[Preliminary version. Please, do not cite without authors' permission.]

How Important is Closeness for Knowledge Flows?[☆]

Kyriakos Drivas[a], Claire Economidou[a,*], Sotiris Karkalakos[a], Mike Tsionas[b]

[a] *Department of Economics, University of Piraeus, Piraeus 184 35, Greece*
[b] *Department of Economics, Athens University of Economics and Business, Athens 104 34, Greece*

Abstract

This paper aims to provide novel insights in the diffusion of knowledge and its consequences for innovation activity. We use a single framework to learn more about the relative scope and intensity of four different channels of knowledge flows that operate via patent citations, inventors' mobility, trade of patents, and trade of goods, which have been analyzed separately from different avenues of the knowledge spillover literature. To jointly study these flows, we propose econometric techniques appropriate to the nature of the data. Using recently developed and detailed data for the states of the US, our findings support that geographic proximity, in terms of distance and border, matters for the spread of knowledge and is more essential to the operation of market-based (patents, inventors, goods) than to the operation of non-market (citations) channels of knowledge flows. The geographic scope of disembodied knowledge flows based on traded patents or citations is far larger than that of embodied knowledge in inventors or goods. Other types of similarities such as technological effort of states and technological proximity of production sectors further shape knowledge diffusion. Finally, knowledge flows, especially disembodied flows, are relevant to a state's innovation production as external accessible R&D gained through these flows has a strong positive effect on a state's innovation activity, as large, for some cases, as that of state's own R&D stock.

Keywords: patents, citations, inventor mobility, trade, knowledge flows, non-linear regression systems
JEL: C11, C33, O31, O51

[☆] We are grateful to Stuart Graham, Alan Marco, Kirsten Apple, Saurabh Vishnubhakat, Galen Hancock and the entire staff of the Office of the Chief Economist at the United States Patent and Trademark Office (USPTO) for their wise assistance and generous support. We thank Adamandia Karaiskaki and Cathrin Linardaki for excellent research assistance. We also thank Jaap Bos, George Delta, Grigoris Emvalomatis, Mark Sanders, Vassilis Sarafidis, Lu Zhang and participants of the ZEW conference on Innovation and Patenting for useful insights. Kyriakos Drivas gratefully acknowledges financial support from the *National Strategic Reference Framework* No: SH1_4083. The usual disclaimer applies.

[*] Corresponding author.

Email addresses: `dribask@unipi.gr` (Kyriakos Drivas), `economidou@unipi.gr` (Claire Economidou), `sotkar@unipi.gr` (Sotiris Karkalakos), `tsionas@aueb.gr` (Mike Tsionas)

1. Introduction

Generation of ideas lies at the center of the economic growth theory (Romer, 1986; Jones, 2005). Generated new technological knowledge that spills over through positive learning externalities comprises a potential source of economic growth for a country behind the technological frontier, therefore contributing to convergence across countries.

The growth-enhancing effects of innovation and its diffusion have been extensively studied in the literature.[1] Branches of the so-called spillover literature have progressed on separate avenues in their analysis of knowledge flows and contributions differ widely on the way knowledge flows are inferred and on the aggregation level. One research avenue, initiated by the seminal work of Jaffe et al. (1993), assess the direction and intensity of knowledge flows by analyzing patent citations flows between entities. The principal assumption driving this research is that citations trace out knowledge flows and technological learning: a citation form a patent to another patent indicates that inventors on the latter patent knew and used the former.[2] An important outpour of research, the patent-citation literature (Branstetter, 2001; Peri, 2005; Mancusi, 2008) performs cross-country analyses of non-market, disembodied knowledge flows based on patent citation flows.[3]

A separate avenue of the spillover literature, the trade-growth literature, emphasizes the role of trade as an important conduit of embodied knowledge flows (Coe and Helpman, 1995; Keller, 2002b). Importing a foreign intermediate good allows a recipient country to capture the R&D-, or 'technology'-content embodied in the traded good and consequently of the origin country. Alternatively, some studies consider foreign direct investment as an adequate proxy for knowledge flows (Blomstrom and Kokko, 1998). The trade-growth literature is, overall, reluctant to incorporate information on patent citations and technological space of the interaction units.[4]

A recent strand of spillover literature documents evidence on learning, across firms and geographic locations, via the mobility of highly-skilled personnel, the inventors of patents. The contribution of research personnel to knowledge diffusion has been acknowledged already by Arrow (1962).[5] But it was until the recent construction of detailed data on inventors' mobility by Lai et al. (2011) that a number of studies began to investigate the job mobility of inventors and its consequences on firms' innovation decisions and the learning effects (Agrawal et al., 2006; Kim et al., 2010).[6]

A rather emerging line of research studies the market of intellectual property and documents evidence from the market of patents. Patent transactions certainly serve strategic, legal, or financial purposes, but could also act as a conduit of acquiring technological knowledge. Investigation of this potential channel of knowledge diffusion has been hampered by data problems. A recent study by Serrano (2011) shows that the transfer of licenses and patents has become an important source of adopting technology for US firms.[7] We now have a detailed and rich data on all patent reassignments of all US-based entities (firms,

[1] See Keller (2004) for a review of the literature.

[2] The development of of detailed data on citations across patents allowed the investigation of flows of knowledge originating in countries, regions, universities, federal labs, and firms.

[3] Prior to the development of patent citations, an earlier related strand of research had examined knowledge diffusion between entities within the same technological group. Learning across entities is inferred via flows of intermediate capital goods (Terlecky, 1980), input-output matrices (Wolf and Nadiri, 1987), and correlations of specialization in technological sectors (Jaffe, 1986).

[4] Rather than considering flows, a parallel related literature infers knowledge spillovers from cross-country (industry) productivity correlations and (innovation) indices of country (industry) performance (Bloom et al., 2002; Keller, 2002a; Griffith et al., 2004; Cameron et al., 2005).

[5] As tacit knowledge (Polanyi, 1966) cannot be transferred into documentation such as papers, patents, and so on, face-to-face interaction between talented individuals is an important mechanism through which knowledge diffuses.

[6] The study of Agrawal et al. (2006) documents knowledge flows to an inventor's prior location are approximately 50% greater than if he had never lived there, suggesting that social relationships, not just physical proximity, are important for determining flow patterns. A recent work of Kim et al. (2010) explores the linkages between inventor' mobility and knowledge flows in nanotechnology sector confirming that the mobility of inventors enhances the citations across patents of firms that the inventor was previously employed. For any two patents in a technological field, A and B, where A and B are assigned to different firms and where A is granted after B, patent A is more likely to cite patent B if the patent A firm employs an inventor who earlier worked for the patent B firm.

[7] A stream of research by Serrano (2006), Serrano (2010), and Serrano (2011) develops and estimates models of costly technology transfer and renewal in the market for innovation and quantifies possible gains from trading patents as well as costs of adopting technology in the market for patents.

labs, government, universities) used so far only in the study of Serrano and in the present study in order to explore knowledge flows from the market of intellectual property (patents).

The present paper aims to provide novel insights in the diffusion of knowledge and its consequences for innovation activity. We study the determinants of different channels of knowledge flows within one framework, which is compatible with different knowledge spillover literature traditions. Then, we examine the effects of external accessible knowledge that operates via various channels on the production of innovation. In doing so, we compile rich and newly detailed data on cross-patent trade, citations, mobility of inventors, and trade of goods to learn more about the relative scope and intensity of knowledge flows across the states of the US.

More particularly, in this paper, we consider diffusion of knowledge that operates through four different channels: the market channel of contracts (i.e. trade of patents), non-market channel (citations of patents), the market channel of job mobility (inventor flows), and the market channel of traded goods and examine the effect of different aspects of space, geographic and technological, on all four knowledge flows. To our knowledge, this is the first attempt in the knowledge spillover literature where different, more or less, established channels of knowledge diffusion, which have been previously analyzed separately from different research avenues, are studied jointly within a common framework of analysis.

The conventional way in the knowledge spillover literature is the study of learning via a single channel of knowledge flows and, accordingly, the estimation of a single equation. Consequently, in comparing knowledge estimates across different channels, one has to resort in "borrowing" existing estimates from different branches (and model specifications) of the spillover literature. Possible inter-dependencies across different channels and omitted factors when one estimates single equations of knowledge flows could hamper the efficiency of the estimates.[8] We propose multivariate system estimation techniques, following the tradition of Tsionas (1999) and Winkelmann (2000), to account the different nature of the data (count and continuous) and potential unobserved heterogeneity. The development and application of novel and appropriate econometric techniques in this literature consists the second contribution of this paper.

The study of knowledge flows across the states of the US and the effect each channel of knowledge dissemination has on a state's production of innovation, constitutes our third and last contribution. The US is one of the most prolific nation as far as innovation activity is concerned. The geographic structure of the US, the contiguous states, the fairly high degree of cultural and institutional homogeneity and migration, the common currency, the large number of states, and the availability of rich and detailed data provides us with clear insights on the mechanisms of knowledge diffusion and on the acclaimed the influence of border and distance in shaping knowledge flows.

Our paper relates to and complements a number of important works. For example, the seminal study of Jaffe et al. (1993) examines the role of geographic distance as the major resistance factor of citation flows. In particular, the study compares the average geographic distance of patents that cite another patent and a random control group of patents that do not cite, controlling for same attributes of these two groups. Results show that firms which are located in the same city as the inventor are more likely than others to gain from knowledge spillovers from innovation. Subsequent studies of Peri (2005), Thompson (2006), and Alcacer and Gittelman (2006) extent the work of Jaffe et al. (1993) in various aspects. Peri (2005), for example, relies on citation function approach, while the other two studies use control group approach to identify localized spillover effects. Despite the significant contribution of the aforementioned studies, it is not always easy to interpret distance or border effect.[9] We also relate to the work of Mowery and Ziedonis (2001), which is the single study so far that explores knowledge flows, between firms and universities, via two channels: the channel of market of intellectual property (patents licenses) and the channel of non-market spillovers (citations) concluding that market (formal) knowledge flows, operating by academic licensing,

[8]Rapid technological change, depicted also in higher patent trade and citation exchange, and expanding trade could become a unified force affecting labor markets, including the market of inventors (Autor and David Dorn, 2013). In this paper, to derive proper comparisons, we jointly estimate knowledge flows that operate via different channels with employing system(s) of seemingly unrelated equations of knowledge flows. System estimates are more efficient in finite samples than single equation estimates (Zellner, 1962), as common elements (e.g. technology shock) omitted in each equation could cause error terms across equations to be correlated with each other and therefore producing less efficient estimates.

[9]For instance, Thompson (2006) does not include a distance measure at all, which may confound the effect of distance and border.

are more bounded by geographic distance compared to non-market (informal) flows exemplified by patent citations. However, as the authors state, their sample is small and compiled by patents from only three major US universities. Using system(s) of gravity-like specification of knowledge flows that operate via different channels, we try do not to conflate border with distance effects and further to uncover all possible influences of various distance internals. Along with geographic space, we also consider technological space and technological effort of the states to avoid overestimation of geography effect.

We apply our modeling approach to US states over the period 1993-2006 with two key questions in mind: (i) Is closeness important for knowledge flows across the states of the US? (ii) Do knowledge flows contribute to innovation production?

The evidence we provide is straightforward: we find the gravity model to perform quite well in explaining ideas and inventor exchange as goods trade. We capture the variance of cross-border exchange of knowledge with a parsimonious set of variables. Results show that geography, in terms of distance and contingency, greatly matters for the spread of knowledge as it has been massively documented in the spillover literature. The very significant impact of border and distance on knowledge exchange across states of the US is at first sight quite surprising, especially for knowledge flows based on ideas (traded patent and citation flows), and puzzling: unlike goods and people (inventors), ideas are 'weightless', and distance cannot just proxy only transportation costs. Along with transportation costs, distance is positively correlated with informational frictions. In other words, distance (and border) acts as a barrier to interaction, various micro-cultural affinities, and networking of economic agents. The continuing controversy over whether this localization bias is attributed to transaction costs, informational asymmetries, familiarity effects or other frictions remains pretty much also in this present paper, as we make no attempt to explore the causes of this bias. Results illuminate the relative geographic effectiveness and intensity of various channels of knowledge flows, which are generated in different markets (and non-markets).

In particular, we find that disembodied knowledge, generated from ideas, which are patented and then traded and cited, are less geographically restricted and, therefore, their effective reach is beyond that of knowledge embodied in goods trade or inventors' flows as the latter involve movements of goods and people, respectively. Further, non-market channel knowledge flows are more far-reached than market-based flows, with knowledge flows via inventors' mobility to be the most geographically confounded among the market-based flows. Finally, with respect to other types of closeness, technological effort proximity of states and production structure similarities, these greatly enhance knowledge interactions across states. Our results remain robust for different variable definitions, sub-samples, and alternative specifications.

The implications of our findings for the theoretical literature are potentially relevant. Although theoretical studies (Grossman and Helpman, 1991; Rivera-Batiz and Romer, 1991) emphasize the important consequences of disembodied knowledge flows over trade of goods flows, there has been little effort, on the empirical side, to thoroughly explore this issue. Along with other important studies, this paper makes an effort towards analyzing knowledge diffusion via different channels. We find that knowledge flows are relevant to a state's innovation production, as external accessible R&D gained through different flows has a strong positive effect on a state's innovation activity, which is as large, for some cases, as that of state's own R&D stock. Further, the effective reach of disembodied knowledge flows, exemplified via citation and traded patent flows, is larger than knowledge embodied in the trade of goods and inventors' mobility, confirming thus their importance for technology transfer and economic growth.

