources of agglomeration economies. In Henderson J. V. & Thisse, J. F. (Eds.), *Handbook of Regional and Urban Economics*, vol. 4, in: Cities and Geography, Elsevier, New York, pp. 2119–2171.

Rosenthal, S. S., & Strange, W. C. (2008a). The attenuation of human capital spillovers, *Journal of Urban Economics*, 64(2): 373-389.

Rosenthal, S. S., & Strange, W. C. (2008b). Agglomeration and hours worked. *The Review of Economics and Statistics*, 90(1): 105-118.

Rothwell, R., & Dodgson, M. (1994). Innovation and size of firm. In Dodgson, M. & Rothwell R. (Eds.) *The Handbook of Industrial Innovation*. Aldershot Hants: Edward Elgar, pp.310-324.

Rothwell, R., & Zegveld, W. (1982). Innovation and the small and medium sized firm. Working paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1496714

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

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

Sforzi, F. (1990). The quantitative importance of Marshallian industrial districts in the Italian economy. In *Industrial districts and inter-firm co-operation in Italy*, pp.75-107.

Shefer, D., & Frenkel, A. (2005). R&D, firm size and innovation: an empirical analysis, *Technovation*, 25(1): 25-32.

Shields, M., Barkley, D. & Emery, M. (2009). Industry clusters and industry targeting. In Goetz S.J., Deller, S, & Harris, T. (Eds). *Targeting Regional Economic Development*, pp. 35-44.

Simonen, J., & McCann, P. (2008). Firm innovation: The influence of R&D cooperation and the geography of human capital inputs. *Journal of Urban Economics*, 64(1): 146-154.

Singh, J., & Marx, M. (2013). Geographic constraints on knowledge spillovers: Political borders vs. spatial proximity. *Management Science*, 59(9): 2056-2078.

Staber, U. (2001). Spatial proximity and firm survival in a declining industrial district: the case of knitwear firms in Baden-Wurttemberg. *Regional Studies*, 35(4): 329–341.

Swan, P. L. (1970). Market structure and technological progress: The influence of monopoly on product innovation. *The Quarterly Journal of Economics*, 84(4): 627-638.

Teichert, N. (2013). *Innovation in general purpose technologies: how knowledge gains when it is shared*. KIT Scientific Publishing, Karlsruhe.

United Nations. (2015). World urbanization prospects: the 2014 revision. https://esa.un.org/unpd/wup/publications/files/wup2014-report.pdf

Van Dijk, B., Den Hertog, R., Menkveld, B., & Thurik, R. (1997). Some new evidence on the determinants of large-and small-firm innovation. *Small Business Economics*, 9(4): 335-343.

Vásquez-Urriago, Á. R., Barge-Gil, A., Rico, A. M., et al., 2014. The impact of science and technology parks on firms' product innovation: empirical evidence from Spain. *Journal of Evolutionary Economics*, 24(4): 835-873.

Voytek, K. & Ledebur, L. (1991). Is industry targeting a viable economic development strategy. In Bingham, R. D. & Mier, R. (Eds.), *Dilemmas of urban economic development*, pp.171-194. Thousand Oaks: Sage.

Wagner, J. (2004). Are young and small firms hothouses for nascent entrepreneurs? Evidence from German micro data. Working paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=494202

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

Wang, C. C., Lin, G. C., & Li, G. (2010). Industrial clustering and technological innovation in China: New evidence from the ICT industry in Shenzhen. *Environment and Planning A*, 42(8): 1987-2010.

Winters, J. V. (2014). Foreign and native-born STEM graduates and innovation intensity in the United States, Working Paper. http://ftp.iza.org/dp8575.pdf

Yang, X., & Ng, Y. K. (1993). *Specialization and economic organization: A new classical microeconomic framework*. Amsterdam, the Netherlands: North-Holland.

Young, A. A. (1928). Increasing returns and economic progress. *The Economic Journal*, 38 (152): 527–542.