Exploring the spatially varying innovation capacity of the US counties in the framework of Griliches' knowledge production function: a Mixed GWR approach

Dongwoo Kang and Sandy Dall'erba

REAL 15-T-7 August, 2015

Exploring the spatially varying innovation capacity of the US counties in the framework of Griliches' knowledge production function: a Mixed GWR approach

Dongwoo Kang

Korea Labor Institute, 370, Sicheong-daero, Sejong-si, 339-007, Korea. E-mail: dwkang1982@gmail.com

Sandy Dall'erba

Department of Agricultural and Consumer Economics and Regional Economics Applications Laboratory, University of Illinois at Urbana-Champaign, IL 61801, USA. E-mail: dallerba@illinois.edu

Grant: This study was supported by the National Science Foundation Grant (SMA-1158172). Any opinions, findings and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Abstract: Griliches' knowledge production function has been increasingly adopted at the regional level where location-specific conditions drive the spatial differences in knowledge creation dynamics. However the large majority of such studies relies on a traditional regression approach that assumes spatially homogenous marginal effects of knowledge input factors. This paper extends the authors' previous work (Kang and Dall'erba, 2015) to investigate the spatial heterogeneity in the marginal effects by using nonparametric local modeling approaches such as Geographically Weighted Regression (GWR) and Mixed GWR with two distinct samples of the US Metropolitan Statistical Area (MSA) and non-MSA counties. The results indicate a high degree of spatial heterogeneity in the marginal effects of the knowledge input variables, more especially for the local and distant spillovers of private knowledge measured across MSA counties. On the other hand, local academic knowledge spillovers are found to display spatially homogenous elasticities in both MSA and non-MSA counties. Our results highlight the strengths and weaknesses of each county's innovation capacity and suggest policy implications for regional innovation strategies.

Keywords: Knowledge production function, knowledge spillovers, spatial heterogeneity, Mixed Geographically Weighted Regression (MGWR)

JEL classifications: C21, O31, R11

1 Introduction

Compared to Griliches' (1979) original firm-focused knowledge production function, successive developments in the literature have recognized the critical role of localized knowledge spillovers (Audretsch and Feldman 2004). It has led to a surge of studies focusing on a regional approach of knowledge creation (Anselin et al. 1997; Audretsch and Feldman 1996; Bode 2004; Jaffe 1989). Shifting observation from the firm-level to the regional level does not lead to a mere aggregation of individual firm units. Instead, it implicitly assumes that innovative activities and local knowledge spillovers between individual firms are not negligible and prevalent within the spatial unit (Audretsch and Feldman 2004). Another motivation for using aggregated units is that the role of region-specific characteristics is crucial in knowledge creation. Each region has a specific institutional environment organized by frequent interactions and reciprocal trust between entrepreneurs, universities and government agencies within the region (Döring and Schnellenbach 2006). And since such an institutional environment plays an important role in intra-regional collective synergies and externalities (Harris 2011), we should expect the mechanisms that shape intra-regional knowledge creation to be as spatially heterogeneous across regions as institutional environments are.

However, spatial variations in the capacity to innovate are not limited to intra-regional activities only. Firms in one location benefit from knowledge created in distant sources (Rosenkopf and Almeida 2003) to the point where remote partners can be a firm's strongest partner for innovation (Asheim and Isaksen 2002; Gertler and Levitte 2005; Trippl et al. 2009). Yet, at the regional level, the ability to capitalize on such external sources of knowledge creation, i.e. the region's absorptive capacity (Cohen and Levinthal 1990), is still subject to local conditions such as the quality of its institutions, its geography and the level of agents' interactions within the region (Agrawal et al. 2010). As such, one should expect the absorptive capacity and the dynamics of knowledge creation in any region to be different from what they are in all the remaining regions.

