AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

The role of interregional and inter-sectoral knowledge spillovers on regional knowledge creation across US metropolitan counties

Dongwoo Kang

Korea Labor Institute, Sejong National Research Complex

E-mail: dwkang1982@kli.re.kr

Sandy Dall'erba

Department of Agricultural and Consumer Economics and Regional Economics Applications Laboratory, University of Illinois at Urbana-Champaign

E-mail: dallerba@illinois.edu

Kun Peng

Department of Agricultural and Consumer Economics and Regional Economics Applications Laboratory, University of Illinois at Urbana-Champaign

E-mail: kunpeng2@illinois.edu

Selected Paper prepared for presentation at the 2017 Agricultural & Applied Economics Association Annual Meeting, Chicago, Illinois, July 30-August 1

Copyright 2017 by [Kang et al.]. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

Abstract: We examine the heterogeneous determinants of knowledge production across 5 manufacturing sectors of the US metropolitan counties. In addition to the traditional knowledge inputs, we capture knowledge spillovers based on an innovative matrix of patent creation-patent citation separating intra-regional from short- and long-distance interregional effects. Furthermore, we identify the singular role of MAR vs. Jacobian externalities. Our results show that an aggregated approach would mask the variation in the marginal effects detected across sectors and highlight the key role of specialization and of university R&D in innovation production. Intersectoral and long-distance interregional spillovers, usually ignored in the literature, display a significant impact too.

Keywords: knowledge production function, interregional and inter-sectoral knowledge spillovers, sectoral heterogeneity, panel Tobit model

JEL classifications: C23, O31, R11

1. Introduction

As knowledge accumulation and its spillovers are recognized as important determinants of economic growth (Romer 1986; Grossman and Helpman 1994; Jaffe 1989), the knowledge production function literature has paid an increasing amount of attention to the role and the geography of interregional knowledge spillovers (Audretsch and Feldman 2004; Anselin *et al.* 1997; Acs *et al*. 2002; Bode 2004; Autant-Bernard 2012). Yet, knowledge can diffuse over long distance via various types of channels such as inventors' network (Crescenzi *et al*., 2016), technological proximity (Maggioni *et al.*, 2010; Fischer *et al*., 2006) or labor migration (Almeida and Kogut 1999). As such, a growing number of studies has investigated the role of distant interregional knowledge spillovers (Gertler and Levitte 2005; Peri 2005; Gittelman 2007; Ponds *et al*., 2010).

Furthermore, sectorally aggregated data is commonly used in empirical estimations of regional knowledge production functions (Anselin *et al*., 1997; Fischer and Varga 2003; Bode 2004; Parent and LeSage 2008). Indeed, only a small number of studies have dealt with sectoral heterogeneity: Ponds *et al.* (2010) apply dummies to identify sectoral differences while Bottazzi and Peri (2003) proxy each sector's innovation output by using the share of that sector in total value added. Furthermore, even less studies have investigated the presence of sectoral heterogeneity in the knowledge spillovers. Jaffe (1989) and Anselin *et al.* (2000) highlight that inter-regional spillovers vary from one sector to the next but they do not consider intersectoral spillovers. Autant-Bernard and LeSage (2011) do but their panel data models do not measure how each individual sector is affected by such spillovers since they only report average marginal effects across all sectors. As such, differentiating clearly the effects of intra- vs. inter-sectoral spillovers and of intra- vs. inter-regional spillovers on each sector is necessary to shed new light on regional innovation dynamics.

In order to reach this goal, we focus on five innovative manufacturing sectors that represent about 82% of our patent data drawn from USPTO (2010): 1) Chemical, 2) Drugs & Medical, 3) Mechanical, 4) Computer & Communication, 5) Electrical & Electronic. These sectors are analyzed across 853 metropolitan counties, which gives us more spatial observations than any of the previous studies in this field since they are limited to the US state level or Metropolitan Statistical Area (MSA) level. Furthermore, the sectoral approach we use here obliges us to rely on a more sophisticated econometric approach than studies based on the aggregate level. Indeed, we need to control for the case where no knowledge output is recorded and for cross-sectional unobservable heterogeneity, hence we adopt a panel Tobit model . Last but not least, knowledge spillovers are all based on the NBER US Patent Citation Data File (Hall *et al.*, 2001) as in Sorenson *et al.* (2006) and Sonn and Storper (2008). This dataset allows us to track actual flows of knowledge from the place where they are created to the place(s) where they are cited. Capturing the direction of these flows allows us to identify explicitly the role of externalities on knowledge output, unlike the knowledge spillovers based on geographical proximity (Anselin *et al.*, 1997; Bode 2004) or collaborative work (Ponds *et al.*, 2010; Crescenzi *et al.*, 2016).

The remainder of the paper is organized as follows: Section 2 reviews the literature focusing on intra- and inter-sectoral spillovers, i.e. Marshall-Arrow-Romer (Marshall 1920; Arrow 1962; Romer 1986) and Jacobian externalities (Jacobs 1969), and their role on knowledge creation and innovation. Section 3 describes our knowledge production function, the strategy of modeling intra- and inter-sectoral knowledge spillovers across counties and the relevant data. The estimation results and their interpretation are reported in Section 4 while the last section closes the paper with some concluding remarks.

2. Literature Review

2.1. Intra- and inter-sectoral knowledge spillovers

Most studies on knowledge creation and innovation focus on local knowledge spillovers (Jaffe, 1986; Jaffe *et al.*, 1993; Feldman, 1994; Anselin *et al.*, 1997). Their local extent is usually explained by two types of externalities. The first one is Marshall-Arrow-Romer (MAR) externalities that emphasize industrial specialization within the same sector. Geographical concentration of firms within the same industry can lower the cost of communication and transaction, thereby facilitating knowledge spillovers among firms (Audretsch and Feldman 2004). The second type of local externalities is based on the opposite source: Jacobs (1969) points out that it is industrial diversity that is conducive of knowledge spillovers. Audretsch and Feldman (2004) echo his views by arguing that a diverse knowledge coming from external sectors can complement a specific sector's knowledge. As geographical proximity contributes to the exchange of ideas (Feldman and Kogler 2010) and activities have become more clustered over time (Glaeser *et al.*, 1992), intra- and inter-sectoral spillovers have played an increasing role in the creation of knowledge.

However, recent studies pay more and more attention to not only the aforementioned local dynamics but also to knowledge emanating from distant sources (Owen-Smith and Powell, 2004; Gertler and Levitte, 2005; Trippl *et al.*, 2009; Kang and Dall'erba, 2015). Feldman and Kogler (2010) and Moreno and Miguélez (2012) demonstrate that firms with limited access to distant knowledge pools tend to be less innovative and generate less output than their peers. For Maskell *et al.* (2006), it is the complementarity between the local knowledge pool and distant sources of knowledge that will promote regional innovation growth. Because each region has its own industry-mix and exploits local and distant knowledge pools differently (Feldman and Kogler 2010), one should expect the relative role of distant intra- and inter-sectoral knowledge spillovers on local innovation to vary across sectors also. To our knowledge, no previous study investigates this issue; hence this paper intends to fill this gap.