The remainder of the paper proceeds as follows. Section 2 introduces the framework of our analysis. Section 3 discusses our data. Section 4 presents the results. Section 5 summarizes our findings and concludes.

2. A Framework of Analysis

2.1. Production and Reach of Ideas

Knowledge flows take place when an idea, generated in region, country or institution, is learned by another region, country or institution. As the literature argues (Griliches, 1992), learning from someone else's ideas eventually implies the development of a stock of 'borrowed' or 'accessible' knowledge, proxied by R&D stock. The effect of this learning process can be measured by the impact of this borrowed or accessible R&D on the production of innovation.

We begin by describing the production of innovation activity in a given state i. We follow Peri (2005) and define innovation output in state i to be determined by the R&D activity in state i as well as by the R&D activity produced in states other than state i, but is accessible to state i. Therefore, the production function of innovation in state i can be expressed as follows:

$$Q_{it} = I_{it}(A_{it})^{\gamma}(A_{it}^{\alpha})^{\mu} \tag{1}$$

where Q index of innovative output; I a set of policy factors specific to state i; A is 'own' (within-state) R&D stock accumulated from past and current R&D investments in state i; and A^{α} is the stock of 'external' R&D accumulated in states other than i and accessible (α superscript denotes that) to state i at time t.

If the external and accessible R&D activity manages to perfectly and completely spill over to state i, is described by:

$$A_{it}^{\alpha} = \sum_{j \neq i} A_{jt} \tag{2}$$

In reality, however, the diffusion of innovation across states may be less than complete, the external accessible to state i R&D activity is then described by:

$$A_{it}^{\alpha} = \sum_{j \neq i} \phi_{ij} A_{jt} \tag{3}$$

where, the parameter ϕ_{ij} denotes the share of research results of state j learned by state i and the product $\phi_{ij}A_{jt}$ is the flow-weighted, available and external to state i, R&D stock.

Substituting equation (3) into equation (1) and by taking logs, equation (1) yields:

$$lnQ_{it} = lnI_{it} + \gamma lnA_{it} + \mu ln(\sum_{j \neq i} \phi_{ij} A_{jt}) \tag{4}$$

The parameter ϕ_{ij} is actually a vector or four parameters, $\phi_{ij} = \left[\phi_{ij}^{P}, \phi_{ij}^{C}, \phi_{ij}^{N}, \phi_{ij}^{G}\right]$, with each parameter, ϕ_{ij}^{f} (f = P, C, N, G) to represent the flow-weighted external R&D stock gained through a certain channel, namely via traded patent flows, ϕ_{ij}^{P}, citation flows, ϕ_{ij}^{C}, inventors' mobility flows, ϕ_{ij}^{N}, and trade of goods flows, ϕ_{ij}^{G}.

The empirical task of this paper is twofold. First, we estimate how geography shapes knowledge flows, ϕ_{ij}^{f}, in the states of the US. Then, we use these estimated values to assess the contribution of each flow on innovation production of a state, as described by equation (4).

Knowledge flows between states are shaped by states' geographical characteristics and a number of resistance factors such as technological effort, structure of production of the states. Following the literature (Peri, 2005), we model knowledge flows as below:

$$\begin{gathered}\phi_{ijt}^{f} = \beta_{ij} + \beta_1 State\,Border_{ij} + \beta_2 Nearby\,States\,[500\,miles]_{ij} + \\ \beta_3 Distance\,[500 - 1000\,miles]_{ij} + \beta_4 Distance\,[1000 - 1500\,miles]_{ij} + \\ \beta_5 Distance\,[1500 - 2000\,miles]_{ij} + \beta_6 Distance\,[2000 - 2500\,miles]_{ij} + \beta_7 Z_{ijt} + \epsilon_{ij}\end{gathered} \tag{5}$$

where ϕ_{ij}^{f} is (one of the four types of) knowledge flows between two states i (destination) and state j (origin); βij is origin and destination state fixed effects; $State\,Border$ is a dummy for adjacency and takes the value of 1 if states share a common border, 0 otherwise; $Nearby\,States\,[500\,miles]$ is a dummy for nearby area and takes the value of 1 if states are located within an area of 500 miles and do not share a common border; $Distance\,[500-1,000\,miles]$, $Distance\,[1,000-1,500\,miles]$, $Distance\,[1,500-2,000\,miles]$, and $Distance\,[2,000-2,500\,miles]$ are distance classes of 500 to 1,000 miles, 1,000 to 1,500 miles, 1,500 to 2,000 miles, and 2,000 to 2,500 miles, respectively; the vector Z contains controls, and ϵ is an iid error term.

The coefficients β_1 till β_6 provide a characterization of how geographic factors shape the flows of knowledge across states. The coefficient of the first dummy, $State\,Border$, captures how much knowledge exchange takes places between states that share a common state border compared to in-state knowledge exchange. Irrespective of the border effect, the second dummy, $Nearby\,States\,[500\,miles]$, captures the effect of geographic nearness of states that do not share common border compared to within-state knowledge flows. The coefficients of the rest of the distance dummies, examine whether states that have been located in various distance classes exchange exhibit different knowledge interactions in comparison to within-state knowledge interactions. We opted for this distance taxonomy, i.e., batches of 500 miles, because the longest distance in miles between two neighboring states is approximately 500 (517,705 miles to be precise), which is the distance between the centers of Colorado and Oklahoma. Then, we proceed with batches of 500 miles till the distance between East and West Coast is exhausted. One would expect that increasing geographic distance would reduce exchange knowledge across the states due to the presence of spatial transaction costs, signaling that knowledge flows are bounded in space and characterized by spatial declining effect. The localization of knowledge flows has been considerably tested in the spillover literature, which almost unanimously documents that physical distance does matter and spillovers are constrained geographically (Jaffe et al., 1993; Peri, 2005; Thompson, 2006; Alcacer and Gittelman, 2006; Belenzon and Schankerman, 2011).

However, states that are located nearby each other may exchange more knowledge with each other simply because they have similar technological efforts and/or production structures. Not allowing for technological and structural differences may lead to an over estimation of the geography effect. Therefore, we also consider, along with the geographic closeness, the technological and structural closeness between states. Vector Z includes the $TechnologicalDistance$, which captures the differences in technological efforts (investing in R&D and Scientists) between two states, and the $StructuralCloseness$, between two states, which is the correlation of patent profiles, when their citations are allocated to various technological fields and reflects the degree of sectoral production similarity.

More particularly, technological closeness purports at examining whether flows of ideas among states depend on the magnitude of innovative activity. One may suggest that states with high innovative activity are also those with most intense knowledge exchanges.[10] A state's ability to understand and exploit external knowledge depends on a state's experience in research; an idea analogous to absorptive capacity introduced by Cohen and Levinthal (1989). The higher the research experience, the higher the absorption of external flow of ideas and consequently the exchange of ideas among states. The technological effort of a state is proxied by the R&D activity.[11] Human capital (number of researchers) is another measure of technological effort.[12] We, therefore, proxy technological effort of a state as the share of R&D activity in the state to the number of scientists of that state. Thus, the technological distance between two states i and j

[10] The importance of technological similarity in shaping knowledge flows has been stressed in the literature (Bottazzi and Peri, 2003; Fischer and Griffith, 2008; Paci and Usai, 2009).

[11] According to innovation-driven models of growth (Cohen and Levinthal, 1989; Grossman and Helpman, 1991; Aghion and Howitt, 1997), R&D has two roles or 'faces'. The first role, which is the stimulation of innovation, has received the most attention in the existing empirical literature. The second role is in facilitating the imitation of others' discoveries. Rigorous econometric work assessing the statistical significance and quantitative importance of the 'two faces of R&D' has been provided (among others) by Hall and Mairesse (1995), Griffith et al. (2004), Cameron et al. (2005), and Kneller and Stevens (2006), who document that R&D is statistically and economically important in the catch up process as well as in directly stimulating innovation.

[12] Human capital can affect the absorption of existed advanced technologies (Abramovitz, 1986; Benhabib and Spiegel, 1994). Further, it accounts for aspects of innovation not captured by the innovation sector (e.g., R&D), including 'learning-by-doing' and 'on-the-job-training' (Romer, 1989; Redding, 1996).

for a given year, t, is calculated as:

$$TechnologicalDistance = \mid ln \frac{R\&D_i}{Scientists_i} - ln \frac{R\&D_j}{Scientists_j} \mid$$

Structural closeness aims at examining whether ideas flows occur between states with comparable production structures, that is, states specialized in similar sectors. This is because researchers are expected to benefit more from other researchers who work in the same or related sectors (Bode, 2004; Paci and Usai, 2009). We expect to find a positive association between intensity of knowledge flows between two states specialized in similar sectors. To measure structural proximity between two states, we follow Jaffe (1986).[13] We first classify each patent, according to their primary US Classification, in one of the 37 technology fields, as defined in Hall et al. (2001).[14] Then, for each state, we create a patent profile by taking the vector of shares of patents issued in technology field $Sh_i = (sh_{i1}, sh_{i2}, ..., sh_{i37})$, for a given year. Then, the structural similarity between two states i and j for given year, t, is the uncentered correlation of their patent profiles and is calculated as:

$$StructuralCloseness = \frac{{sh_i}' \; sh_j}{\sqrt{\sum_{s=1}^{37} sh_{is}^2 \sum_{s=1}^{37} sh_{js}^2}}$$

This is basically the correlation of patent portfolios between state i and state j and, therefore, captures the degree of similarity between state i and state j. The constructed index ranges from zero (minimum closeness), which implies that the production structures are orthogonal to one (maximum closeness), which implies identical sectoral structure (patenting in similar sectors) in two states.

2.2. Estimation Strategy

We first proceed with the joined estimation of trade patent flows, citation flows and inventors' mobility flows. All aforementioned flows are count data and there is sufficient long information. A first candidate model for our case is the non-linear seemingly unrelated (SUR) Poisson model, introduced by King (1989). A serious limitation of such model, however, is its inability to account for over-dispersion or extra-Poisson variation in the data. To account for seemingly unrelated count data with over-dispersion, Winkelmann (2000) proposes an alternative model, which does not abandon the basic convolution structure of the seemingly unrelated Poisson, but rather generalizes some of its assumption to allow of over-dispersion. The proposed model has negative binomial marginals and is referred to as seemingly unrelated negative binomial model. Building on the work of Tsionas (1999) and Winkelmann (2000), we further account for potential unobserved heterogeneity in the system and therefore estimate a tri-variate system of seemingly unrelated negative binomial regressions allowing for unobserved heterogeneity.

Second, we include trade flows - despite the limited information we have on these flows - in our estimation and, therefore, enhance the tri-variate system of count dependent variables with one more equation, the trade of goods, which is continuous. In doing so, we estimate a non-linear mixed SUR model for count and continuous responses applying Bayesian techniques developed in this paper, which is deleopd and used first time in the literature in the present paper. The next two sub-sections present these methodologies.

As an exercise and in comparison with the literature, we also estimate single (univariate) equations for each channel of knowledge flows, for traded patents, citations, inventors' mobility and trade of goods flows. We use negative binomial estimation techniques for the first three flows and OLS with state and year fixed effects for the latter.[15]

[13]There are different ways to measure structural distance. For a review, see, Los (2000).

[14]Hall et al. (2001) had categorazided US classifications in 36 broad technology fields; however, in the 2006 NBER update, there was an addition of a 37th technology field in the area of Computers and Communication Technologies.

[15]Negative binomial estimation is also used in similar to ours contexts (Peri, 2005; Perkins and Neumayer, 2011; Furman and Stern, 2011).

2.2.1. Multivariate Negative Binomial Regression

We introduce multivariate negative binomial regression along with techniques for statistical inference. Our point of departure is Winkelmann (2000). To summarize Winkelmann's approach suppose $y_s = [y_{s1}, ..., y_{sM}]'$ is an $M \times 1$ vector of count variables, for a particular observation, $s = 1, ..., n.$, where n is the number of observations $ixjxt$. For a single count variable, say Y_s, is well-known that the negative binomial (NB) specification arises from the Poisson, if we assume: $Y_s|v_s \sim \text{Po}(\lambda v_s)$ and $v_s \sim \text{Ga}(\alpha, \alpha)$, where "Po" denotes the Poisson and "Ga" the gamma distributions. Then, it can be shown that Y_s follows a NB distribution with mean λ and variance $\lambda + \alpha^{-1}\lambda^2$. In the multivariate context, Winkelmann (2000) proposes the following model:

$$y_s = y_s^* + u_s \iota_M \tag{6}$$

where the scalar random variable u_i follows a NB, ι_M is a vector of ones in $\Re^M$ and $y_{sm}^*|v_{sm} \sim \text{Po}\left(\lambda_{sm} v_{sm}\right)$, $v_{sm} \sim \text{Ga}\left(\alpha, \alpha\right)$, m = 1,..., M

For u_s the NB assumption leads to the formulation:

$$u_s|v_{so} \sim \text{Po}\left(\lambda_{so} v_{so}\right), v_{so} \sim \text{Ga}\left(\gamma, \gamma\right)$$

One can introduce covariates by assuming:

$$\lambda_{sm} = \exp\left(x_s' \beta_m\right), m = 1, ..., M$$

where x_s is a $k \times 1$ vector. Using the re-parametrization $\mu_{sm} = \lambda_{sm}/\sigma$ the mean remains the same and the variance becomes $\text{Var}\left(y_{sm}^*\right) = \lambda_{sm}\left(1 + \sigma\right)$. The covariance matrix of y_s is $\text{Cov}\left(y_s\right) = \left(\Lambda_s + \gamma\iota\iota'\right)\left(1 + \sigma\right)$, where $\Lambda_s = diag[\lambda_{i1}, ..., \lambda_{sm}]$.