Nevertheless, most empirical studies of the regional knowledge production function presume spatially homogenous marginal effects of the knowledge input factors (Anselin et al. 1997; Jaffe 1989; Ó hUallacháin and Leslie 2007; Ponds et al. 2010). Indeed, in the traditional regression approach, the coefficient associated with a variable corresponds to its average impact across the entire sample, which may mask a positive impact in some areas and a less positive or even negative one in other areas (McMillen and Redfearn 2010). In order to measure the potential presence of heterogeneity, we adopt a geographically weighted regression (GWR) approach to obtain local coefficients (Brunsdon et al. 1996; Brunsdon et al. 1998; Fotheringham et al. 2002). It should allow us to uncover each county's innovation strengths and weaknesses and suggest place-tailored innovation policies (Fritsch and Stephan 2005; Stough 2003).

However one should not expect the role of every knowledge input variable to significantly vary spatially. If that is the case, then the results of GWR could lead to inefficient and incorrect conclusions (Wei and Qi 2012). As such we examine the degree of spatial stationarity in the marginal effects of input variables and compare the result of GWR with those of Mixed Geographically Weighted Regression (MGWR) (Fotheringham et al. 2002). GWR or MGWR are not widely used in the knowledge production function literature. We attribute it to the fact that most of the attention has focused on modeling the mechanisms of spatial knowledge spillovers (for the overview on spatial modeling of knowledge spillovers, see Autant-Bernard 2012). In addition, the collinearity problem of GWR that makes this technique inapplicable to small sample of less than 400 observations (Bárcena et al. 2014; Paez et al. 2011; Wheeler and Tiefelsdorf 2005) is inconvenient for the states or Metropolitan Statistical Areas (MSA) level data usually used in regional knowledge production functions (Anselin et al. 2000; Ó hUallacháin and Leslie 2007; Peri 2005).

Given the aforementioned background, this paper explores the spatial heterogeneity in the mechanisms of regional knowledge creation across 3,109 continental US counties. As a smaller observation unit than states or MSA, counties allow us to provide more details about spatial variations in the US knowledge creation mechanism, as well as to adopt the GWR approach. Our large sample also helps us not to disregard non-metropolitan counties of which innovative capacity is admittedly lower than their metropolitan counterparts but certainly not negligible. The rest of the paper is composed as follows: Section 2 reviews the theoretical literature pinpointing the sources of spatial heterogeneity and lists the ways it has been modeled previously. Section 3 describes our knowledge production function and the relevant data. The calibration methods of GWR and MGWR are also explained in this section. The results and their interpretation appear in Section 4 while the last section closes the paper with some concluding remarks.

2 Literature review

2.1 Sources of spatial heterogeneity in knowledge creation and innovation

Firms, universities and government agencies are the main actors of knowledge creation and innovation. As such the individual capacity of these agents is an important determinant of knowledge creation in the region they belong to. Yet the region's characteristics constituted by interactions between the agents are also highly relevant to intra-regional innovation outputs (Döring and Schnellenbach 2006; Harris 2011). For instance, Marshall (1920) points out that socio-cultural and institutional assets such as collective identity and expertise develop gradually within an industrial district and its "industrial atmosphere" is critical to facilitating localized knowledge spillovers and the creation of innovation. The literature of collective learning (Camagni 1991; Lazaric and Lorenz 1998; Lorenz 1992) also emphasizes the role of regional "innovative milieus" in promoting regional learning and innovation. Local cultural and institutional rules and civic engagement facilitate collaborations between firms thereby contributing to stronger intra-regional trust, inter-firm networking (Keeble et al. 1999) and a greater innovative capacity at the regional level (Storper 1997).