2.2. Empirical studies on regional and sectoral knowledge spillovers

Only a handful of studies focus on the differences in regional knowledge production across sectors. Using US state level data, the seminal work of Jaffe (1989) investigates the influence of university research on corporate patents across 4 different sectors. This study finds that the Drugs, Chemical and Mechanical sectors benefit from local intra-sectoral university research. Based on more detailed MSA data, Anselin *et al.* (2000) also investigate how local (within MSA) university research spills over to 4 industrial sectors and, unlike the previous study, highlight that interregional (beyond the MSA boundaries) spillovers of university research play a key role for some sectors. The authors find that the Drugs and Chemical sectors do not benefit from university research while the Electronic and Instruments sectors enjoy significant local spillovers. The Machinery sector is the only one that appears to benefit from global research spillovers.

One important element that is missing from the aforementioned studies is the presence of spillovers of private R&D, whether they take place within or across sectors, as they have been found to promote innovation as well (Wallsten, 2001; Orlando, 2004). Autant-Bernard and LeSage (2011) circumvent to this shortcoming by accounting for the capacity of both private and public intra- and interregional R&D spillovers to promote knowledge across 11 sectors of 94 French regions. Their conclusions indicate that Jacobian externalities dominate MAR externalities when they emanate from private R&D efforts. This result holds true whether they look at intra- or interregional spillovers. Local Jacobs and MAR externalities have roughly a similar role on innovation when they come from public R&D efforts while at the interregional level only the MAR externalities matter. No sectoral-level details are provided in Autant-Bernard and LeSage (2011) as the sectors are pooled together in their panel model. In addition, the spatial extent of the knowledge spillovers is modeled on geographical contiguity only so that neither the actual geographical extent nor the directionality of the knowledge flows are captured in their work. This paper remedies to these problems by tracking the various types (intra- vs interregional, intra-vs intersectoral, private vs. public R&D) of knowledge flows that exist. The relevant data and specific modeling strategy are described in the next section.

3. Empirical Model and Data

3.1. Regional Knowledge Production Function and Tobit model

Our empirical model relies on a regional approach of Griliches' (1979) knowledge production function using US county-level panel data. The knowledge production function is assumed to follow a Cobb-Douglas functional form as depicted in Equation (1) where y_{iht} is the knowledge output of sector h in county i at time t, $x_{k,iht}$ is the k$^{\text{th}}$ knowledge input, β_k is the elasticity of the output with respect to the corresponding input. α_{ih} and ε_{iht} represent the county-level unobservable heterogeneity of sector h and an error term respectively.

$$y_{iht} = \prod_k x_{k,iht}^{\beta_k} \cdot e^{\alpha_{ih}} \cdot e^{\varepsilon_{iht}} \qquad (1)$$

The logarithm transformation of Equation (1) leads to a log-linear model that is widely used in previous empirical studies of the knowledge production function (Anselin *et al.*, 1997; Acs *et al.*, 2002; Fischer and Varga, 2003; Bode, 2004). As usual in the literature, we use patent data as a proxy for knowledge output (Parent and LeSage, 2008; Autant-Bernard and LeSage, 2011; Kang and Dall'erba, 2015) and decide to work with patent applications (Cincera, 1997; Ramani *et al.*, 2008) instead of granted patents because the former are closer in time to knowledge creation. Since we focus on the knowledge created across 5 manufacturing sectors, we classify the patent applications based on the North American Industry Classification System (NAICS) defined in 2002 (Table 1).[1]

[Table 1]

Since the minimum value of several observed patent data is zero, we rely on a Tobit model for our empirical estimation (Autant-Bernard and LeSage, 2011). Equation (2) represents our empirical model and (*i*, *h*, *t*) are the indices of county *i*, sector *h*, and time *t*. $\boldsymbol{Patent}^{*}_{\boldsymbol{iht}}$ is the unobservable latent value of patent application and $\boldsymbol{Patent}_{\boldsymbol{iht}}$ is the observed patent application so that: $ln\boldsymbol{Patent}_{\boldsymbol{iht}} = \begin{cases} ln\boldsymbol{Patent}^{*}_{\boldsymbol{iht}} & if \;\; \boldsymbol{Patent}^{*}_{\boldsymbol{iht}} > 0 \\ 0 & if \;\; \boldsymbol{Patent}^{*}_{\boldsymbol{iht}} \leq 0 \end{cases}$.

As usual in a Tobit model, we assume the error term (ε_{iht}) follows a normal distribution. A similar distribution assumption applies to the county-level heterogeneity (α_{ih}). According to Kang and Dall'erba (2015), the metropolitan regions have a greater propensity to innovate and their knowledge production mechanism is different from that of the non-metropolitan regions.

Therefore, we focus on the metropolitan counties only. There were 853 of them across the 3,109 continental US counties in 2000.

The knowledge output at the county level is measured by patent application data from the US Patent and Trade Office (USPTO 2010). In order to allocate the patent data across counties, we use the fractional counting method suggested by Jaffe *et al.* (1993). When a patent is created by N inventors, it is assumed that 1/N fraction of the patent is attributed to each inventor. Each 1/N fractional patent is geocoded to its associated county based on the address of the inventor. As a result, the patent data is not an integer value but a rational number.[2]