Relative to the multivariate Poisson distribution there is an over-dispersion parameter given by $1 + \sigma$ and as $\sigma \to 0$ the distribution approaches the multivariate Poisson.

Unlike Winkelmann (2000), we introduce further unobserved heterogeneity in the following form:

$$\log \lambda_{sm} = x_i' \beta_m + \varepsilon_{sm}, m = 1, ..., M \tag{7}$$

where $\epsilon_s \sim N_M(0, \Sigma)$.

To develop Markov Chain Monte Carlo (MCMC) methods for inference, we write the distribution of y_s conditional on all latent variables $\mathcal{U}_s = (v_s, u_s, \lambda_s)$ and parameters θ of the model:

$$\begin{array}{r} p(y_s|\mathcal{U}_s, \theta) = \frac{\alpha^{\alpha M}}{\Gamma(\alpha)^M} \prod_{m=1}^{M} \left\{ \exp\left(-\lambda_{sm} v_{sm}\right) \frac{\left(\lambda_{sm} v_{sm}\right)^{y_{sm} + u_s}}{\left(y_{sm} + u_s\right)!} v_{sm}^{\alpha - 1} \exp\left(-\alpha v_{sm}\right) \right\} \\ \frac{\gamma^\gamma}{\Gamma(\gamma)} \exp\left(-\lambda_{so} v_{so}\right) \frac{\left(\lambda_{so} v_{so}\right)^{u_s}}{u_s!} v_{so}^{\gamma - 1} \exp\left(-\gamma v_{so}\right) \\ (2\pi)^{-m/2} |\Sigma|^{-1/2} \left\{ \prod_{m=1}^{M} \lambda_{sm}^{-1} \right\} \exp\left\{ -\frac{1}{2} \left(\log \lambda_s - X_s \beta\right)' \Sigma^{-1} \left(\log \lambda_s - X_s \beta\right) \right\} \end{array} \tag{8}$$

where $X_s = I_M \otimes x_s'$ and $\beta = \left[\beta_1', ..., \beta_M'\right]'$.

The first line is the joint distribution of λ_{sm} and v_{sm} ($m = 1, ..., M$) conditional on u_s, the second line provides the distribution of u_s, while the third line gives the distribution of λ_s conditional on the observed covariates.

Therefore, we estimate equation (7) for the case of three count variables, patent trade, citations, and inventor flows (tri-variate seemingly unrelated negative binomial regressions):

$$\log \lambda_{sm} = x_i' \beta_m + \varepsilon_{sm}, m = 1, 2, 3 \tag{9}$$

where x_s' is the vector of geographic and technological factors as defined in equation (5), β the vector of the corresponding coefficients, and ϵ_s error term.

2.2.2. Mixed Continuous and Count Responses

Suppose that we have an additional $R \times 1$ vector of responses, y_{ir}, $r = M+1, ..., M+R$. The most reasonable way to handle the matter in the multivariate situation is to extend (7) in the following form:

$$\begin{array}{r} \log \lambda_{im} = x_s' \beta_m + \varepsilon_{sm}, m = 1, ..., M \\ y_{sr} = x_s' \beta_r + \varepsilon_{sr}, r = M+1, ..., M+R \end{array} \tag{10}$$

We redefine $\varepsilon_s = [\varepsilon_{s1}, ..., \varepsilon_{sM}, \varepsilon_{s,M+1}, ..., \varepsilon_{s,M+R}]'$ and assume:

$$\epsilon_s \sim N_{M+R}(0, \Sigma)$$

where Σ is an $(M+R) \times (M+R)$ covariance matrix.

The major change in the distribution of observables $p(y_s|U_s, \theta)$ is in the third line of equation (8), which should now be:

$$(2\pi)^{-m/2} |\Sigma|^{-1/2} \left\{ \prod_{m=1}^{M} \lambda_{sm}^{-1} \right\} \exp\left\{ -\tfrac{1}{2} (\Psi_s - X_s\beta)' \Sigma^{-1} (\Psi_s - X_s\beta) \right\}$$

where $\Psi_s = [\log \lambda_s', y_{s,M+1}, ..., y_{s,M+R}]'$.

Bayesian analysis for the multivariate Poisson regression model, developed by Tsionas (1999) can be applied in this case as well. The analysis is organized around MCMC methods for inference.

We therefore estimate equation (10) for three count variables, patent, citation, and inventors' mobility flows, and one continuous, trade of goods (four-variate mixed system of count and continuous responses):

$$\begin{array}{r} \log \lambda_{sm} = x_s' \beta_m + \varepsilon_{sm}, m = 1, 2 \\ y_{sr} = x_s' \beta_r + \varepsilon_{sr}, r = 1 \end{array} \tag{11}$$

where y is knowledge flows, x_s' is the vector of geographic and technological factors as defined in equation (5), β the vector of the corresponding coefficients, and ϵ_s error term.

3. Data Description and Analysis

Our empirical analysis is based on 48 states of the US (excluding Alaska, District of Columbia and Hawaii) for the period 1993 to 2006.

The primary source of patent data is a recently compiled dataset by the office of the chief economist of the United States Patent and Trademark Office (USPTO) referred as Patent Assignment Dataset, which contains assignments (transactions) of US issued patents between entities registered at the USPTO.[16] A typical assignment is characterized by a unique identifier (i.e., reel frame), the names of the buyer (i.e., assignee), the seller (i.e., assignor), the date that the transaction agreement was signed (execution date), and the patent numbers or patent applications that are traded per assignment.[17] In constructing our patent dataset, we faced two main challenges when employing assignment data. The first relates to the fact that entities are not required to disclose transactions to the USPTO. However, for legal and perhaps accounting reasons, they have incentives to do so.[18] A challenge associated with using assignment data is that it is still likely that a number of transactions have not been disclosed to the USPTO due to negligence or to strategic behavior. In any case, we do not expect this to be systematic for aggregated transactions across geographical areas. An additional challenge is associated with excluding 'routine' transactions. In the

[16] In the US, when entities transfer US issued patents to other entities, they disclose such transactions to the USPTO. The latter are called assignments.

[17] There is also a field in the assignment data in which entities can disclose the justification for the transfer. However, the justification, in most cases, is a generic one (i.e. assignment of assignor's interest). Therefore, it is really difficult to extract information from that field.

[18] For instance, in a potential litigation the courts will need to know clearly which firm or organization holds the intellectual property in question.

US, only an individual can file for a patent application. Subsequently, this individual may re-assign the patent application (or patent) to her firm or institution where she is employed. These transactions are also included to the dataset. Thus, the challenge here is to isolate the economically meaningful re-assignments and discard otherwise. Taking these two challenges into consideration, we end up having 128,578 patents issued between 1988 and 2006 and have been traded from 1993 to 2006 between US located entities and are associated with 65,558 transactions for which we have address information for both the assignor and assignee.

There is a many-to-many relationship between patents and transactions. That is, transactions may contain more than one patents and a patent may be transacted more than once. To construct the flows of patents, we aggregate the number of patents that have been traded from entities located in origin state to entities in destination state for any year. For patents, which are traded more than once, each transaction is registered as a new transaction and, therefore, counted accordingly.[19]

Patent citation data originate from the National Bureau of Economics Research (NBER) *Patent Data Project*, which is publicly available and described in detail by Hall et al. (2001).[20]. The database contains citations of all US issued patents up until 2006. We construct bilateral citation flows between states for the period 1993-2006 by considering citations made from 1993 to 2006 to all US patents issued from 1988 to 2006. We then distinguish citations into citations of traded patents and citations of non-traded patents and accordingly construct citation flows of traded patents and citation flows of non-traded patents. Information about the nature of a patent (whether it is traded or not) is retrieved from the USPTO. To construct bilateral citation flows across states, we consider the location of the patent owner (written on the patent document wrapper) for all (citing and cited) patents.

Information on inventors' mobility flows, defined as the number of firms or states a patent inventor changes during his lifetime when he every time files for a new patent, is obtained from the Disambiguation and Co-authorship Networks of the US Patent Inventor Database, which is publicly available and described in detail by Lai et al. (2011).[21] The original NBER database includes authorship and firm and state level data but did not identify unique inventors over time. This is a non-trivial task because the USPTO does not require consistent and unique identifiers for inventors. This database contributes an improved disambiguation algorithm and a database with which the co-authorship networks of utility patent inventors is constructed.

Bilateral merchandise trade flows are extracted from the Bureau of Transportation Services *Commodity Flow Surveys* and for the years 1993,1997, and 2002.

Finally, data on geographic characteristics of the states, as well as, data to construct technological and structural closeness are obtained from a range of sources. The geographic distance (in miles) between two states is the distance between each state's geographical center calculated as the crow flies. This information is obtained from Google Maps.[22] Information on a state's R&D expenditure and number of scientists (i.e. science, engineering, and health researchers) to construct the technological effort of a state, is extracted from from the National Science Foundation *Science and Engineering State Profiles*. Using state level R&D spending data and the perpetual inventory method as in Guellec and van Pottelsberghe de la Potterie (2004) we also construct R&D stocks that will be used to estimate elasticities in the innovation function.[23] Further, to construct the index of structural proximity, we need to allocate patents into different technological fields. Patents' primary US Classifications are retrieved from NBER.

[19]For patents traded more than once, and for robustness purposes, we construct two alternative measures of patent flows. The first is called, 'first flow' and considers, for each patent, only its first assignment, ignoring the rest of its transactions. The second measure is called 'last flow' and for each patent excludes all the intermediate transactions and records only the assignment between the first and last entity. For example, for a certain patent which is sold from California to New York state and then from New York state to Texas, the measure 'all flows' registers both transactions, while the other two measures register only one transaction: 'first flow' registers the transaction between California to New York state, and 'last flow' registers the one between California to Texas.

[20]The database is available at: http://sites.google.com/site/patentdataproject.

[21]Information on the data is provided at http://hdl.handle.net/1902.1/15705 UNF:5:9kQaFvALs6qcuoy9Yd8uOw== V1.

[22]See http://www.freemaptools.com.

[23]Following the literature, we have tried different depreciation percentages, e.g., 15%, and 20%. The resulted R&D stocks are highly correlated.

Table 1 below provides the descriptive statistics of the variables included in our model:

Table 1: Summary Statistics

variables	Observations	Mean	St. Dev.	Min	Max
Patent Trade Flows	32256	4.71	41.63	0	3262
Citation Flows	32256	210.76	1316.56	0	85287
$Citation_T$ *Flows*	32256	35.46	252.26	0	18003
$Citation_{NT}$ *Flows*	32256	126.45	897.03	0	61205
Inventor Flows	32256	13.07	174.01	0	10669
Trade Flows	6227	3396.93	16467.95	0.96	535263
State Border	2256	0.10	0.29	0	1
Nearby States [500 *miles*]	2256	0.12	0.32	0	1
Distance $[500 - 1,000\, miles]$	2256	0.32	0.47	0	1
Distance $[1,000 - 1,500\, miles]$	2256	0.25	0.43	0	1
Distance $[1,500 - 2,000\, miles]$	2256	0.13	0.34	0	1
Distance $[2,000 - 2,500\, miles]$	2256	0.08	0.27	0	1
TechnologicalDistance	32256	0.63	0.50	0	3
StructuralCloseness	32256	0.70	0.18	0.05	1

Note: *Patent Trade Flows*, *Citation Flows*, $Citation_T$ *Flows*, and $Citation_{NT}$ *Flows*, and *Inventor Flows* are occurrences; *Trade Flows* is merchandise trade flows in millions of constant (2000) US dollars, *State Border* is dummy (1 if states for common border, 0 otherwise), different distance classes of 500 miles are dummies (1 if states locate with the class, 0 otherwise); *TechnologicalCloseness* ranges from 0 (similar) to 3 (dissimilar) and *StructuralCloseness* from 0 (dissimilar) to 1 (similar).

According to Table 1, states, on average, trade 4.71 patents per year and exchange 210.76 citations, out of which 35.46 are of traded patents and 126.45 of non-traded patents. In other words, citation flows associated with traded patents comprise in our sample 16.82% of total citation flows. On average, states trade, per year, goods of 3.4 billion dollars value. For every state pair, in a given year, there are, on average, 13 occurrences of inventor mobility. On average, each state pair is 10% likely to be neighboring with each other. Furthermore, 12% of all possible pairs of states (i.e., 48*48-48=2256) are closer than 500 miles and do not share common state border, 32% are located in a distance of 500 to 1,000 miles, 25% in a distance of 1,000 to 1,500 miles, 13% between 1,500 and 2,000 miles and 8% within a range of 2,000 to 2,500 miles. In terms of technological effort, states, on average, appear to be less distant than the maximum potential distance that they could have. Finally, they also appear to be, on average, quite close in terms of technological specialization in their production.