Similar ideas are found in the literature of learning region (Asheim 1996; Florida 1995; Morgan 1997; Simmie 2011; Storper 1993). In the knowledge-based economy, global companies depend on their home regions' local knowledge assets and infrastructures which are subject to region-specific institutional and cultural norms (Florida 1995). As such, the regional learning process of generating and transferring knowledge is affected by local social capital, i.e. the institutional and cultural context of local networks, trust and conventions (Asheim 1996; Morgan 1997; Storper 1993). Studies on regional innovation system have also highlighted the critical role of institutional infrastructures on intra-regional innovative activities (Asheim and Isaksen 2002; Asheim and Gertler 2006; Cooke et al. 1997; Oughton et al. 2002). Since these institutional infrastructures have evolutionary properties and path dependence (Asheim and Gertler 2006; David 1994; Zysman 1994), their regional characteristics are different across regions (Simmie 2011). Therefore, heterogeneous region-specific conditions are a source of spatial heterogeneity in intra-regional knowledge creation.

In addition, heterogeneous region-specific conditions are related with the regional capacity of exploiting external knowledge sources. Firms can exploit benefits of knowledge spillovers not only from their local knowledge pool but also from distant external sources (Asheim and Isaksen 2002; Gertler and Levitte 2005; Rosenkopf and Almeida 2003; Trippl et al. 2009) and each region has its own capacity to absorb knowledge created elsewhere (Cohen and Levinthal 1990) as it depends on the unique combination of the ability of local individual agents (Mukherji and Silberman 2013). Verspagen and Schoenmakers (2004) demonstrate that this capacity is proportional to the existing stock of local knowledge while Döring and Schnellenbach (2006) and Agrawal et al. (2010) show that it is also depends on the locality's institutions, sectoral structure and historical similarity with its partners. As such heterogeneous regional characteristics lead to a spatially heterogeneous capacity to exploit external knowledge sources.

2.2 Modeling spatial heterogeneity in the empirical studies of knowledge production function

Various approaches have been used in the literature to model spatial heterogeneity in the regional innovation dynamics. The most common approach relies on capturing a set of control variables of which values differ by region. Since the contribution of Jaffe (1989), geographically-aggregated R&D expenditures and human capital are commonly used in regional knowledge production functions (Acs et al. 2002; Bode 2004; Parent and LeSage 2008). An index of specialization or diversity is also commonly found (Feldman and Audretsch 1999; Ó hUallacháin and Leslie 2007; Parent and LeSage 2008) and the sign of its associated estimates helps researchers concludes on whether Marshall-Arrow-Romer (Arrow 1962; Marshall 1920; Romer 1986) or Jacobian externalities (Jacobs 1969) drive regional innovation the most as both matter theoretically (Fung and Chow 2002; Glaeser et al. 1992; Henderson 2003; Jaffe 1986). In addition, Anselin et al. (1997) and Bode (2004) propose to control for the share of small firms since they have a comparative advantage in exploiting knowledge generated from university laboratories (Acs et al. 1994). On the other hand, large firms have a greater impact on the local labor market (Acs and Armington 2004), thus their contribution to agglomeration effects and consequently to knowledge creation could be greater than small firms. As such, the relative presence of large or small firms needs to be accounted for to shed some light on its net effect on regional knowledge creation.

The recent contributions of Mukherji and Silberman (2013) and Capello and Lenzi (2014) focus on the regional entrepreneurship culture and social capital respectively due to their role on the local absorptive capacity. In the absence of actual measurement of these regional characteristics, the level of entrepreneurship is proxied by the self-employment rate, the rate of firm birth and deaths and the share of employment in young firms. Following the innovative milieus theory (Camagni 1991; Lazaric and Lorenz 1998; Lorenz 1992), the level of trust, cooperation and collective actions within a region are used as a measure of social capital. In addition, several studies attempt to measure the level of intra-regional knowledge spillovers directly. Jaffe (1989) uses the geographical coincidence of university and commercial R&D laboratories within a US state, while Kang and Dall'erba (2015) rely on intra-regional patent citation flows at the county level.