$$
\begin{aligned}
&ln\boldsymbol{Patent_{iht}} = \\
&\beta_0 + \beta_1 ln\boldsymbol{Private_{iht}} + \beta_2 ln \sum_{\bar{h}\neq h}^{n} \boldsymbol{p_{ii\bar{h}h}}\, \boldsymbol{Private_{i\bar{h}t}} + \beta_3 \ln \boldsymbol{Univ_{iht}} + \ \beta_4 \ln \boldsymbol{Graduate_{iht-1}} \\
&+\beta_5 ln\boldsymbol{Emp_{iht-1}} + \beta_6 ln\boldsymbol{Large_{it-1}} + \beta_7 ln\boldsymbol{Diversity_{it-1}} + \beta_8\, ln\, \boldsymbol{Intra_{iht}} \\
&+\beta_9\, ln \sum_{i\neq j}^{N} \boldsymbol{p_{ijhh}}\, \boldsymbol{Private_{jht}} + \beta_{10}\, ln \sum_{\bar{h}\neq h}^{n} \sum_{i\neq j}^{N} \boldsymbol{p_{ij\bar{h}h}}\, \boldsymbol{Private_{j\bar{h}t}} + \beta_{11}\, ln \sum_{i\neq j}^{N} \boldsymbol{p_{ijhh}}\, \boldsymbol{Univ_{jht}} \\
&+\beta_{12}\, ln \sum_{i\neq j}^{N} \mathbf{P}_{\boldsymbol{ijhh}}\, \boldsymbol{Private_{jht}} + \beta_{13}\, ln \sum_{\bar{h}\neq h}^{n} \sum_{i\neq j}^{N} \boldsymbol{P_{ij\bar{h}h}}\, \boldsymbol{Private_{j\bar{h}t}} + \beta_{14}\, ln \sum_{i\neq j}^{N} \mathbf{P}_{\boldsymbol{ijhh}}\, \boldsymbol{Univ_{jht}} \\
&\qquad + \alpha_{ih} \\
&+\varepsilon_{iht}\ where\ \alpha_{ih} \sim N(0, \sigma_\alpha^2), \varepsilon_{iht} \sim N(0, \sigma_\varepsilon^2)\ and\ i = 1, \cdots, 853;\ h = 1, \ldots, 5; t = 2001, \ldots, 2008
\end{aligned}
\tag{2}
$$

The stock of knowledge is a main factor of the knowledge production function (Griliches, 1979). Here, the county-level knowledge stock is modeled through lagged expenditures in Research and Development (R&D) using the perpetual inventory method (Equation 3) as in

Mancusi (2008). In the equation, S_{iht} and RD_{iht} represent the stock of knowledge and the R&D expenditure in county *i*, sector *h* at time *t*. All R&D expenditures are converted in constant 2008 US dollars using each sector's Producer Price Index from the US Bureau of Labor Statistics.[3] We assume a 15% depreciation rate (δ) following Okubo *et al.*, (2006) and Mancusi (2008). In order to calculate the knowledge stock of the initial year, we approximate the industry specific growth rate of R&D expenditures (g) by the average of the annual growth rate over 1990-1999 across the US continental counties. This approach is used for each individual sector as in Mancusi (2008).

$$S_{iht} = (1-\delta) \cdot S_{iht-1} + RD_{iht-1} \;\; and \;\; S_{ih1990} = \left(\frac{RD_{ih1990}}{\delta+g}\right) \qquad (3)$$

We model two types of regional knowledge stocks of private and academic R&D. The private knowledge stock ($\boldsymbol{Private_{iht}}$) is approximated by the R&D expenditures of private companies collected from Standard and Poor's COMPUSTAT (Standard & Poor's, 2011). We use the address of these companies and their NAICS codes to allocate the R&D expenditures across counties and sectors. The regional academic knowledge stock ($\boldsymbol{Univ_{iht}}$) is measured by the amount of R&D spent across universities and colleges according to the National Science Foundation's Survey of R&D expenditures (National Center for Science and Engineering Statistics, 2013). In order to match this type of expenditure to a specific county and industrial sector, we use the address of the relevant institutions and the academic fields recipient of the R&D expenditures. The matching between academic fields and industrial sectors is based on Feldman (1994, p.58) and is reported in Table 2. Since one academic field can contribute to several sectors, we sum all academic R&D expenditures relevant to a sector under study for making the regional academic knowledge stock. For example, if we assume that there is an academic R&D input in Electrical

Engineering, then it contributes the exact amount to all three sectors of Mechanical, Computer, and Electrical equally. Anselin et al. (2000) follow the same process. If $\widehat{\beta_1}$ and $\widehat{\beta_3}$ are significant and positive, they will provide evidence that intra-regional MAR externalities induced by private and university R&D respectively are taking place within sector *h*.

[Table 2]

It is well known that human capital plays an important role in knowledge creation (Audretsch and Feldman, 2004). In order to measure the level of human capital available by county and industrial sector ($\boldsymbol{Graduate_{iht-1}}$), we use the total number of Graduate (Master's and Doctoral) or professional degree holders who are 25 years and over. The use of a one year lag is common in the knowledge production literature to alleviate any possible endogeneity problem (Ponds *et al.* 2010; Nesta and Saviotti 2005). The data comes from the 2000-2007 Integrated Public Use Microdata Series (IPUMS) developed by Ruggles *et al.* (2010). Since IPUMS classifies the occupation of the workers according to NAICS, we can easily allocate the number of degree holders by sector. IPUMS is surveyed based on the Public Use Microdata Area (PUMA), thus we match the location of the PUMA with that of the counties based on their 2000 US Census boundaries.[4]

In addition, we control for several region-specific conditions. Regional differences in the economic size of each sector are captured by the total number of employees in each sector ($\boldsymbol{Emp_{iht-1}}$). This variable is constructed based on the same method and data as the human capital variable. We also control for the share of large firms in a regional economy ($\boldsymbol{Large_{it-1}}$). Since small firms capitalize better than large firms on the knowledge created in university laboratories according to Acs *et al*. (1994), more small firms in a region can be conducive to regional knowledge creation. On the other hand, as large firms contribute to the higher level of agglomeration in a local economy (Acs and Armington 2004), their presence could be more beneficial to regional knowledge creation. In order to examine the relative role of small or large firms on regional knowledge creation, we include the share of establishments with at least 500 employees in our model. This cut-off is used by the 2000?? County Business Patterns to define small businesses and it has been used for similar purposes by Acs and Audretsch (1988), Anselin et al. (1997).

The degree of industrial diversity is also included to control for the general economic structure of each locality ($\boldsymbol{Diversity_{it-1}}$). It is measured through the index developed by Duranton and Puga (2000). The presence of this variable is necessary to capture the net effect of MAR vs. Jacobs externalities on innovation as more diverse places appear, by definition, to benefit more from the latter type. The calculation of this variable is reported in Equation (4) where s_{iht} is the share of industry h in county *i's* employment at time *t*, and s_{ht} the share of industry h in employment at the national level. The number of employees is measured across 13 industries over 2000-2007[5]. Obviously, this variable changes over time and space but not by sector.

$$Diversity_{it} = 1/\sum_h |s_{iht} - s_{ht}| \quad (4)$$

3.2. Regional and Sectoral Knowledge Spillovers

In addition to the above intra-regional and intra-sectoral characteristics, we account for 1) *intra-regional and inter-sectoral spillovers*, 2) *interregional and intra-sectoral spillovers*, 3) *interregional and inter-sectoral spillovers*. Intra-regional and intra-sectoral spillovers are already accounted for in $\boldsymbol{Private}$ and $\boldsymbol{Univ}$ since they capture expenses within the same county and sector as the dependent variable. Moreover, note that the intersectoral knowledge spillovers emanate from private R&D expenses only. Those deriving from university R&D spending are not included since one academic field can contribute to innovation across several sectors.