Below, Figure 1 shows the production of innovation in the US over the period 1993-2006. As it is apparent, intense innovation activity is concentrated in few states in the US. More than 60% of production of patents takes place in five states, California (CA) and New York (NY), which stand out among the top producers, followed by Texas (TX), Illinois (IL), and New Jersey (NJ). The least involved states in producing innovation are Alaska (AK), Hawaii (HI), North Dakota (ND), South Dakota (SD), and Wyoming (WY). Moreover, states which are patent production leaders are also top performers across all measures of innovation, citing, and trading activity.

Table A.1 in the Appendix reports summary statistics per state. States like California (CA), New York (NY), Illinois (IL), Massachusetts (MA), Michigan (MI), and Texas (TX) are top patent producers, traders of patents and goods, and receive the most citations (of either traded or non-traded patents). These states also enjoy the presence of most scientists in the country and exhibit exhibit the most occurrences of inventors' mobility. In the opposite side of the spectrum are the states of Alaska (AK), Hawaii (HI), the Dakotas (SD and ND) and Wyoming (WY). Finally, the average US state spends about 2% of state GDP in R&D. Overall, summary statistics per state reveal a large variety of patterns.

Figure 1: Patent Production per State

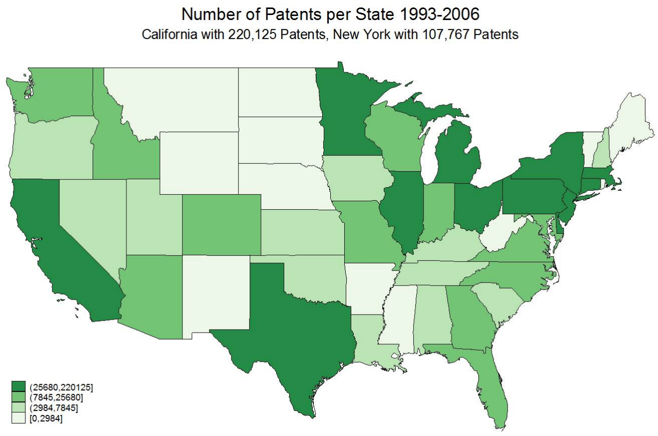
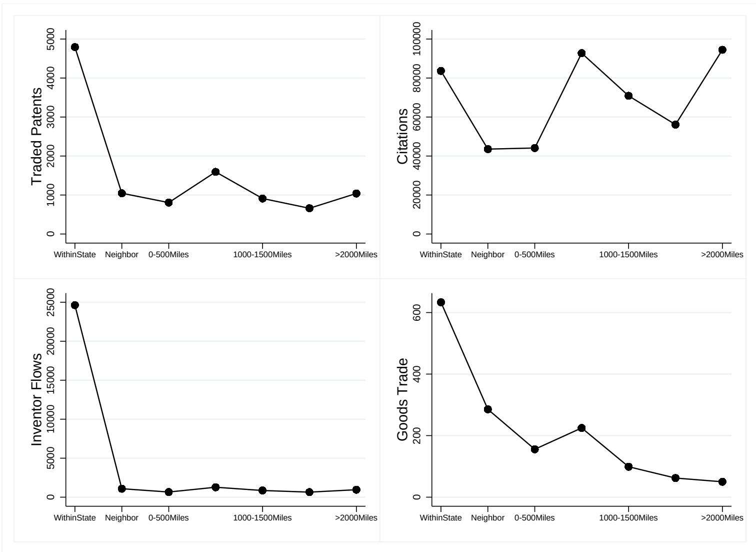


Figure 2 below shows the estimated decay of knowledge flows moving out of a state, out of the nearby neighbor, and out by steps of 500 miles. The four decay functions reported are the knowledge flows based on traded patents, citations, inventors' mobility and trade of goods.

Figure 2: Decay of Knowledge Flows Due to Geographical Barriers



The evidence from the figure above confirms the significance of state border and geographic distance in shaping knowledge flows. As one observes, the decay effect is apparent: state border and distance restricts all different channels of knowledge flows; however, they have a much larger impact on market-based flows, especially on inventors' mobility and trade of goods flows, which are sharply and fast, reduced, than on non-market (citation) flows. While inventors' mobility slightly pick up again, trade flows continue to diminish. In contrast, the geographic reach of knowledge flows via traded patents and citations dies out rather smoothly and pick up again, after some thousand miles, due to the intensity of ideas exchange between East and West (California). Finally, all flows for distance between 500 and 1,000 miles show a pick, which is probably due to flow exchange with Illinois, Texas, and Michigan.

4. Empirical Results

This section presents our results. We first examine whether closeness shapes knowledge flows that operate via four different channels across the states of US Then, we examine whether knowledge flows have an effect on state's innovation production.

4.1. *How Important is Closeness for Knowledge Flows?*

We investigate flows originating from an average state and then flows originating from the most innovating states of the US Accordingly, we present our findings.

Knowledge Flows from the Average State

Table 2 reports the results. Columns (i) to (iv) report univariate negative binomial estimates from equation (5) when it is separately estimated for traded patent (*Patent Flows*), citation (*Citation Flows*), inventor (*Inventor Flows*), and trade of goods (*Goods Flows*) flows, respectively. Then, traded patent, citation, and inventor flows are jointly estimated and the negative binomial estimates of the multivariate (tri-variate) SUR as expressed in (9) are reported in column (v). Lastly, columns (vi) reports the estimated coefficients of the multivariate (four-variate) mixed count and continuous SUR as expressed in (11) for the years 1993, 1997, and 2002 (as bilateral trade flows are available only for these years). Standard errors are in parentheses.

As Table 2 shows, all types of closeness are important for knowledge flows across the states of the US Single (univariate) equation estimates are, in most cases, quite close to their SUR counterparts, with the latter to be somewhat larger than the former. However, the standard errors are smaller in the SUR approaches, compared to those of the univariate for almost all parameter estimates. The main advantage of the system over single equation estimates is clearly an increase in the efficiency of the SUR estimates. We further test for equality of coefficients between single equation and SUR estimation.[24]. The values of χ^2 of each SUR indicate that the coefficients of univariate are different from the multivariate SUR. Lastly, we also rejected the hypothesis that the coefficients of single regressions as well as those of SURs are zero. Below, we embark on discussing our findings based on SUR estimates.

Each geographic coefficient in Table 2 captures the difference between knowledge flows diffused in geographic space to knowledge flows within a state, which is our benchmark by model construction. To convert each value to percentage change, one needs to use the exponential formula. Beginning with column (v), the coefficient of *State Border* for the case of the traded patent flows implies that states, which are neighbors and, therefore, share common border, exchange about 90% (= $e^{-2.31}$) less knowledge to what they would exchange within their borders. In other words, on crossing a state border, knowledge diminishes to about 10% (= 1 - $e^{-2.31}$) compared to within state level of knowledge. Irrespective of the border, distance also shapes knowledge flows that operate via traded patents. States that are nearby (in the vicinity of 500 miles), but do not share common borders, exchange 92.4% (= $e^{-2.58}$) less knowledge than what they would exchange within themselves as the coefficient of *Nearby States* [500 *miles*] indicates. Put it differently, knowledge crossing nearby but not adjacent states, diminishes to 7.6% (= 1 - $e^{-2.58}$) to its initial (within state) level. Further, there is about 0.5% reduction of knowledge flows for each 500 miles traveled farther than 500 miles. Consequently, the geographic scope of knowledge flows, based on patent trade, is limited mainly by the state border, which is an important hurdle to out-of-state knowledge transmission, and by distance. Significant knowledge reduction takes place already within a district of 500 miles. Any further increase of distance has, practically, no additional reduction effect. An interesting finding, however, emerges when distance between states becomes long enough, specifically bigger than 2000 miles, as there is a slight increase of knowledge flows of 0.4% compared to the average effect of all previous distance intervals. This seemingly controversial finding is due to the "California effect'. Despite of its distant, from a typical state, location, California is an exceptional producer and trader of patents.

[24]If p is a set of coefficients from SUR and p_0 the counterpart from single equation, let $d = p - p_0$ and suppose V is the covariance of d, $V = cov(d)$, then $d' inv(V) * d \sim X^2$

A somewhat different picture emerges when knowledge diffuses via citation flows, as this channel appears to be less geographically restricted compared to knowledge flows that operate via patent trade. State border and geographic distance let a small fraction of initial knowledge to transcend, but this portion is significantly larger than that based on trade patent flows. Specifically, less knowledge, about 76% (= $e^{-1.42}$), is exchanged between bordered states to what is exchanged within state border, while knowledge interactions in a neighborhood of 500 miles are by 79% (= $e^{-1.55}$) less compared to in-state level of knowledge. In other words, when knowledge, based on citations, crosses a state border, is reduced to 24% (= $1 - e^{-1.42}$) of the in-state level of knowledge, and to 21% (= $e^{-1.55}$) on crossing an area of 500 miles. Furthermore, exchange of knowledge between states that are located from 500 to 2,000 miles is by 80% (mean of estimates of different classes of *Distance*) less compared to knowledge exchange in a state, with an additional reduction of 1.75% for each 500 miles traveled. When distance becomes larger than 2,000 miles, knowledge flows increase, compared to the flows exchanged within previous distance intervals, by almost 8%, which is again due to the California effect.

The opposite seems to be the case for the inventors' mobility flows, which are heavily shaped by geography. As the last set of estimates in column (v) reveal, only 2.4% (= $1 - e^{-3.71}$) of knowledge carried by inventors crosses a state border, which is indicative of the strong border effect on embodied knowledge. In addition, only 1.7% (= $1 - e^{-4.07}$) of inventors's knowledge crosses the vicinity of 500 miles, and this percentage remains unaltered for any farther traveled distance, implying that the die-out effect is large and sharp. Nonetheless, there is an increase of these flows again after a distance of 2,000 miles, but the size of this effects is diminutive.

Our tri-variate SUR estimates have shown so far that market-based knowledge flows that operate either via patent trade or inventors' mobility flows are more geographically restricted compared to non-market knowledge spillovers that operate via citation flows. In particular, the geographic scope of the latter is 2 to 3 times bigger than that of traded patents and 10 to 12 times larger than that of inventors' mobility. Between the two market-based channels, mobility of inventors is 5 to 6 times more geographically localized than patent trade flows. The potential need for the patent buyer to maintain contact with the inventor attaches a more nuance role to geography in shaping patent trade than shaping citations and consequent knowledge generated. Furthermore, as knowledge diffusion via traded patents does not require movements of people, the effective reach of traded patent and citation flows, and therefore subsequent diffusion of knowledge, should be beyond that of inventors' mobility.

From the preceding analysis one could conclude that the drop in learning, due to state border and geographic distance, is quite substantial across different channels of knowledge flows. However, states that are located close to each other may exchange more knowledge with each other simply because they have similar technological efforts and/or production structure. We, therefore, consider along with geographic closeness, technological and structural closeness of states. The coefficient of $TechnologicalDistance$ captures the difference in the technological efforts (investing in R&D and Scientists) between two states, while the coefficient of $StructuralCloseness$ reflects the degree of technological specialization similarity of states in their production sectors. Examining the coefficients of both indices, we find that technological effort dissimilarity between states tends to decrease knowledge flows. Specifically, a state receives 4% (via citation flows) up to 28% (via mobility of inventors) more knowledge from a state which pours the same technological effort as itself ($TechnologicalDistance = 0$) than from a state with dissimilar technological efforts ($TechnologicalDistance = 1$). Technological proximity between states is more essential to the market-based knowledge flows, particularly to inventors' mobility flows. As the literature stresses, investment in R&D and human capital makes a region attractive to talented individuals like inventors (Lucas, 1988). Moreover, two states with similar technological sectors ($StructuralCloseness = 1$), exchange 12% (via inventor mobility) up to 118% (via traded patent flows) more knowledge flows compared to states with completely dissimilar ($StructuralCloseness = 0$) technological sectors. Structural specialization between states appears to matter more for disembodied knowledge flows that operate via traded patents and patent citations, as researchers are expected to benefit more from other researchers who work in the same or related technologies (Bode, 2004; Peri, 2005; Paci and Usai, 2009).