An increasing number of spatial econometric studies incorporate interregional knowledge spillovers to deal with the spatial dependence and heterogeneity inherent in regional knowledge creation and innovation (Autant-Bernard 2012). In order to model spillovers capturing the diffusion of tacit knowledge due to face-to-face interactions, many studies rely on knowledge created within a distance cut-off or among physically contiguous neighboring regions (Anselin et al. 1997; Anselin et al. 2000; Autant-Bernard and LeSage 2011; Bode 2004). In addition, several studies examine the role of interregional knowledge spillovers taking place over long distance through channels such as technological proximity (Fischer et al. 2006; Maggioni et al. 2010; Maurseth and Verspagen 2002). Peri (2005) uses a matrix of interregional patent creation-citation flows in Western Europe and North America. Ponds et al. (2010) model the spillovers across Dutch regions based on research collaboration between universities and firms.

As an alternative way to model the spatial heterogeneity of the innovation activities, random and fixed effects are also introduced. For instance, in the framework of a spatial interaction model, Fischer et al. (2006) adopt heteroskedastic error terms assumed to reflect heterogeneous knowledge flows. Parent and LeSage (2008) choose to model spatially structured random effects assumed to reflect interregional knowledge spillovers in conjunction with non-spatially structured heteroskedastic variance terms. Mukherji and Silberman (2013) focus on a region's absorptive capacity that they model through destination fixed-effect coefficients in the frame of a spatial interaction model. The estimated regional absorptive capacity is then used to explain innovation productivity.

However, the aforementioned modeling approaches presume that the average marginal impacts of the knowledge input factors are spatially homogenous. Since spatially heterogeneous regional characteristics affecting regional innovation activities are hard to control perfectly, local marginal impacts of knowledge inputs would reflect the relevant influence of regional characteristics. Therefore, the assumption of homogeneous parameters over space in the empirical model of knowledge production function may be too restrictive, thereby leading to locally biased misspecification (McMillen and Redfearn 2010). As a result, this paper adopts GWR and MGWR approaches to explore spatial heterogeneity in the marginal effects of the knowledge input variables.

3 Model and data

3.1 Regional knowledge production function

The prevalent knowledge production function at the firm level is a Cobb-Douglas function as in Griliches (1979). It proposes the innovation output to be associated to the knowledge stock and human capital inputs (Audretsch and Feldman 2004). At the regional level, the conceptual framework expands the firm level approach by aggregating geographically the knowledge stock and human capital level (Acs et al. 2002; Bode 2004; Jaffe 1989). In addition, several region-specific conditions can be included in the regional model as described in the previous section. Among them, we select a set of variables that can be measured for our observation units, i.e. the 3,109 continental US counties, and build on recent efforts to separate clearly intra-regional spillovers (Kang and Dall'erba, 2015) from local spillovers (Anselin et al. 1997; Bode 2004) and long distance spillovers (Peri 2005; Ponds et al. 2010). Eq. (1) presents our empirical model of regional knowledge production of which variables have been chosen based on past literature (see section 2) and the authors' previous work (Kang and Dall'erba, 2015).

$$ln\boldsymbol{Patent_i} = \beta_0 + \beta_1 ln\boldsymbol{Private_i} + \beta_2 ln\boldsymbol{Univ_i} + \beta_3 ln\boldsymbol{Graduate_i} + \beta_4 ln\boldsymbol{Diversity_i} + \beta_5 ln\boldsymbol{Large_i} + \beta_6 ln\boldsymbol{Intra_i} + \beta_7 \ln \boldsymbol{Local.Private_i} + \beta_8 \ln \boldsymbol{Local.Univ_i} + \beta_9 ln\boldsymbol{Distant.Private_i} + \beta_{10} ln\boldsymbol{Distant.Univ_i} + \beta_{11} ln\boldsymbol{Size_i} + \varepsilon_i \quad (1)$$