The three types of spillovers above are modeled based on NBER US Patent Citation Data File (Hall *et al*., 2001). Since this data allow us to track the patent creation-citation flows between all 3,109 US continental counties as well as the industrial sector of both the cited and citing patents, we can construct 25 (5×5 sectors) technological network matrices across (3,109×3,109) counties. The fractional counting method is used here too so that we capture all 1/(O×D) knowledge flows between the number of inventors at the place origin O and their peers at the destination D for any pair of origin-destination sectors. This patent creation-patent citation matrix is noted $\boldsymbol{p_{ij}}$ (respectively $\boldsymbol{P_{ij}}$) to capture the expected amount of knowledge emanating from up to 50 miles (resp. from 50 miles on) from *i*. Patent citation data are available over 1975-1999 only. To the best of our knowledge, more recent data do not exist. Here, we use the 1990-1999 data only since our R&D expenditures are measured over 1990-2007. However, one could argue that the citation flows of the early 1990's are not necessarily correlated with the more recent citation flows that lead to knowledge creation in 2008, the final measurement of our dependent variable. As such, we use the patent citation flows over 1990-1999 to model the 2001-2004 spillovers of knowledge stocks and the patent citation flows over 1995-1999 for the 2005-2008 spillovers. In the first case, we decide to include the most recent patent citation flows from 1995-1999 since their effects are much greater than those during 1990-1999 due to the 15 percent depreciation rate. And that we drop the patent citation flow over 1990-1994 is supported by the same idea that old R&D obsolesces over time. Since we focus on knowledge creation in the metropolitan counties only, we use the (853×3,109) sub-matrix of $\boldsymbol{p_{ij}}$ (or $\boldsymbol{P_{ij}}$) to capture the knowledge flows from all 3109 counties to the 853 metropolitan counties.

This matrix is used to model first the *intra-regional and inter-sectoral spillovers* from private knowledge stock ($\boldsymbol{p_{ii\bar{h}}Private_{i\bar{h}t}}$) using Equation (5) below. This variable captures intra-regional Jacobian externalities as $\boldsymbol{p_{ii\bar{h}}}$ is the (i^{th}, i^{th}) element of $\boldsymbol{p_{ij\bar{h}}}$:

$$\boldsymbol{p_{ii\bar{h}}Private_{i\bar{h}t}} = \begin{cases} \sum_{\bar{h}\neq h}^{n} \boldsymbol{p_{ii\bar{h}h(1990-1999)}}\ \boldsymbol{Private_{i\bar{h}t}} \text{ for } t = 2001, \cdots, 2004 \\ \sum_{\bar{h}\neq h}^{n} \boldsymbol{p_{ii\bar{h}h(1995-1999)}}\ \boldsymbol{Private_{i\bar{h}t}} \text{ for } t = 2005, \cdots, 2008 \end{cases} \quad (5)$$

This matrix is normalized by its column sums to represent the frequency of the citation flows from sector $\bar{h}$ to h within the MSA county i.

The interregional and intra-sectoral spillovers (interregional MAR externalities) are modeled as in Equations (6) and (7). We distinguish the singular role of nearby (matrix $\boldsymbol{p_{ij}}$) and distant (matrix $\boldsymbol{P_{ij}}$) interregional knowledge spillovers. The former have a spatial extent limited to 50 miles as in Anselin *et al*. (1997) since it corresponds to the average daily US commuting distance (Smallen 2004; Rapino and Fields 2013). A robustness test will be performed for 75 miles too. Distant interregional spillovers (equation 7) correspond to those taking place from 50 miles to any further counties. Distances are based on the great circle distance between the centroids of counties *i* and *j*. Nearby and distant interregional spillovers of academic knowledge ($\boldsymbol{p_{ijhh}\ Univ_{jht}}$and $\boldsymbol{P_{ijhh}Univ_{jht}}$) are defined in the same way but with the amount of R&D spent across universities and colleges instead.

$$\boldsymbol{p_{ijhh}Private_{jht}} =$$

$$\begin{cases} \sum_{i\neq j}^{N} \boldsymbol{p_{ijh h(1990-1999)}}\ \boldsymbol{Private_{jht}} \cdot 1(d(i,j) \leq 50\ miles) \text{ for } t = 2001, \cdots, 2004 \\ \sum_{i\neq j}^{N} \boldsymbol{p_{ijhh(1995-1999)}}\ \boldsymbol{Private_{jht}} \cdot 1(d(i,j) \leq 50\ miles) \text{ for } t = 2005, \cdots, 2008 \end{cases} \quad (6)$$

$$\boldsymbol{P_{ijhh}Private_{jht}} =$$

$$\begin{cases} \sum_{i\neq j}^{N} \boldsymbol{P_{ijhh(1990-1999)}}\ \boldsymbol{Private_{jht}} \cdot 1(d(i,j) > 50\ miles)\ \text{ for } t = 2001, \cdots, 2004 \\ \sum_{i\neq j}^{N} \boldsymbol{P_{ijhh(1995-1999)}}\ \boldsymbol{Private_{jht}} \cdot 1(d(i,j) > 50\ miles)\ \text{ for } t = 2005, \cdots, 2008 \end{cases} \quad (7)$$

Interregional Jacobian externalities are built on the same model as equations (6) and (7) but they are captured through the normalized patent creation-citation flows matrix from the other 4 sectors to sector h $\left(\sum_{\bar{h}\neq h}^{n}\sum_{i\neq j}^{N} \boldsymbol{p_{ij\bar{h}h}}\right)$. Finally, we capture differences in the knowledge absorption capacity of each county by calculating the share of intraregional flows in the intrasectoral patent creation-citation flows across any county (i.e. $100\times\boldsymbol{p_{iihh}}$). We note this variable $Intra_{ih}$.

Tables 3 shows descriptive statistics of the aforementioned variables by sector. It is obvious that the Chemical and Drugs & Medical (hereafter Drugs) sectors generate relatively less patents than the other sectors while the high-tech sectors of Computer & Communication (hereafter Computer) and Electrical & Electronic (hereafter Electrical) display the largest mean value. Since most explanatory variables have a minimum value of zero, we add one to all of them before log transformation.