Table 2: Determinants of Knowledge Flows in the U.S

	Sigle Equation Estimates[(I)]				Tri-variate SUR Estimates[(II)]			Four-variate SUR Estimates[(III)]			
	PatentFlows[a]	*CitationFlows*[b]	*InventorFlows*[c]	*GoodsFlows*[d]	*Patent Flows*[a]	*CitationFlows*[b]	*InventorFlows*[c]	*Patent Flows*[a]	*CitationFlows*[b]	*InventorFlows*[c]	*GoodsFlows*[d]
	(i)	(ii)	(iii)	(iv)		(v)			(vi)		
State Border	-3.01***	-1.66***	-4.42***	-2.26***	-2.31***	-1.42***	-3.71***	-2.25***	-1.31***	-3.77***	-2.26***
	(0.122)	(0.043)	(0.036)	(0.083)	(0.002)	(0.000)	(0.004)	(0.008)	(0.000)	(0.010)	(0.010)
Nearby States [500 *miles*]	-3.59***	-1.76***	-5.13***	-3.25***	-2.58***	-1.55***	-4.07***	-2.45***	-1.48***	-4.27***	-3.23***
	(0.122)	(0.043)	(0.037)	(0.080)	(0.003)	(0.000)	(0.006)	(0.009)	(0.000)	(0.016)	(0.016)
Distance $[500-1,000\, miles]$	-3.75***	-1.87***	-5.41***	-3.96***	-2.66***	-1.59***	-4.19***	-2.52***	-1.46***	-4.6***	-3.94***
	(0.125)	(0.043)	(0.036)	(0.079)	(0.002)	(0.000)	(0.004)	(0.006)	(0.000)	(0.010)	(0.010)
Distance $[1,000-1,500\, miles]$	-4.05***	-1.99***	-5.65***	-4.58***	-2.74***	-1.67***	-4.26***	-2.59***	-1.52***	-4.55***	-4.57***
	(0.142)	(0.043)	(0.034)	(0.081)	(0.003)	(0.000)	(0.004)	(0.010)	(0.000)	(0.009)	(0.009)
Distance $[1,500-2,000\, miles]$	-4.06***	-2.06***	-5.73***	-4.91***	-2.76***	-1.77***	-4.32***	-2.62***	-1.64***	-4.51***	-4.89***
	(0.134)	(0.044)	(0.041)	(0.083)	(0.003)	(0.000)	(0.006)	(0.011)	(0.000)	(0.013)	(0.013)
Distance $[2,000-2,500\, miles]$	-4.24***	-1.88***	-5.64***	-4.98***	-2.74***	-1.45***	-4.30***	-2.61***	-1.29***	-4.30***	-4.97***
	(0.134)	(0.044)	(0.043)	(0.088)	(0.002)	(0.000)	(0.003)	(0.006)	(0.000)	(0.008)	(0.008)
TechnologicalDistance	-0.29***	-0.14***	-0.33***	-0.13***	-0.25***	-0.01***	-0.33***	-0.22***	0.01***	-0.84***	-0.13***
	(0.078)	(0.012)	(0.019)	(0.023)	(0.002)	(0.000)	(0.004)	(0.008)	(0.000)	(0.010)	(0.010)
StructuralCloseness	1.07***	1.19***	0.60***	0.031	0.78***	0.25***	0.11***	0.76***	1.02***	0.82***	0.05***
	(0.195)	(0.048)	(0.063)	(0.093)	(0.006)	(0.000)	(0.011)	(0.024)	(0.000)	(0.029)	(0.029)
Wald chi2(115)	10928.76	272080.70	69083.07								
$\chi^2(30)$						4590			5790		
Observations	32,256	32,256	32,256	6,227	32,256	32,256	32256	6,227	6,227	6,227	6,227

All regressions include time dummies and origin and destination state fixed effects. Standard errors, reported in parentheses, are heteroscedastic robust standard errors only for the single equation estimates; Coefficients of *constant* term are omitted; *State Border* is dummy for adjacency and is 1 if states share common border and 0 otherwise; *Nearby States* [500 *miles*] is a dummy for geographic nearness and is 1 when states are located nearby, within 500 miles and do not share common state border, and 0 otherwise; *Distance* $[500-1,000\, miles]$, *Distance* $[1,000-1,500\, miles]$, *Distance* $[1,500-2,000\, miles]$, and *Distance* $[2,000-2,500\, miles]$ are dummies for distance classes of 500 mile length and take the value of 1 if states are located between 500 to 1,000, 1,000 to 1,500, 1,500 to 2,000, and 2,000 to 2,500 miles, respectively and 0 otherwise; *TechnologicalCloseness* is the degree of proximity of the technological effort (*R&D/Scientists*) of states; *StructuralCloseness* is the degree of technological similarity of production sectors of states; (***): significance at 1% level.

[I] Specification (I) reports single (univariate) negative binomial equation estimates (columns i, ii, ii, iv).

[II] : Specification (II) reports negative binomial estimates of tri-variate SUR (column v).

[III] : Specification (III) reports estimates of four-variate SUR of mix count and continuous responses (column vi) for the years 1993, 1995, and 2002 (due to trade flows data availability).

[a] *Patent Flows* are patent trade flows of the average state.

[b] *Citation Flows* are citation flows of the average state.

[c] *Inventor Flows* are inventor flows of the average state.

[d] *Goods Flows* are trade (imports) flows of the average state.

We extent our discussion by including in our analysis one more market-based channel of embodied knowledge flows, that of trade (merchandise) of goods. To perform such comparison with the rest of the flows, we develop and estimate a multivariate (four-variate) SUR model for mixed count and continues responses. In doing so, we are able to discuss, in a common framework, acclaimed conduits of knowledge diffusion such as citation and trade of goods flows, which were previously examined by separate literature strands, and properly compare their estimates.

Estimates reported in column (vi) of Table 2 are remarkably similar to their counterparts reported in column (v). On crossing a state border knowledge diminishes to 11% (= 1 - $e^{-2.25}$) based on trade patent flows, 27% (= 1 - $e^{-2.31}$) based on citation flows, 2% (= 1 - $e^{-3.77}$) based on inventors mobility, and to 10% (= 1 - $e^{-2.26}$) based on goods flows compared to in-state level of knowledge. Irrespective of border, distance, on average, diminishes knowledge via traded patent flows to 8% (= 1 - $e^{-2.56}$, where -2.56 is the mean of the five distance interval coefficients), 23% (= 1 - $e^{-1.48}$) via citation flows, 1% (= 1 - $e^{-4.44}$) via inventor mobility flows, and 2% (= 1 - $e^{-4.32}$) for trade flows. All knowledge flows sharply die out once they cross a state border and within an area of 500 miles with any additional 500 mile traveled distance to exert only an infinitesimal reduction on them. The California effect is discernible to all flows, but trade of goods flows. As expected, goods travel less in space than ideas, based either on patent trade flows or citation flows. Movements of inventors are reduced within the first 500 miles and but then slightly increase when distance gets large, larger than 2,000 miles. As California acts as an attractor of high quality scientists, inventors take these extra miles in order to enjoy California's research-facilitating environment. In addition, technological effort and structural proximities continue to shape knowledge flows. For the newly added channel of trade, a state receives 12% more knowledge from a state with the same technological effort as itself ($TechnologicalDistance = 0$) than from a state with dissimilar technological efforts ($TechnologicalDistance = 1$). In addition, two states with similar production specialization structure ($StructuralCloseness = 1$) exchange 5% more knowledge flows compared to states with completely dissimilar ($StructuralCloseness = 0$) technological sectors.

We further re-estimate the system allowing this time only for one type of closeness, that of geographic nearness, as the latter is the most commonly investigated form of closeness in the trade literature.[25] The geographic effect, when estimating the model without technological and structural closeness, becomes, as expected, somewhat stronger. For example, the effect of state border and nearby area further decrease knowledge flows by approximately 0.3%, while there is no change in the estimates of distance classes.[26] The geographic (border and distance) effects on trade flows we find are not very far away from those reported in the study of Wolf (2000), which examines "home bias" in trade flows in the US.

Rather than estimating the effect of geography on trade flows, Peri (2005, p. 317) borrows estimates of distance and border from the studies of Anderson and Van Wincoop (2003) and Feenstra (2003) to assess the geographic localization of citation to that of trade flows. As the author states, it is the first and single attempt, so far, in the literature that evidence from different strands of spillover literature is discussed. By comparing the effects of distance and border on citation and trade flows, the study concludes that border and distance reduce physical trade 4 to 5 times more than they reduce citation flows. By jointly estimating the effect of geographic characteristics on different channels of knowledge flows we find that on crossing state border, knowledge based on citation flows is 3 times bigger than knowledge based on trade flows. In addition, the geographic stretch of citation flows is 16 to 18 times larger than that of trade flows. Our (distance) effects are smaller compared to those reported in Peri (2005), but reasonable, if one considers that we investigate flows within a country and not across word regions as the aforementioned study does.

[25] A large volume of literature has documented the negative impact of geographic distance and borders on the flows of physical trade. In a seminal paper, McCallum (1995) finds that Canadian provinces trade up to 22 times more with each other than with US states initiating a strand of literature on international border effect. Subsequent studies, Anderson and Van Wincoop (2003) revisited the US-Canadian border effect with new micro-founded estimates, which considerably reduced the border effect. A parallel, smaller literature has documented that border effects also exist within a country, known as the domestic border effect or home bias. A number of studies, for example, for US (Wolf, 2000; Nitsch, 2000), EU (Nitsch, 2000), Germany (Wolf, 2009), and EU at industry-level (Chen, 2004) document significant intra-national border effects.

[26] The trade flows estimates of $State\,Border$, $Nearby\,States\,[500\,miles]$, $Distance\,[500-1,000\,miles]$, $Distance\,[1,000-1,500\,miles]$, $Distance\,[1,500-2000\,miles]$, and $Distance\,[2,000-2,500\,miles]$ are: -2.33 (0.008), -3.33 (0.015), -4.04 (0.007), -4.67 (0.005), -4.99 (0.011), and -5.05 (0.005), respectively. Number in parentheses are standard errors.

Taken together with our earlier results, local bias on the state level appears to be quite sturdy for all kind of knowledge flows. The very significant impact of distance on knowledge exchange across states of the US is at first sight quite surprising, especially for knowledge flows based on ideas (traded patent and citation flows), and puzzling: unlike goods and people (inventors), ideas are weightless, and distance cannot just proxy only transportation costs. Further, state border appears to equally confound trade of ideas (via traded patents) and trade of goods and certainly this effect cannot attributed to higher transportation costs when an idea, through the market of patents, crosses a border. Instead, distance and border could be seen as informational barriers, and therefore can serve as proxies for all types of informational frictions: agents within a state tend to know much more about each other and each other's business and technologies, either because of direct interactions between their citizens or because of better media coverage. In consequence, distance and border act as barriers to interaction, various micro-cultural affinities, and networking of economic agents. In our case, trade of patents requires as much as intensive information as trade of goods and this is reflected on the size of the border coefficient for these two types of trade, which is similar and quite high.

In addition to the general and well-documented finding of geographic confoundedness of flows, our results show that disembodied knowledge, generated from innovative ideas, which are patented and then traded and cited, are less geographically restricted and, therefore, their effective reach is beyond that of knowledge embodied in trade and inventors' flows as the latter channels involve movements of goods and people, respectively. Further, non-market channel knowledge flows are more far-reached than market-based flows, with knowledge flows via inventors' mobility to be the most geographically confounded among the market-based flows. Specifically, ideas based on citation flows are 10 to 12 (or 5 to 6, based on traded patent flows) times less restricted by a state border than knowledge flows based on inventors' mobility and 3 (or equal to, based on traded patent flows) times less restricted than knowledge flows based on trade of goods. In addition, the geographic scope of knowledge based on citation flows is about 16 to 18 (or 6 based on traded patent flows) times larger than knowledge flows based on inventors' mobility and 14 to 15 (or 5 based on traded patent flows) times bigger to knowledge flows based on trade of goods.

To get a better sense of the size of our coefficients, we compare our findings with evidence reported in the spillover literature. As there is no study to jointly estimate the effects of closeness on different channels of knowledge flows, we resort in studies which are at least conceptually closer to ours. The study of Mowery and Ziedonis (2001) is the most related one, as it examines the effects of geography (distance) on patent licenses and citation flows. Mowery and Ziedonis (2001) document that geographic distance matters more for university-generated knowledge flows that operate through the market of contracts (patent licenses) to university-generated knowledge flows that operate through non-market channels and that the spread of knowledge flows in space of the latter is about 3 times bigger than that of the former; a relationship which is also supported by our estimates. Contrary to some arguments stated in the literature (Audretsch and Stephan, 1996, p. 651), we too find that geographic proximity is more essential to the operation of market contracts compared to the operation of informal, non-market flows based on citation flows.[27]

We would further like to compare our findings with those from the patent-citation literature. Cross-study comparisons are not always easy due to different measures of distance, omission of border, and different level of analyses employed. Such differences create inevitably variation in the estimated effects of geography on knowledge flows. However, we can recover some effects that can be compared with ours. Note that we only can compare our citation estimates with those reported in the patent-citation literature as all studies in this strand consider solely one channel of knowledge diffusion, that of patent citations. We begin with the seminal work of Jaffe et al. (1993), which studies the degree of knowledge spillovers using patent citations across the states of the US The authors report a drop of 50% to 60% in the citation flows when they transcend a state's border. We find a drop on citation flows when crossing a state border in the range of 73% to 76%. Peri (2005) examines knowledge flows across world's regions and reports a reduction of 21% when knowledge flows cross a region border and a 3% drop for each 1,000 km traveled. We too find a reduction of 24% to 27% on crossing a state border and a 0.5% drop for every 500 miles traveled within

[27] Audretsch and Stephan (1996) studied interactions between university-based scientists and biotechnology firms based on disclosures in firms' initial public offering documents about academic researchers' roles in the firms.

the US. Further, our estimates are not widely apart also from estimates reported in a number of recent studies that use finer level of data disaggregation analysis and/or exclusively focus on knowledge flows across universities.[28] From the exposition of related evidence, it is reassuring to conclude that that size of our geography estimates is comparable to most of the aforementioned studies and reasonable.

The localization robustness of weightless ideas documented in our study, matches also findings from quite different lines of research. For example, studies in the financial trade literature, using 'gravity-like' models have investigated whether geographic distance imposes a hurdle on financial asset transactions, which are weightless compared to goods. In fact, Portes et al. (2001) and Portes and Rey (2005) examine the determinist of cross-border assets (corporate bonds equities and treasury bonds) and show that gravity model explains financial asset transactions a least as well as goods trade transactions and further document a very strong negative effect of distance on all asset flows (-0.826 to - 0.763).