The knowledge output is measured using the average of total patent applications over 2003-2005 by county (***Patent***). The patent data comes from the US Patent and Trade Office (USPTO 2010). The inventor's address is used to assign patent data to US counties of the 2000 US Census using the fractional county method as in Jaffe et al. (1993). For example, when N inventors apply for a patent together, it is assumed that 1/N faction of the patent is attributed to each inventor. Then each 1/N fractional patent is geocoded to the US counties corresponding to the address of the inventors. The regional knowledge stock in private sector (***Private***) is approximated by the discounted sum of the companies' R&D expenditures over 1995-2002 (measured in 2003 constant dollars using the Producer Price Index of the US Bureau of Labor Statistics). It is calculated using the perpetual inventory method on the lag polynomials of the R&D expenditures (Griliches 1992) coupled with a 15% annual depreciation rate. Data comes from Standard and Poor's COMPUSTAT database (Standard & Poor's 2011). The address of the companies is used for allocating the R&D expenditures across counties. We create the stock of knowledge in the academic sector (***Univ***) by the same method as ***Private.*** It is based on the academic R&D expenditures found in the NSF Survey of R&D expenditures at universities and colleges (National Center for Science and Engineering Statistics 2013).

In order to measure the regional level of human capital (***Graduate***), we rely on the share of Graduate or professional degree holders 25 years old and over. The data comes from the 2000 US Census. In addition, we control for several regional economic conditions. As a relative diversity index of the regional industrial structure (***Diversity***), we use the index of Duranton and Puga (2000) as in Eq. (2) where S_{ij} is the share of industry *j*'s employment in county *i* and S_j represents the share of national employment in industry *j*. The relevant data is based on the 13 industry system of the 2000 US Census.

$$Diversity_i = 1/\sum_j \left|s_{ij} - s_j\right| \quad (2)$$

The share of large firms with 500 employees or more (***Large***) is measured to evaluate the influence of the composition of firm sizes in the regional economy. Its data comes from the 2000 County Business Patterns. As a control variable for the differences in regional economic size, the total number of employment in a county (***Size***) is included in the estimation model. The data comes from the 2000 US Census.

For the modeling of intra- and inter-regional knowledge spillovers, we rely on the patent creation-citation flows as seen in the spatial interaction modeling literature (Fischer et al. 2006; Maggioni et al. 2010). By constructing a matrix of patent citation patterns across regions, we can track the interregional knowledge spillovers (Fischer et al. 2006). In order to make a matrix of the patent creation-citation flows across our 3,109 counties, we use the patent citation records over 1995-1999 from the NBER US Patent Citation Data File (Hall et al. 2001). The patent citation flows are constructed using the same geocoding process as for the dependent variable. For example, when O investors create patent A and the patent is cited by patent B generated by D inventors, we assume that there are (O×D) knowledge flows from patent A to patent B. In addition, each flow is expected to capture a fraction 1/(O×D) of the knowledge diffused from the counties of the cited patent inventors (origin counties) to the counties of the citing patent inventors (destination counties). Finally, we aggregate the fractional knowledge flows between counties to construct a (3,109×3,109) patent citation flow matrix. We call it the P matrix and the column-standardized P matrix is used to define the variables of interregional spillovers. However, we intentionally separate the interregional spillovers into two categories. Based on the US daily commuting patterns (Smallen 2004) and following Anselin et al. (1997), Acs et al. (2002) and Mukherji and Silberman (2013), we use 50 miles as the distance cut-offs delimiting local spillovers. Spillovers over this distance cut-off are qualified as long-distance spillovers. Both are defined further below. In order to assess the robustness of our results, the distance cut-off of 75 miles will be examined also.