[Table 3]

4. Estimation Results

Table 4 shows the Maximum Likelihood estimation results of the panel Tobit models with a 50 mile distance cut-off. The results are robust to a 75 mile cut-off, which indicates that they are not sensitive to the definition of the spatial extent of daily commuting in the US. Models 1 to 5 report the results for each of our 5 sectors. They all indicate that the stock of private knowledge leads to significant and positive intra-regional MAR and Jacobian externalities on regional knowledge creation. However, the former displays a greater elasticity than the latter at the 5% significance level (one-tailed test) for all sectors. It suggests that regional specialization leads to more innovation than otherwise (Henderson, 2003), a result that is confirmed by the negative elasticity of diversity among the significant results. As expected, we also find a significant role of spending in higher education within the sector and the county of interest. Its elasticity is not statistically different from that of the intra MAR spillovers of private knowledge at the 5% level for 3 sectors and it is even greater in the Chemical and Drugs sectors. Meyer-Krahmer and Schmoch (1998) suggest that it is because basic research is very important in the Chemical sector and Anselin *et al*. (2000) suggest that it is because of the time lag between basic research and commercialization in the pharmaceutical industry.

Everything else held constant, a stronger share of intra-county citation (***Intra***) leads to greater knowledge creation for all sectors. This result confirms our expectations as the importance of face-to-face contacts in the transmission of knowledge has already been well documented (Jaffe, 1986; Jaffe *et al*., 1993; Audretsch and Feldman, 1996; Rodríguez-Pose, 2001; Sonn and Storper, 2008). We also find a significant positive role of the number of graduate degree holders and employees in a sector on the patenting activity of the same sector. These results confirm the economies of scale that can be achieved with spatial agglomeration in conjunction with specialization and MAR externalities as found earlier. While the elasticity of human capital ranges from 0.039% to 0.048% across all sectors, the role of employment is particularly acute in the Mechanical sector compared to the other sectors. Finally, we find that the greater is the presence of large establishments the more knowledge is created in any sector. Acs and Armington (2004) indicate that large firms lead to a greater local labor pool which contributes to agglomeration economies and, as seen earlier, innovation.

[Table 4]

The results related to all types of interregional spillovers appear in the middle part of table 4. The Chemical and Drugs sectors (Models 1 and 2) show that interregional (above and below 50 miles) private and academic MAR spillovers foster innovation. This result indicates that geographical proximity is not a requirement to transfer basic research knowledge in these sectors (Mansfield 1995; Meyer-Krahmer and Schmoch 1998) and, as demonstrated by Gittelman (2007), that the collaboration network of the US biotechnology industry is spread geographically. Indeed, even if the elasticity of the spillovers emanating from private R&D investments diminishes with increasing distance, the difference is not significant at the 5 percent level. Interestingly, we do not observe the same linear distance decay with respects to the role of academic spillovers. It is obviously high within the county where innovation is measured, but then long-distance academic spillovers seem more important than the short-distance ones (below 50 miles).

The results for the Mechanical sector are reported in Model 3. While the sign and significance level of the intra-regional factors are similar to what they are for the previous two sectors, some differences on the interregional spillovers appear. In this sector, the interregional MAR spillovers of private knowledge are significant within 50 miles only, which is in tune with Anselin *et al.* (2000). On the other hand, interregional Jacobian spillovers of private knowledge and interregional spillovers of academic knowledge are effective over 50 miles only. The latter result confirms both Jaffe (1989) and Anselin *et al*. (2000) who find no evidence of localized spillovers of university research in the Machinery sector using state and MSA level data respectively.

Our sectoral analysis focuses on the Computer and Electrical sectors in columns 4 and 5 respectively. While the role of the intra-regional characteristics – including spillovers – is similar to what other sectors display, the various types of interregional spillovers have a unique contribution on sectoral innovation in these two sectors. Indeed, both sectors experience a significant role of localized and distant MAR interregional spillovers of private and academic knowledge. Anselin *et al.* (2000) find significant local university research spillovers in similar sectors (Electronics and Instruments sectors). On the other hand, interregional Jacobian spillovers of private knowledge do not play any significant role in either sector at the 5% level. It seems that, in these sectors, geographical proximity is a critical aspect to benefit from Jacobian spillovers. On the other hand, we find in each of these two sectors that short- and long-distance spillovers of private R&D spending have a significant role on local knowledge creation; yet, their difference is not statistically different at 5%. The long-distance knowledge pipelines these sectors rely on to innovate is just as important as those built on interactions with close neighbors.

5. Conclusion

The regional knowledge production literature has given an increasing amount of attention to the role of spatial spillovers on knowledge creation. However, the bulk of empirical studies rely on an aggregated approach that masks the differences in the importance of the various types of MAR and Jacobian externalities on knowledge creation across sectors. The few exceptions (e.g. Jaffe, 1989; Anselin *et al*., 2000) have highlighted the presence of sectoral heterogeneity in the size of the localized knowledge spillovers emanating from university research. However, they have missed the opportunity to also investigate how inter-sectoral and distant interregional knowledge spillovers matter. This paper corrects these shortcomings by examining the heterogeneous role of intra- and interregional as well as intra- and inter-sectoral knowledge spillovers across 5 US manufacturing sectors that cover 82% of the patents recorded in USPTO.

Our estimation results show that both intra-sectoral (MAR) and inter-sectoral (Jacobian) spillovers are important determinants of knowledge production. But generally intra-regional MAR spillovers display a larger return than the corresponding Jacobian spillovers. Moreover, we find that the elasticity of academic knowledge spillovers is as great as the one of private knowledge across 3 sectors and it is even greater in the Chemical and Drugs sectors. We also discover that in the Mechanical, Computer and Electrical sectors, intra-regional MAR spillovers play a greater role than the interregional spillovers emanating from nearby counties. This result implies that frequent face-to-face contacts are still an important factor in knowledge spillover creation. However, we find no statistical evidence of a difference in the role of short-vs. long-distance MAR spillovers of private knowledge except in the Chemical and Mechanical sectors. This result implies that once the distance between intellectual partners is greater than a daily commuting extent, physical interactions due to proximity can easily be substituted for other forms of contacts. Ponds et al. (2010) find that the impact of academic research on the industry knowledge production stemming from the university-industry collaboration networks is not limited by geographical proximity.

Our estimation results suggest three implications for policy-makers interested in more efficient innovations strategies. First, more attention to the sectoral heterogeneity is needed as each economic sector relies on knowledge spillovers that have their own specific source and spatial structure. Recommendations based on the traditional aggregated approach mask this heterogeneity. Second, we have also find that specialization is more innovation-prone than diversity. It is reflected in the greater role of intra- and interregional MAR spillovers of private knowledge as well as in the positive impact of the presence of graduate degree holders and employees in the sector of interest. While specialization was already brought to the fore in Jaffe (1986) and Henderson (2003), their work did not include the necessary presence of interregional knowledge externalities. Last but not least, promoting academic knowledge and its distant interregional spillovers is a strategy that is beneficial for innovation (Ponds *et al.*, 2010). We have shown that it is relevant across all the sectors we have explored here and that the role of spillovers does not decrease with distance. Therefore, policy makers need to facilitate academic R&D and help build long-distance networks of university-industry collaboration.