Finally, the decay of knowledge flows due to geography resistance factors depicted in Figure 2 seem to match our results. The decay effect is stronger for market-based flows, especially on inventors' mobility and trade of goods and milder for citation flows. The California effect is present for all but goods trade.

Knowledge Flows from Top Innovator States

Next, we examine whether the geographic scope of flows from top innovator states is wider than the average state flows. To explore this aspect, we consider knowledge flows originating *only* from the top innovator states when estimating equation (9).

We select the top innovators to be states with high R&D spending (share to GDP) and states that produce, combined, more than half of the US total R&D activity. The states of California (CA), Massachusetts (MA), Michigan (MI), New Jersey (NJ), and New York, (NY), Texas (TX), Illinois (IL), Pennsylvania (PA), Maryland (MD), Washington (WA), and Ohio (OH) are among the top 10 states in R&D spending and account for about 70% of the total US R&D activity in our sample. Therefore, they may act as innovation leaders.

Table 3 depicts the results for the top 10 innovator states. Column (i) reports tri-variate SUR estimates, while column (ii) the four-variate SUR estimates. One reads the estimated coefficient in the same way as explained in the previous section.

Estimates show that top leaders' knowledge flows based, on citations and trade flows, are approximately 1.5 and 1.2, respectively, less geographic localized, both in terms of state border and distance, than average state flows, while for the rest two channels of knowledge flows the geographic scope of the leaders' flows is similar to the average state flows. As before, similarities in technological effort and technological specialization still play an important role - especially the latter, for knowledge diffusion via citation and inventors' mobility flows, while structural proximity rather decreases knowledge flows that operate via trade.

Our results confirm the broader reach of leaders' flows only for two out of four channels. The analysis of leaders' flows confirms that independent of their origin, i.e., whether they flow from an average or innovator state, knowledge flows based on citation flows are the least geographically confounded flows with the largest spatial reach.

[28] Maurseth and Verspagen (2002), for example, uses citations between European-granted patents across 122 european regions and finds that the effect of (log) distance ranges from -29% to -38%, while the border effect varies from 53% to 56%. Further, Singh and Marx (2011) studies the geographic reach of knowledge spillovers across US cities and metropolitan areas and concludes that neighboring states increase knowledge interactions by 35.6% compared to states that are not. Lastly, Belenzon and Schankerman (2011) examine 184 research-oriented universities in the United States and find the effect of distance on university patent citations flows ranges from -16.5% to -32%.

Table 3: Determinants of Knowledge Flows in the US

	Tri-variate SUR Estimates[(I)]			**Four-variate SUR Estimates**[(II)]			
	Patent Flows[a]	*CitationFlows*[b]	*InventorFlows*[c]	*Patent Flows*[a]	*CitationFlows*[b]	*InventorFlows*[c]	*GoodsFlows*[d]
		(v)			(vi)		
State Border	-2.42***	-0.67***	-3.53***	-2.34***	-.0537***	-3.55***	-1.27***
	(0.003)	(0.000)	(0.008)	(0.013)	(0.000)	(0.0159)	(0.016)
Nearby States $[500\,miles]$	-2.90***	-0.87***	-4.11***	-2.85***	-0.07***	-4.1***	-2.31***
	(0.004)	(0.000)	(0.011)	(0.016)	(0.000)	(0.0265)	(0.027)
Distance $[500 - 1000\,miles]$	-2.99***	-0.91***	-4.32***	-2.82***	-0.751***	-4.31***	-2.89***
	(0.003)	(0.000)	(0.008)	(0.009)	(0.000)	(0.0167)	(0.017)
Distance $[1000 - 1500\,miles]$	-3.09***	-.096***	-4.42***	-2.95***	-0.78***	-4.43***	-3.36***
	(0.004)	(0.000)	(0.007)	(0.019)	(0.000)	(0.0137)	(0.014)
Distance $[1500 - 2000\,miles]$	-3.10***	-1.05***	-4.50***	-3.00***	-0. 876***	-4.54***	-3.62***
	(0.004)	(0.000)	(0.009)	(0.016)	(0.0004)	(0.0207)	(0.0207)
Distance $[2000 - 2500\,miles]$	-3.10***	-0.74***	-4.49***	-2.94***	-0..636***	-4.45***	-3.73***
	(0.003)	(0.000)	(0.005)	(0.0096)	(0.0003)	(0.011)	(0.011)
TechnologicalDistance	-0.03***	0.01***	-0.20***	-0.05***	.0871***	0.0891***	-0.10***
	(0.004)	(0.000)	(0.008)	(0.016)	(0.0005)	(0.0189)	(0.0189)
StructuralCloseness	0.53***	0.63***	-0.34***	0.76***	1.28***	.439***	-.0957***
	(0.011)	(0.000)	(0.025)	(0.042)	(0.0010)	(0.0529)	(0.0529)
Observations	32,256	32,256	32,256	1,341	1,341	1,341	1,341
$\chi^2(30)$		3980			4040		

All regressions include time dummies and origin and destination state fixed effects. Standard errors reported in parentheses; Coefficients of *constant* term are omitted; *State Border* is dummy for adjacency and is 1 if states share common border and 0 otherwise; *Nearby States* $[500\,miles]$ is a dummy for geographic nearness and is 1 when states are located nearby, within 500 miles and do not share common state border, and 0 otherwise; *Distance* $[500 - 1,000\,miles]$, *Distance* $[1,000 - 1,500\,miles]$, *Distance* $[1,500 - 2,000\,miles]$, and *Distance* $[2,000 - 2,500\,miles]$ are dummies for distance classes of 500 mile length and take the value of 1 if states are located between 500 to 1,000, 1,000 to 1,500, 1,500 to 2,000, and 2,000 to 2,500 miles, respectively and 0 otherwise; *TechnologicalCloseness* is the degree of proximity of the technological effort (*R&D/Scientists*) of states; *StructuralCloseness* is the degree of technological similarity of production sectors of states; (***): significance at 1% level.

[I] : Specification (I) reports negative binomial estimates of tri-variate SUR (column i).

[II] : Specification (II) reports estimates of four-variate mix count and continuous SUR (column ii) for the years 1993, 1995, and 2002.

[a] *Patent Flows* are traded patent flows originating for the 10 most innovative states.

[b] *Citation Flows* are citation flows originating for the 10 most innovative states.

[c] *Inventor Flows* are inventor flows originating for the 10 most innovative states.

[d] *Goods Flows* are trade (imports) of goods flows originating for the 10 most innovative states.

Robustness

To sharp the robustness of our results, we have performed several checks. First, we excluded the very distant states with the most zeros, Alaska and Hawaii. The exclusion of Alaska and Hawaii barely changes the results. Then, we excluded California, which in terms of innovation performance could act as an outlier.[29] Results, available upon request, mildly change, but overall conclusions drawn from Table 2 hold. A notable difference is that, the (long) distance effect, due to California, disappears. Second, we considered alternative definitions of traded patent flows, for example, 'first flow' and 'last flow', instead of (all) traded patent flows and re-estimate the SUR model. Results are similar to the ones discussed. Third, we classify citations of patents into two groups: citations of traded patents and citations of non-traded patents and estimate accordingly two tri-variate SUR models. All previously drawn conclusions hold. As expected, there is some difference on the effect of geography on these two flows. Citations of traded patents are found to be about 1.3 times more geographically localized compared to citations of non-traded patents. Lastly, we estimated knowledge flows for two sub-periods (1986-1996 and 1997-2006) in order to examine whether the importance for any type of closeness changes over time. The estimated effects of all types of closeness remain virtually the same. Overall, results do not change in any significant way across different specifications, sub-samples and alternative proxies of knowledge flows.[30]

*4.2. **Do Knowledge Flows Contribute to Innovation Production?***

We have established that knowledge flows across states depend on several characteristics of closeness. The existence of these flows, however, does not necessarily support the existence of externalities of knowledge on innovation. Available knowledge originating in other regions may bring, along with new ideas, a reduction in innovation possibilities thus generating a zero or even negative net effect on the productivity of researchers in innovation. Therefore, no clear prior exists on the sign and magnitude of innovation elasticities in the innovation-generating equation (4).

The second task of this paper is to assess the effect of external available knowledge on state's innovation activity. The flow-weighted external available R&D stocks, when estimating equation (4), are constructed in two ways: either by using the estimated $\hat{\phi}^f_{ijt}$ from the gravity-like equation (5) and therefore susceptible to any criticism addressed to the model, or by using the actual raw flows, ϕ^f_{ijt}, which do not rely too much on the modeling of the knowledge flows across states. The latter, are further standardized as following: $type\,of\,flow_{ijt}/Q_{jt}/type\,of\,flow_{iit}/Q_{it}$. The numerator represents flows between origin and destination state divided by the amount of innovation (proxied by number of patents) in the origin state, Q_j. It is, in other words, the relative share of knowledge that is exchanged compared to knowledge produced in the origin state. The denominator captures the size of flows within the destination state divided by the amount of innovation produced in the that state. The overall ratio denotes the relative importance of external knowledge flows to in-state knowledge, accounting for the innovation activities of the two exchanging states.

The dependent variable of equation (4) is the innovation output Q, which is the count of patents granted to a state in year t weighted for the patent citations, taking into account the grant year and the technology field of the patent. More specifically, every patent is assigned to an issued year and technology field. We have 14 years and 37 technology groups; therefore, each patent is classified in one out of 14x37=518 groups. Each patent in every group is then weighted by the number of citations it has in the group's distribution. The weighting scheme is $w_1 = 0.1$, if citation belongs to the 1st quintile, $w_2 = 0.2$ for the 2nd, $w_3 = 0.3$ for the 3rd, and $w_4 = 0.4$ for the fourth. We then sum these values up for every state at year t and get our weighted measure of innovation output. To estimate equation (4), we use OLS controlling for time effects.

Table 4, below, reports estimated elasticities of state's own R&D stock and external available to a state R&D stock, gained through the four aforementioned channels using the actual values of flows.[31] Columns

[29] See Table A.1 in the Appendix for descriptive statistics on states' profiles.

[30] On a separate note and due to lack of finer level data, we cannot study these four channels of knowledge flows jointly together at the sectoral level. Ideally, we also like to consider foreign direct invest (FDI) flows as an additional channel of (disembodied) knowledge but lack of inter-state bilateral data prevented us.

[31] We report results from raw flow-weighted R&D stocks. Results based on estimated flow-weighted R&D stocks are quite similar.

(i) and (ii) report, along with state's own R&D stock, external accessible flow-weighted R&D stocks originating from all states, while columns (iii) and (iv) consider external accessible flow-weighted R&D stocks originating only from top 10 innovator states in the US. In other words, we consider them as the only source of relevant flows in constructing A_{ij}^{α} in equation. This allows us to minimize potential endogeneity in estimating the coefficient μ of A_{ij}^{α}. In fact, columns (iii) and (iv) do not include the top 10 states in the regressions as receivers of R&D spillovers, but rather include only the remaining states. External accessible R&D stock in columns (iii) and (iv) are thus measured as A_{ij}^{α}=$\sum_{j \in Top\,10}(\phi_{ij}A_{ij})$.

Table 4: Elasticities of Innovation Function

	[direct measures of ϕ_{ij} to construct flows (weights)]			
	Flows from All States[a]		*Flows from Top 10 States*[b]	
	(i)	(ii)	(iii)	(iv)
$lnR\&D_{own}$	0.57***	0.43***	0.51***	0.23**
	(0.031)	(0.086)	(0.041)	(0.106)
$lnR\&D_{cites}$	0.29***	0.39***	0.23***	0.50***
	(0.046)	(0.102)	(0.050)	(0.102)
$lnR\&D_{patents}$	0.21***	0.12**	0.18***	0.12**
	(0.023)	(0.049)	(0.024)	(0.049)
$lnR\&D_{inventor}$	-0.03	-0.01	0.01	0.05
	(0.042)	(0.090)	(0.040)	(0.082)
$lnR\&D_{trade}$		0.19*		0.18**
		(0.110)		(0.086)
Observations	612	136	446	100
R-squared	0.88	0.88	0.81	0.83

All OLS regressions include time dummies. Robust standard errors reported in parentheses; Coefficients of *constant* term are omitted; $lnR\&D_{own}$ is state's own R&D stock; $lnR\&D_{cites}$ is external available to a state weighted by citation flows R&D stock; $lnR\&D_{patents}$ is external available to a state weighted by traded patent flows R&D stock; $lnR\&D_{inventors}$ is external available to a state weighted by inventors' mobility flows R&D stock; $lnR\&D_{trade}$ is external available to a state weighted by trade flows R&D stock; (***), (**), and (*): significance at 1%, 5%, and 10% level, respectively.

[a] All states were including as senders (origin) of knowledge flows. All states were including as receivers (destination) of knowledge flows.

[b] Only the top 10 innovators were including as senders (origin) of knowledge flows. Only the remaining 38 states were including as receivers (destination) of knowledge flows. Top 10 innovators are: California (CA), Massachusetts (MA), Michigan (MI), New Jersey (NJ), and New York, (NY), Texas (TX), Illinois (IL), Pennsylvania (PA), Maryland (MD), Washington (WA), and Ohio (OH).