[1] This classification is not directly compatible with the industrial classification of the patent citation data of Hall *et al*. (2001) because the latter are classified based on the US Patent Classification Systems (USPCS). As a result, we rely on the USPCS-NAICS 2002 concordance file developed by the US Patent and Trademark Office (USPTO) to transfer all the data in NAICS format.

[2] The geographical allocation of the patent data could also be based on the address of the assignee(s). However, large companies use the address of their headquarter to file patents (Fischer and Varga, 2006), which is not always the place where research took place. We are aware that using the inventor's address to geocode the creation and citation of patent data can lead to a similar problem (Autant-Bernard and LeSage, 2011), but the size of the error is smaller as we assume that inventors live close to their place of work.

[3] The average of the Producer Price Index (PPI) of the NAICS sectors reported in Table 1 is used to calculate the annual PPI of each of our five manufacturing sectors.

[4] If a PUMA area consists of more than one county, we allocate the number of degree holders of each sector proportionally to the counties' total number of degree holders (for all industrial sectors) of which data come from the 2000 US Census. This approach is used for the years 2000, 2005, 2006, 2007. For the years 2001-2004, IPUMS provides personal information at the state level only. As a result, we first calculate the total sum of degree holders by state and by sector and then distribute it across counties proportionally to their average number of degree holders over the years 2000 and 2005.

[5] The 13 industries are based on the 2000 US Census classification: 1) Agriculture, forestry, fishing and hunting, and mining, 2) Construction, 3) Manufacturing, 4) Wholesale trade, 5) Retail trade, 6) Transportation and warehousing, and utilities, 7) Information, 8) Finance, insurance, real estate, and rental and leasing, 9) Professional, scientific, management, administrative, and waste management services, 10) Educational, health and social services, 11) Arts, entertainment, recreation, accommodation and food services, 12) Other services (except public administration, 13) Public administration or Industries not classified. We use the 2000 US Census for the diversity index of 2000 and the County Business Patterns for the index over 2003-2007. For the year 2002, the Census Bureau does not provide the number of employees for several sectors. We fill these data with the corresponding values from the 2003 County Business Patterns data and use an average of 2000 and 2002 as a proxy for 2001.

Table 1 Classification of Industrial sectors

Sector	NAICS 2002	Description
Chemical	325	Chemicals
	3251	Basic Chemicals
	3252	Resin, Synthetic Rubber, and Artificial and Synthetic Fibers and Filaments
	3253, 3255, 3256, 3259	Other Chemical Product and Preparation
Drugs & Medical	3254	Pharmaceutical and Medicines
	3391	Medical Equipment and Supplies
Mechanical	333	Machinery
	336	Transportation Equipment
	3361, 3162, 3363	Motor Vehicles, Trailers and Parts
	3364	Aerospace Product and Parts
	3365, 3366, 3369	Other Transportation Equipment
Computer & Communication	334	Computer and Electronic Products
	3341	Computer and Peripheral Equipment
	3342	Communications Equipment
Electrical & Electronic	3344	Semiconductors and Other Electronic Components
	3345	Navigational, Measuring, Electromedical, and Control Instruments
	3343, 3346	Other Computer and Electronic Products
	335	Electrical Equipment, Appliances, and Components

Note: Classification is based on the 4 digit codes of NAICS 2002. When 4 digit codes are not available for patents and private R&D expenditures, we use their 3 digit codes. This classification is defined to make the patent data compatible with the patent citation data of Hall et al. (2001).

Table 2 Matching Academic Fields with the Manufacturing Sectors

Sector	Field of Science and Engineering
Chemical	Engineering-Chemical eng. Physical sciences-Chemistry
Drugs & Medical	Engineering-Bioengineering/ biomedical eng. Life sciences-Biological sciences Life sciences-Medical sciences
Mechanical	Engineering-Aeronautical & astronautical eng. Engineering-Civil eng. Engineering-Electrical eng. Engineering-Mechanical eng. Engineering-Metallurgical & materials eng. Engineering-Other Physical sciences-Astronomy Physical sciences-Physics Physical sciences-Other Mathematical sciences
Computer & Communication	Engineering-Aeronautical & astronautical eng. Engineering-Electrical eng. Physical sciences-Astronomy Physical science-Physics Mathematical sciences Computer sciences
Electrical & Electronic	Engineering-Aeronautical & astronautical eng. Engineering-Electrical eng. Physical sciences-Astronomy Physical sciences-Physics Mathematical sciences Computer sciences

Table 3 Descriptive Statistics: Five manufacturing sectors (NT=6,824)

	Chemical		Drugs & Medical		Mechanical		Computer & Communication		Electrical & Electronic	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Patent_{h}	12.7	39.1	12.5	44.2	18.1	46.8	33.7	162.7	42.4	185.4
Private_{h}	98135	949415.7	304737.5	2653715.1	247268.7	3646690.3	82349.5	1113119.8	136660.2	1756760.1
$p_{ii\bar{h}h}\text{Private}_{\bar{h}}$	161834.7	1672363.9	50301.2	661173.7	126178.3	1576988.5	113087.8	1253533.5	88088.7	1003104.4
Univ_{h}	13787.4	52115.4	137802.2	621499	48804	220857.9	16169.6	83064.7	22575.4	116220.9
Graduate_{h}	91.4	255.9	177.9	556.1	306	1257.1	131.2	674.3	235.7	1139.9
Emp_{h}	742	1637.7	914.2	2128.9	3253.4	7691.3	595.5	2089.7	1548.5	4662.8
Intra_{h}	13.3	18.9	7.1	13.4	10.8	15.5	4.8	11.3	8.4	14
Large	14.2	4.4	14.2	4.4	14.2	4.4	14.2	4.4	14.2	4.4
Diversity	3.4	1.5	3.4	1.5	3.4	1.5	3.4	1.5	3.4	1.5
$p_{ij(50)hh}\text{Private}_{h}$	10046.7	78051.2	37030.8	222613.5	44785.3	611634.8	7097.9	75391.7	9393.3	70431.2
$p_{ij(50)\bar{h}h}\text{Private}_{\bar{h}}$	74313	431528.4	35632.2	284128	57333	568899.8	73029.6	726062.3	59228.7	496647.9
$p_{ij(50)hh}\text{Univ}_{h}$	1049.1	4369.7	9567.1	71615.7	3554.6	19277.4	982.4	10498.3	1590.6	15244.2
$p_{ij(75)hh}\text{Private}_{h}$	11492.1	80753.4	44158.4	240621.2	48165.4	617403.4	7966.2	76794.7	10563.5	71560.6
$p_{ij(75)\bar{h}h}\text{Private}_{\bar{h}}$	92671.5	478177.8	43778.3	309721.7	69890.3	650826.3	86461.9	761632	73674.2	602713.6
$p_{ij(75)hh}Univ_{h}$	1241.3	4670.3	11249.3	74110.4	4082.6	19980.5	1109.1	10868.2	1874.4	15899
$P_{ij(50)hh}\text{Private}_{h}$	50212.1	121558.7	216210.2	732990.8	130098	311783.4	58544.4	241057.5	89425.3	388009.9
$P_{ij(50)\bar{h}h}\text{Private}_{\bar{h}}$	499218.7	1724010	444909.1	2584978.8	383092.7	1544913.5	540082.1	2415983.5	523612.4	1940453.1
$P_{ij(50)hh}\text{Univ}_{h}$	9586.5	21847.2	108833.1	366316.7	35115.6	99157.3	14325.3	75677.5	18830.4	73944.7
$P_{ij(75)hh}\text{Private}_{h}$	48766.7	118812.2	209082.5	718665.4	126717.8	309031	57676.2	240018.9	88255.1	386432.6
$P_{ij(75)\bar{h}h}\text{Private}_{\bar{h}}$	480860.2	1672180	436763	2578045.4	370535.4	1455878.7	526649.8	2402668.2	509166.9	1903774.9
$P_{ij(75)hh}\text{Univ}_{h}$	9394.3	21518.6	107150.8	363527.5	34587.7	98288.8	14198.6	75287.6	18546.6	73630.2

Table 4 Panel Tobit model with 50 miles distance cut-off

Dep.: ln Patent	Model 1 Chemical			Model 2 Drugs & Medical			Model 3 Mechanical			Model 4 Computer & Communication			Model 5 Electrical & Electronic		
	Marginal Effects	(S. E.)		Marginal Effects	(S. E.)		Marginal Effects	(S. E.)		Marginal Effects	(S. E.)		Marginal Effects	(S. E.)	
ln Private$_h$	0.068	(0.011)	**	0.055	(0.010)	**	0.075	(0.009)	**	0.108	(0.012)	**	0.102	(0.009)	**
ln $p_{ii\bar{h}h}$Private$_{\bar{h}}$	0.038	(0.009)	**	0.027	(0.010)	**	0.024	(0.008)	**	0.021	(0.009)	*	0.030	(0.007)	**
ln Univ$_h$	0.122	(0.009)	**	0.082	(0.008)	**	0.068	(0.007)	**	0.107	(0.009)	**	0.100	(0.008)	**
ln Graduate$_h$	0.048	(0.010)	**	0.044	(0.011)	**	0.039	(0.009)	**	0.040	(0.011)	**	0.040	(0.008)	**
ln Intra$_h$	0.110	(0.019)	**	0.118	(0.023)	**	0.067	(0.016)	**	0.055	(0.026)	*	0.089	(0.018)	**
ln Emp$_h$	0.051	(0.016)	**	0.076	(0.015)	**	0.317	(0.025)	**	0.026	(0.010)	**	0.073	(0.015)	**
ln Large	0.459	(0.116)	**	0.669	(0.113)	**	0.259	(0.100)	**	0.566	(0.110)	**	0.440	(0.097)	**
ln Diversity	-0.213	(0.073)	**	-0.061	(0.071)		-0.123	(0.060)	*	0.028	(0.068)		-0.043	(0.056)	
ln $p_{ij(50)hh}$Private$_h$	0.044	(0.011)	**	0.048	(0.012)	**	0.021	(0.009)	*	0.060	(0.015)	**	0.048	(0.008)	**
ln $p_{ij(50)\bar{h}h}$Private$_{\bar{h}}$	0.031	(0.008)	**	0.018	(0.008)	*	0.009	(0.006)		0.012	(0.008)		0.007	(0.006)	
ln $p_{ij(50)hh}$Univ$_h$	0.033	(0.011)	**	0.029	(0.011)	**	0.016	(0.009)		0.041	(0.014)	**	0.028	(0.010)	**
ln $P_{ij(50)hh}$Private$_h$	0.019	(0.010)	*	0.032	(0.011)	**	0.000	(0.007)		0.066	(0.013)	**	0.057	(0.006)	**
ln $P_{ij(50)\bar{h}h}$Private$_{\bar{h}}$	0.020	(0.005)	**	0.003	(0.005)		0.018	(0.004)	**	0.002	(0.005)		0.008	(0.005)	
ln $P_{ij(50)hh}$Univ$_h$	0.043	(0.012)	**	0.035	(0.012)	**	0.031	(0.008)	**	0.070	(0.015)	**	0.017	(0.006)	**
	Estimate	(S. E.)		Estimate	(S. E.)		Estimate	(S. E.)		Estimate	(S. E.)		Estimate	(S. E.)	
Intercept	-4.069	(0.439)	**	-5.565	(0.467)	**	-2.797	(0.296)	**	-3.806	(0.391)	**	-2.099	(0.310)	**
σ_α	1.497	(0.049)	**	1.538	(0.053)	**	1.008	(0.032)	**	1.356	(0.045)	**	1.141	(0.037)	**
σ_ε	1.320	(0.015)	**	1.358	(0.016)	**	0.866	(0.009)	**	1.147	(0.013)	**	0.826	(0.009)	**
Log likelihood	-10741.8			-10384.7			-9603.2			-10365.4			-9214.2		
Total counties	853			853			853			853			853		
Time periods	8			8			8			8			8		
Total observation	6,824			6,824			6,824			6,824			6,824		
Left-censored obs.	1,820			2,113			688			1,529			955		

Note: * P-value< 5%, ** P-value< 1%. Standard errors are calculated by the delta method.

References

Acs Z, Armington C (2004) Employment growth and entrepreneurial activity in cities. *Regional Studies* 38**:** 911-927

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

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

Almeida P, Kogut B (1999) Localization of knowledge and the mobility of engineers in regional networks. *Management Science* 45**:** 905-917

Anselin L, Varga A, Acs Z (1997) Local geographic spillovers between university research and high technology innovations. *Journal of Urban Economics* 42**:** 422-448

Anselin L, Varga A, Acs Z (2000) Geographical spillovers and university research: A spatial econometric perspective. *Growth and Change* 31**:** 501-515

Arrow KJ (1962) The economic implications of learning by doing. *The Review of Economic Studies* 29**:** 155-173

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

Audretsch DB, Feldman MP (2004) Chapter 61 knowledge spillovers and the geography of innovation. In: Henderson J.V., Thisse J.-F. (Eds) *Handbook of regional and urban economics*. Elsevier, San Diego, CA

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

Autant-Bernard C (2012) Spatial econometrics of innovation: Recent contributions and research perspectives. *Spatial Economic Analysis* 7**:** 403-419

Autant-Bernard C, LeSage JP (2011) Quantifying knowledge spillovers using spatial econometric models. *Journal of Regional Science* 51**:** 471-496

Bode E (2004) The spatial pattern of localized R&D spillovers: An empirical investigation for germany. *Journal of Economic Geography* 4**:** 43-64

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

Crescenzi R, Nathan M, Rodríguez-Pose A (2016) Do inventors talk to strangers? On proximity and collaborative knowledge creation. *Research Policy* 45 (1):177-194.