Despite of potential worsening of the endogeneity problem when external accessible R&D stock originates form all states, estimates are pretty similar across different specifications. Results strongly support that state's own ($lnR\&D_{own}$), as well as, external accessible R&D stocks are significant contributors to a state's innovation production.

More specifically, we find that 1% increase of state's own R&D, is associated with an increases in the production of innovation by, approximately, 0.23% to 0.57%. Others' states R&D effort positively contributes to a state's innovation production by 0.23% to 0.50% via the channel of citation flows, by 0.12% till 0.21% via the channel of traded patent flows, and 0.18% to 0.19% via the channel of trade of goods, whereas a mixed picture emerges for the channel of inventor's mobility flows which is negatively associated with state's innovation when all states flows are considered and positively when flows originate from the top innovator states. In both cases, however, estimates are statistically insignificant and close to zero. It appears that government policies of a typical state to either apply strict regulations on inventors' mobility (Matt et al., 2009) or simply to employ less people in R&D could have a zero or even negative effect on productivity of researchers in innovation.

Summing up, we find that knowledge flows are relevant to a state's innovation production as external accessible R&D gained through different flows has a strong positive effect on a state's innovation activity, which is as large, for some cases, as that of state's own R&D stock. Second, the effective reach of disembodied knowledge flows, exemplified via citation and traded patent flows, is larger than that of embodied knowledge in trade of goods and inventors' mobility. Such finding corroborates with theoretical studies of endogenous growth (Grossman and Helpman, 1991; Rivera-Batiz and Romer, 1991) which emphasize the important consequences of disembodied knowledge flows for technology transfer and economic growth. Lastly, available flow-weighed R&D that reach a state via market channels has smaller effect on state's patent production than flow-weighted R&D that diffuses without market-based mechanisms. This finding is consistent with conclusions drawn from Tables **??** in the previous section, where formal, market-based flows are found to be more localized and less far-reached in space compared to informal, non-market flows, as the former require movements of goods or people or some degree of geographic proximity between the inventor (seller) and the buyer of a patent.

As a further check, we run all regressions in Table 4 lagging all variables on the right-hand side by one period to overcome potential immediate feedback effect; results did not change in any significant way.

Overall, our estimates of own R&D elasticity (23% - 57%) are in the vicinity of existing estimates reported in the international spillover literature, and in particular in the studies of Peri (2005) (60%-80%), Branstetter (2001) (72%), Pakes and Griliches (1980) (61%), Bottazzi and Peri (2007) (78%), and in several other studies. Similarly, our estimates of external accessible R&D gained mainly through citations (23% - 50%) are also close to what the literature reports, Peri (2005) (40%-50%) and Bottazzi and Peri (2007) (55%).

A caveat that applies to all studies that estimate the importance of external available ideas to the innovation production of a region, including ours too, is that one does not know for sure how much of this available external knowledge is actually useable to the destination state. Jaffe et al. (1998) make a point to the case: In studying NASA's patents they find that two thirds of citations are associated with knowledge spillovers, and a survey of patentees by Jaffe et al. (2000) reports that half of citations are or might be related to knowledge spillovers. This issue deserves some further inquiry.

5. Conclusion

This paper offers novel insights in knowledge diffusion across US states and its consequences for innovation activity. We use a single framework to learn more about the relative scope and intensity of four different channels of knowledge flows that operate via (i) trade of patents, (ii) citations of patents, (iii) inventors' mobility, and (iv) trade of goods, which have been analyzed separately from different avenues of the knowledge spillover literature. To jointly study these flows, we develop novel econometric techniques appropriate to the nature of the data.

Using recently developed detailed data for the states of the US, our findings support that geographic proximity, in terms of distance and contingency, matters for the spread of knowledge as it has been massively documented in the literature. Overall, our findings confirm that disembodied knowledge, generated from innovative ideas, which are patented and then traded and cited, are less geographically restricted and, therefore, their effective reach is beyond that of knowledge embodied in trade or inventors' flows as the latter involve movements of goods and people, respectively. Further, non-market channel knowledge flows are the more far-reached than market-based flows, with knowledge flows via inventors' mobility to be the most geographically confounded among the market-based flows. Finally, with respect to other types of closeness, technological effort proximity of states and production structure similarities further enhance knowledge interactions across states. Our results remain robust for sub-samples and alternative specifications.

Our findings shed some light on the effective reach of different channels of knowledge flows, which are generated in different markets (and non-markets), but examined within a common framework. The implications of our findings for the theoretical trade-growth literature are potentially relevant. Although theoretical studies of this field (Grossman and Helpman, 1991; Rivera-Batiz and Romer, 1991) emphasize important consequences of disembodied knowledge flows for technology transfer and economic growth, there has been little effort, on the empirical side, to thoroughly explore this issue. Along with other important studies, this paper makes an effort towards this direction and confirms the important geographic scope of disembodied knowledge flows. We further find that knowledge flows, especially disembodied knowledge flows, are relevant to a state's innovation production as external accessible R&D gained through these flows has a strong positive effect on a state's innovation activity, as large, for some cases, as that of state's own R&D stock.

The next research challenge is to further explore the causes of local at the state level bias. Adding explicit measures of potential inform and formal barriers to the model (gravity) specification provides a promising avenue. Learning more about the causes of bias is important to assess welfare implications of knowledge flows. Another issue deserving some inquiry is to evaluate the share of external available knowledge which indeed becomes actual knowledge and contributes to the production of new ideas.

References

Abramovitz, M., 1986. Catching up, forging ahead, and falling behind. Journal of Economic History 46 (2), 385–406.

Aghion, P., Howitt, P., 1997. Endogenous Growth Theory. MIT Press, Massachusetts, USA.

Agrawal, A., Cockburn, I., McHale, J., 2006. Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships. Journal of Economic Geography 6(5), 571–591.

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

Anderson, J., Van Wincoop, E., 2003. Gravity with gravitas: A solution to the border puzzle. American Economic Review 93(1), 170–192.

Arrow, K., 1962. The rate and direction of inventive activity: Economic and social factors. In: Nelson, R. (Ed.), Economic Welfare and the Allocation of Resources for Invention. Princeton University Press, Princeton NJ, pp. 609–625.

Audretsch, D., Stephan, P., 1996. Company-scientist locational links; the case of biotechnology. American Economic Review 86(3), 641–652.

Autor, D. H., David Dorn, G. H., 2013. The geography of trade and technology shocks in the united states. American Economic Review Papers and Proceedings, forthcoming.

Belenzon, S., Schankerman, M., 2011. Spreading the word: Geography, policy and knowledge spillovers. CEPR Discussion Paper No. 8002, Forthcoming in Review of Economics and Statistics.

Benhabib, J., Spiegel, M. M., Oct. 1994. The role of human capital in economic development: Evidence from aggregate cross-country data. Journal of Monetary Economics 34(2) (2), 143–173.

Blomstrom, L., Kokko, A., 1998. Multinational corporations and spillovers. Journal of Economic Surveys 12(3), 247–277.

Bloom, N., Griffith, R., Reenen, J. V., 2002. Do r&d tax credits work? evidence from a panel of oecd countries 1979-1997. Journal of Public Economics 85(1), 1–31.

Bode, E., 2004. The spatial pattern of localized r&d spillovers: an empirical investigation for germany. Journal of Economic Geography 4(1), 43–64.

Bottazzi, L., Peri, G., 2003. Innovation and spillovers in regions: Evidence from european patent data. European Economic Review 47(4), 687–717.

Bottazzi, L., Peri, G., 2007. The international dynamics of r&d and innovation in the long run and in the short run. Economic Journal 117(518), 486–511.

Branstetter, L. G., 2001. Are knowledgespillovers international or intranational in scope?: Microeconometric evidence from the u.s. and japan. Journal of International Economics 53(1), 53–79.

Cameron, G., Proudman, J., Redding, S., 2005. Technological convergence, r&d, trade and productivity growth. European Economic Review 49(3), 775–809.

Chen, N., 2004. Intra-national versus international trade in the european union: Why do national borders matter? Journal of International Economics 63(1), 93–118.

Coe, D., Helpman, E., 1995. International R&D spillovers. European Economic Review 39(5), 859–887.

Cohen, W. M., Levinthal, D. A., 1989. Innovation and learning: The two faces of r&d. Economic Journal 99(397), 569–596.

Feenstra, R., 2003. Advanced International Trade: Theory and Evidence. Princeton, NJ: Princeton University Press.

Fischer, M. M., Griffith, D. A., 2008. Modeling spatial autocorrelation in spatial interaction data: An application to patent citation data in the european union. Journal of Regional Science 48(5), 969–989.

Furman, J. L., Stern, S., 2011. Climbing atop the shoulders of giants: The impact of institutions on cumulative research. American Economic Review 101(5), 1933–1963.

Griffith, R., Redding, S., van Reenen, J., 2004. Mapping the two faces of r&d: Productivity growth in a panel of oecd industries. Review of Economics and Statistics 86(4), 883–895.

Griliches, Z., 1992. The search for r&d spillovers. Scandinavian Journal of Economics 94 Supplement, 29–47.

Grossman, G., Helpman, E., 1991. Innovation and Growth in the Global Economy. MIT Press, Massachusetts, USA.

Guellec, D., van Pottelsberghe de la Potterie, B., 2004. From r&d to productivity growth: Do the institutional settings and the source of funds of r&d matter? Oxford Bulletin of Economics and Statistics 66(3), 353–378.

Hall, B., Jaffe, A., Trajtenberg, M., 2001. The nber patents citations data file: Lessons, insights and methodological tools. NBER Working Paper No. 8498.

Hall, B., Mairesse, J., 1995. Exploring the relationship between r&d and productivity in french manufacturing firms. Journal of Econometrics 65(1), 263–293.

Jaffe, A. B., 1986. Technological opportunity and spillovers of r&d: Evidence from firms' patents, profits, and market value. American Economic Review 76(5), 984–1001.

Jaffe, A. B., Fogarty, M. S., Banks, B. A., 1998. Evidence from patents and patent citations on the impact of nasa and other federal labs on commercial innovation. Journal of Industrial Economics 46(2), 183–205.

Jaffe, A. B., Trajtenberg, M., Fogarty, M. S., 2000. The meaning of patent citations: Report on the nber/case-western reserve survey of patentees. NBER Working Papers No. 7631.

Jaffe, A. B., Trajtenberg, M., Henderson, R., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. Quarterly Journal of Economics 108(3), 577–598.

Jones, C. I., 2005. Growth and ideas. In: Aghion, P., Durlauf, S. (Eds.), Handbook of Economic Growth, 1st Edition. Vol. 1. Elsevier, Ch. 16, pp. 1063–1111.

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

Keller, W., 2002b. Trade and the transmission of technology. Journal of Economic Growth 7(1), 5–24.

Keller, W., 2004. International technology diffusion. Journal of Economic Literature 42(3), 752–782.

Kim, J., Lee, S. J., Marschke, G., 2010. Inventor mobility and knowledge transmission in nanotechnology. Institute of Economic Research, Korea University in its series Discussion Paper Series No. 1004.

King, G., 1989. A seemingly unrelated poisson regression model. Sociological Methods and Research 17(3), 235–255.

Kneller, R., Stevens, P., 2006. Frontier technology and absorptive capacity: Evidence from oecd manufacturing industries. Oxford Bulletin of Economics and Statistics 68(1), 1–21.

Lai, R., Amour, A. D., Yu, A., Sun, Y., Torvik, V., Fleming, L., 2011. Disambiguation and co-authorship networks of the u.s. patent inventor database (1975-2010). http://hdl.handle.net/1902.1/15705 UNF:5:9kQaFvALs6qcuoy9Yd8uOw== V1 [Version].

Los, B., 2000. Technology and knowledge: from the firm to innovation systems. Edward Elgar, Cheltenham, Ch. The empirical performance of a new interindustry technology spillover measure.

Lucas, R., 1988. On the mechanics of economic development. Journal of Monetary Economics 22(1), 3–42.

Mancusi, M. L., 2008. International spillovers and absorptive capacity: A cross-country cross-sector analysis based on patents and citations. Journal of Internationa Economics 76(2), 155–165.

Matt, M., Strumsky, D., Fleming, L., 2009. Mobility, skills, and the michigan non-compete experiment. Management Science 55(6), 875–889.

Maurseth, P. B., Verspagen, B., 2002. Knowledge spillovers in europe: A patent citations analysis. The Scandinavian Journal of Economics 104(4), 531–545.

McCallum, J., 1995. National borders matter: Canada-u.s. regional trade patterns. American Economic Review 85(3), 615–623.

Mowery, D., Ziedonis, A., 2001. The geographic reach of market and non-market channels of technology transfer: Comparing citations and licences of university patents. NBER Working Paper No. 8568.

Nitsch, V., 2000. National borders and international trade: Evidence from the european union. Canadian Journal of Economics 33(4), 1091–1105.

Paci, R., Usai, S., 2009. Knowledge flows across the european regions. Annals of Regional Science 43(3), 669–690.

Pakes, A., Griliches, Z., 1980. Patents and r&d at the firm level: A first report. Economics Letters 5(4), 377–381.

Peri, G., 2005. Determinants of knowledge flows and their effect on innovation. Review of Economics and Statistics 87, 308–322.

Perkins, R., Neumayer, E., 2011. Transnational spatial dependencies in the geography of non-resident patent filings. Journal of Economic Geography 11(1), 37–60.