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

Duranton G, Puga D (2001) Nursery cities: Urban diversity, process innovation, and the life cycle of products. *The American Economic Review* 91**:** 1454-1477

Feldman MP (1994) *The geography of innovation* (Economics of science, technology and innovation, No. 2). Springer Netherlands, Dordrecht

Feldman MP, Kogler DF (2010) Chapter 8 - stylized facts in the geography of innovation. In: Bronwyn HH, Nathan R (eds) *Handbook of the economics of innovation*. North-Holland,

Fischer MM, Scherngell T, Jansenberger E (2006) The geography of knowledge spillovers between high-technology firms in europe: Evidence from a spatial interaction modeling perspective. *Geographical Analysis* 38**:** 288-309

Fischer MM, Varga A (2003) Spatial knowledge spillovers and university research: Evidence from austria. *The Annals of Regional Science* 37**:** 303-322

Fischer MM, Varga A (2006) 10 spatial knowledge spillovers and university research: Evidence from austria. In: Fischer MM (ed) *Innovation, network, and knowledge spillovers: Selected essays*. Springer, New York

Gertler MS, Levitte YM (2005) Local nodes in global networks: The geography of knowledge flows in biotechnology innovation. *Industry and Innovation* 12**:** 487-507

Gittelman M (2007) Does geography matter for science-based firms? Epistemic communities and the geography of research and patenting in biotechnology. *Organization Science* 18**:** 724-741

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

Grossman GM, Helpman E (1994) Endogenous innovation in the theory of growth. *The Journal of Economic Perspectives* 8**:** 23-44

Hall BH, Jaffe AB, Trajtenberg M (2001) The nber patent citations data file: Lessons, insights and methodological tools. *NBER Woring Paper Series 8498*. National Bureau of Economic Research, Cambridge, MA

Henderson JV (2003) Marshall's scale economies. *Journal of Urban Economics* 53**:** 1-28

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

Jaffe AB (1986) Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value. *The American Economic Review* 76**:** 984-1001

Jaffe AB (1989) Real effects of academic research. *The American Economic Review* 79**:** 957-970

Jaffe AB, Trajtenberg M, Henderson R (1993) Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics* 108**:** 577-598

Kang D, Dall’erba S (2015) An examination of the role of local and distant knowledge spillovers on the us regional knowledge creation. *International Regional Science Review*, 39(4), 355-385.

Maggioni MA, Uberti TE, Usai S (2010) Treating patents as relational data: Knowledge transfers and spillovers across italian provinces. *Industry and Innovation* 18**:** 39-67

Mancusi ML (2008) International spillovers and absorptive capacity: A cross-country cross-sector analysis based on patents and citations. *Journal of International Economics* 76**:** 155-165

Mansfield E (1995) Academic research underlying industrial innovations: Sources, characteristics, and financing. *The Review of Economics and Statistics* 77**:** 55-65

Marshall A (1920) *Principles of economics*. Macmillan, London

Maskell P, Bathelt H, Malmberg A (2006) Building global knowledge pipelines: The role of temporary clusters. *European Planning Studies* 14**:** 997-1013

Meyer-Krahmer F, Schmoch U (1998) Science-based technologies: University–industry interactions in four fields. *Research Policy* 27**:** 835-851

Moreno R, Miguélez E (2012) A relational approach to the gegogrpahy of innovation: A typology of regions. *Journal of Economic Surveys* 26**:** 492-516

National Center for Science and Engineering Statistics (2013) Guide for public use data files national science foundation’s higher education research and development survey: Fiscal year 2011. *Yearly*. NSF National Center for Science and Engineering Statistics, Arlington, Virginia

Nesta, L., & Saviotti, P. P. (2005). Coherence of the knowledge base and the firm's innovative performance: evidence from the US pharmaceutical industry. *The Journal of Industrial Economics*, *53*(1), 123-142.

Okubo S, Robbins CA, Moylan CE, Sliker BK, Schultz LI, Mataloni LS (2006) R&D satellite account: Preliminary estimates. Bureau of Economic Analysis, US Department of Commerce, Washington, DC

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

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

Parent O, LeSage JP (2008) Using the variance structure of the conditional autoregressive spatial specification to model knowledge spillovers. *Journal of Applied Econometrics* 23**:** 235-256

Peri G (2005) Determinants of knowledge flows and their effect on innovation. *The Review of Economics and Statistics* 87**:** 308-322

Ponds R, Oort Fv, Frenken K (2010) Innovation, spillovers and university–industry collaboration: An extended knowledge production function approach. *Journal of Economic Geography* 10**:** 231-255

Rapino MA, Fields AK (2013) Mega commuters in the U.S.: Time and distance in defining the long commute using the american community survey. *Working Paper 2013-03*. United States Census Bureau,

Rey SJ (2000) Integrated regional econometric+input-output modeling: Issues and opportunities. *Papers in Regional Science* 79**:** 271-292

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

Ruggles S, Alexander JT, Genadek K, Goeken R, Schroeder MB, Sobek M (2010) Integrated public use microdata series: Version 5.0 [machine-readable database]. Minneapolis: University of Minnesota

Smallen D (2004) 3.3 million americans are stretch commuters traveling at least 50 miles one-way to work. United States Department of Transportation. https://www.rita.dot.gov/bts/sites/default/files/rita_archives/bts_press_releases/2004/bts010_04/html/bts010_04.html.

Sonn JW, Storper M (2008) The increasing importance of geographical proximity in knowledge production: An analysis of us patent citations, 1975 - 1997. *Environment and Planning A* 40**:** 1020-1039

Sorenson O, Rivkin JW, Fleming L (2006) Complexity, networks and knowledge flow. *Research Policy* 35**:** 994-1017

Standard & Poor's (2011) Standard & Poor's Compustat® Xpressfeed understanding the data. The McGraw-Hill Companies, Inc,

Trippl M, Tödtling F, Lengauer L (2009) Knowledge sourcing beyond buzz and pipelines: Evidence from the vienna software sector. *Economic Geography* 85**:** 443-462

USPTO (2010) Patents bib: Selected bibliographic information from us patents issued 1969 to present. In: Office tUPT (ed). Alexandria, VA

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