Polanyi, M., 1966. The Tacit Dimension. Univeristy of Chicago Press.

Portes, R., Rey, H., 2005. The determinants of cross-border equity flows. Journal of International Economics 65(2), 269–296.

Portes, R., Rey, H., Oh, Y., 2001. Information and capital flows: The determinants of transcations in financial assets. European Economic Review 45(4-6), 783–796.

Redding, S., 1996. The low-skill, low-quality trap: Strategic complementarities between human capital and r&d. The Economic Journal 106 (435), 458–470.

Rivera-Batiz, L., Romer, P., 1991. Economic integration and endogenous growth. Quarterly Journal of Economics 106(2), 227–244.

Romer, P., 1986. Increasing returns and long-run growth. Journal of Political Economy 94(5), 1000–1037.

Romer, P. M., 1989. Human capital and growth: Theory and evidence. National Bureau of Economic Research Working Paper Series No. 3173.

Serrano, C., 2011. Estimating the gains from trade in the market for innovation: Evidence from the transfer of patents. NBER Working Paper No. 17304.

Serrano, C. J., 2006. The market for intellectual property: Evidence from the transfer of patents. Ph.d., University of Minessota, dAI/A 67-06 Oct. 2006.

Serrano, C. J., 2010. The dynamics of the transfer and renewal of patents. RAND Journal of Economics 41(4), 686–708.

Singh, J., Marx, M., 2011. The geographic scope of knowledge spillovers: Spatial proximity, political borders and non-compete enforcement policy. INSEAD Working Papers Collection Issue 44.

Terlecky, N., 1980. New Developments in Productivity Measurement. University of Chicago Press, Ch. Direct and Indirect Effects of Industrial Research and Development on Productivity Growth of Industries.

Thompson, P., 2006. Patent citations and the geography of knowledge spillovers: Evidence from inventor- and examiner-added citations. The Review of Economics and Statistics 88(2), 383–388.

Tsionas, E., 1999. Bayesian analysis of the multivariate poisson distribution. Communications in Statistics (Theory and Methods) 28(2), 431–451.

Winkelmann, R., 2000. Seemingly unrelated negative binomial regression. Oxford Bulletin of Economics and Statistics 62(4), 553–560.

Wolf, H. C., 2000. Intra-national home bias in trade. Review of Economics and Statistics 82(4), 555–563.

Wolf, H. C., Nadiri, M. I., 1987. Spillover effects, linkage structure, technical progress and r&d. New York University C.V. Starr Center Research Report 87/43.

Wolf, N., 2009. Was germany ever united? evidence from intra- and international trade 1885-1933. Journal of Economic History 69(3), 846–881.

Zellner, A., 1962. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. Journal of American Statistical Association 57(298), 348–368.

Appendix

Table A.1: Summary Statistics per State

State	Traded Patents		Citations (all patents)		Citations (traded patents)		Citations (non-traded patents)		Inventor Mobility		Traded Goods		Scientists		R&D spending		R&D share	
	mean	Std.	mean	Std.	mean	Std.	mean	Std.	mean	Std.	mean	Std.	mean	Std.	mean	Std	mean	Std
AK	9.29	4.58	301.21	192.51	44.86	32.83	174.50	135.53	9.21	5.69	5640.49	1866.15	1.30	0.09	0.20	0.06	0.01	0.00
AL	120.21	50.39	2699.21	1709.94	489.50	309.06	1517.43	1198.78	151.43	81.36	136742.80	14813.17	7.11	1.19	2.15	0.33	0.02	0.00
AR	37.21	19.01	1445.50	878.14	204.14	127.38	851.50	621.82	59.29	37.16	102072.50	4198.25	3.13	0.43	0.40	0.06	0.01	0.00
AZ	195.36	73.67	8839.57	5112.21	1542.21	915.69	5077.36	3783.73	572.57	319.85	106381.90	15401.81	7.74	1.20	3.07	1.05	0.02	0.01
CA	3295.79	1395.30	196680.30	128945.40	35131.07	23653.45	123552.80	95628.45	7812.86	4503.33	609102.80	63047.66	84.56	10.10	48.14	7.72	0.04	0.00
CO	388.50	147.68	14092.21	8649.10	3142.93	1979.45	7728.21	5781.69	815.57	445.13	94217.51	9185.09	13.12	1.66	4.14	0.78	0.03	0.00
CT	439.50	101.38	23692.36	13364.13	3683.14	2141.15	13237.64	9464.81	854.50	427.56	110648.00	5636.09	10.26	1.15	5.12	1.68	0.03	0.01
DE	616.93	236.85	18919.36	10840.57	2619.86	1581.38	11021.07	8092.58	213.36	117.69	32635.48	2642.04	3.92	0.44	1.50	0.40	0.04	0.01
FL	537.21	186.50	24404.07	15604.85	4395.64	2913.54	13805.07	10788.61	938.29	472.28	240354.50	21688.26	17.48	2.12	4.81	0.61	0.01	0.00
GA	279.71	117.16	10800.36	6667.31	2034.29	1322.03	6283.50	4655.46	562.21	305.26	311019.70	5926.84	12.18	1.71	2.81	0.73	0.01	0.00
HI	8.93	3.87	595.07	348.04	72.21	47.50	350.43	258.78	22.93	11.49	6569.39	1004.96	2.79	0.28	0.40	0.09	0.01	0.00
IA	118.14	64.38	4086.14	2350.77	722.86	458.68	2351.71	1665.82	237.00	129.66	124773.90	16114.45	4.92	0.29	1.22	0.21	0.01	0.00
ID	120.00	162.19	14416.86	10977.85	1960.14	1502.96	10478.43	8464.85	250.79	129.49	28709.28	3891.09	2.50	0.36	1.01	0.31	0.03	0.01
IL	843.00	259.04	49019.29	27975.22	6536.50	3793.54	28818.14	20411.10	1482.86	755.64	477459.10	28034.33	23.69	1.45	9.70	1.56	0.02	0.00
IN	230.36	65.77	10975.86	6300.02	1797.07	1050.15	6065.86	4398.97	596.79	342.94	295413.40	48487.71	9.66	0.85	3.61	0.78	0.02	0.00
KS	84.50	38.41	2946.29	1789.39	454.29	278.24	1719.71	1316.83	145.71	74.71	108231.60	8599.44	4.31	0.37	1.40	0.58	0.02	0.01
KY	56.43	23.31	3514.21	2555.13	523.57	428.63	2016.50	1755.54	158.57	85.13	188553.80	40155.48	4.91	0.52	0.81	0.23	0.01	0.00
LA	84.43	31.58	3185.29	1716.89	536.57	285.33	1627.71	1140.31	153.57	80.89	128699.20	12328.39	5.94	0.23	0.66	0.16	0.00	0.00
MA	864.86	324.21	44786.93	27315.71	8679.50	5516.57	25834.00	19059.64	2092.21	1161.44	183286.00	9336.05	29.14	3.81	13.11	2.07	0.05	0.00
MD	269.64	111.28	10666.14	6675.59	2195.29	1492.47	5950.21	4453.52	753.57	421.15	145223.30	17471.29	25.26	3.30	9.26	1.88	0.05	0.01
ME	27.14	23.28	1133.79	773.91	171.43	127.52	696.07	563.06	54.21	29.66	31311.10	4395.33	2.46	0.13	0.28	0.11	0.01	0.00
MI	730.79	413.72	32650.14	19029.46	5091.64	3046.94	18633.57	13986.93	1476.43	785.61	364852.50	45548.45	17.65	1.54	15.26	2.35	0.05	0.01
MN	592.79	267.34	38569.29	25742.93	7231.36	4862.83	21489.29	17651.55	1040.93	603.40	164072.20	20179.74	11.42	1.36	4.40	1.06	0.02	0.00
MO	217.36	64.73	8522.50	4750.13	1314.43	767.64	4722.79	3273.00	442.21	244.81	218572.90	15572.51	9.80	0.53	2.43	0.41	0.01	0.00
MS	50.00	31.41	1318.64	1089.27	231.21	196.45	821.64	834.21	60.07	31.32	89770.17	10231.12	3.38	0.23	0.55	0.28	0.01	0.00
MT	24.07	13.76	1091.93	787.22	179.07	118.23	638.79	585.88	40.14	22.14	16141.39	1686.00	1.98	0.19	0.18	0.06	0.01	0.00
NC	280.21	117.20	12180.86	7487.43	2068.93	1292.47	7421.29	5539.45	779.50	433.22	280150.30	21377.07	17.31	2.47	4.79	1.23	0.02	0.00
ND	16.50	9.10	409.79	222.36	55.43	32.26	218.14	157.51	22.29	12.43	18522.55	3788.67	1.63	0.46	0.20	0.12	0.01	0.01
NE	67.00	28.11	1804.86	1030.36	420.29	280.98	923.57	659.56	69.14	35.26	65771.58	7101.39	2.97	0.11	0.48	0.16	0.01	0.00
NH	126.21	96.43	4715.50	3061.73	965.79	712.50	2697.71	2105.67	314.07	171.72	38112.65	7528.93	2.66	0.36	1.11	0.46	0.03	0.01
NJ	824.36	359.10	47536.29	28242.18	7960.71	5003.76	27719.14	20139.22	1930.36	1080.70	340518.00	6643.18	23.38	1.92	11.59	1.18	0.03	0.00
NM	81.43	134.75	2545.14	1696.02	557.36	393.64	1522.57	1204.32	180.21	97.95	28924.74	1935.51	8.44	0.79	3.75	0.74	0.07	0.01
NV	230.50	231.23	5550.79	4081.78	1223.93	962.13	3390.21	2860.40	115.14	67.16	47653.47	9687.69	2.11	0.33	0.46	0.14	0.01	0.00
NY	1106.86	630.08	81306.71	49067.34	10507.21	6598.90	51859.86	37962.09	2157.93	1114.96	400433.90	17918.45	45.88	2.40	12.86	0.89	0.02	0.00
OH	625.64	196.06	31766.93	18166.71	4685.64	2728.97	17991.79	13159.96	1373.43	719.66	463485.00	57176.91	21.68	1.76	7.39	0.65	0.02	0.00
OK	91.21	40.04	4938.71	2581.55	801.00	449.10	2587.57	1782.38	190.71	91.15	88464.96	8744.00	4.96	0.24	0.67	0.10	0.01	0.00
OR	174.93	78.53	8521.21	5137.34	1646.71	1095.91	4751.50	3528.42	469.43	252.91	98741.59	15711.60	8.12	1.13	2.14	1.00	0.02	0.01
PA	685.57	175.40	28791.21	16446.57	4741.50	2871.76	15676.29	11434.38	1577.29	848.06	387158.10	33121.79	28.08	2.48	9.58	0.92	0.02	0.00
RI	72.57	32.57	2199.21	1284.55	479.57	271.47	1171.29	877.18	149.29	78.33	25858.19	2106.39	2.87	0.32	1.35	0.48	0.04	0.01
SC	82.86	45.09	4174.21	2426.20	573.71	342.91	2315.14	1703.40	235.07	114.04	150595.30	20641.18	5.50	0.48	1.22	0.37	0.01	0.00
SD	8.57	6.80	560.14	370.71	67.79	51.71	361.64	280.71	15.21	10.08	19459.47	3205.89	1.14	0.06	0.09	0.04	0.00	0.00
TN	186.71	58.57	6509.57	4050.97	1241.43	788.76	3652.00	2867.53	363.43	186.61	274214.60	60689.53	9.61	0.69	2.18	0.57	0.01	0.00
TX	842.07	234.75	56453.79	35814.61	8575.36	5546.29	35611.29	27248.10	2165.21	1222.22	494561.20	51263.84	34.72	3.51	11.39	2.42	0.02	0.00
UT	159.64	95.54	7831.50	5084.26	1633.50	1028.50	4435.71	3524.97	316.36	167.10	56557.52	8069.14	5.33	0.45	1.34	0.29	0.02	0.00
VA	359.79	70.78	12311.29	7833.11	2078.29	1384.19	7159.79	5547.17	490.43	255.29	193182.50	9789.51	19.47	2.85	5.35	1.60	0.02	0.00
VT	19.21	11.49	817.86	495.44	131.71	84.10	478.86	355.21	117.36	59.44	19019.74	4182.37	1.91	0.20	0.39	0.08	0.02	0.00
WA	348.86	167.55	20755.36	13573.83	3690.29	2432.55	13462.21	10442.88	788.43	436.74	137307.80	15197.25	15.55	2.21	8.67	2.01	0.04	0.00
WI	273.21	76.57	17288.93	11351.91	2431.93	1535.99	9861.29	8062.87	742.36	373.54	215328.20	27417.95	9.23	0.87	2.77	0.54	0.02	0.00
WV	15.79	13.00	605.07	369.38	129.64	104.45	291.00	210.66	55.93	28.97	51474.45	1367.96	2.28	0.19	0.44	0.07	0.01	0.00
WY	13.21	7.21	378.21	251.12	45.79	30.40	225.93	184.78	21.07	12.42	11545.84	1760.98	0.89	0.09	0.08	0.02	0.00	0.00

First column is states' two-letter abbreviation; *Traded Patents* and *Citations* (of all types of patents), and *InventorMobility* are occurrences; *Traded Goods* are in millions of constant (2000) US dollars, *Scientists* are in thousands, *R&D spending* is in millions of constant (2000) US dollars; and *R&D share* is R&D ratio to state-level GDP.