As a result, the localized interregional knowledge spillovers from the private sector (***Local.Private***) are approximated by the weighted sum of the private R&D expenditures of neighboring counties within 50 miles (or 75 miles) as in Eq. (3) where P_{ij} is the (i^{th}, j^{th}) element of the P matrix and represents the citation flows from county *j* to county *i*. Using the column-standardized P matrix, the variable is interpreted as the aggregated knowledge diffused from external knowledge sources. With the same idea but a cut-off of 50 miles (or 75 miles) at least, the distant interregional knowledge spillovers from the private sector (***Distant.Private***) are defined as in Eq. (4). The localized and distant interregional spillovers from the academic sector (***Local.Univ*** and ***Distant.Univ***) are defined based on Eqs. (3) and (4) respectively but by using the academic knowledge stock.

$$Local.Private_i = \sum_{j \neq i} \frac{P_{ij}}{\sum_i P_{ij}} \cdot Private_j \cdot 1(d(i,j) \leq 50\ miles) \tag{3}$$

$$Distant.Private_i = \sum_{j \neq i} \frac{P_{ij}}{\sum_i P_{ij}} \cdot Private_j \cdot 1(d(i,j) > 50\ miles) \tag{4}$$

Finally, the level of intra-regional knowledge spillovers (***Intra***) is approximated by the share of regional self-citation flows within the same county ($Intra_i = P_{ii} / \sum_r P_{ri}$). P_{ri} is the (r^{th}, i^{th}) element of the (3,109×3,109) P matrix. Table 1 presents the list of these variables as well as their descriptive statistics. Since the metropolitan areas play a significant role in knowledge production and associated spillovers (Feldman and Audretsch 1999; Fischer et al. 2001), we assume the data generating process in the MSAs is different from that of the non-MSAs (Partridge et al. 2008). In addition, Kang and Dall'erba (2015) discover a significant structural difference in the knowledge production process across metropolitan and non-metropolitan counties. As such, we separate the samples of metropolitan (853 counties) and non-metropolitan counties (2,256 counties) based on the definition of Metropolitan Statistical Areas found in the 2000 US Census. However, the variables of localized and distant knowledge spillovers are calculated using all counties since spillovers do not take place only between metropolitan counties and vice versa. Table 1 indicates that, for all variables, the MSA counties display a much greater mean and median values than non-MSA counties. The column "Zeros" reports the number of counties with a zero value for each variable. Due to the presence of zeros in most variables, we added one to each of them before log transformation.

[Table 1]

3.2 Calibration of GWR and Mixed GWR

We start exploring the spatial heterogeneity in the knowledge production process across counties by relying on a GWR approach. A basic GWR model is expressed as Eq. (5) where y_i and x_{ik} are the dependent and the kth explanatory variables respectively, (u_i, v_i) denotes the geographical coordinates of the centroid of county *i* and $\beta_k(u_i, v_i)$ is the kth local coefficient at county *i* (Fotheringham et al. 2002). Eq. (6) presents the vector of GWR local estimates at county *i*. In the equation, $\boldsymbol{Y}$ is the vector of the dependent variable and $\boldsymbol{X}$ is the matrix of explanatory variables including the intercept. $\boldsymbol{W}(u_i, v_i)$ is the diagonal matrix of $diag(w_{i1}, w_{i2}, \cdots, w_{in})$ and each weight element w_{ij} represents the adjacency effects of neighboring counties to county *i* (Partridge et al. 2008). Various kernel functions can be used to define the weight element. In order to choose the best kernel function, we compare the model fits of GWR and of MGWR with four different kernel functions: Gaussian, Exponential, Bisquare and Tricube (Eqs. 7-10).[1] In the kernel functions, *b* denotes the kernel bandwidth and d_{ij} is the great circle distance between the centroids of counties *i* and *j*. Since the distribution of the observations is not uniform, we use an adaptive bandwidth defined as a fixed number of nearest neighbors (Fotheringham et al. 2002).

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \tag{5}$$

$$\widehat{\boldsymbol{\beta}}(u_i, v_i) = [\boldsymbol{X}'\boldsymbol{W}(u_i, v_i)\boldsymbol{X}]^{-1}\boldsymbol{X}'\boldsymbol{W}(u_i, v_i)\boldsymbol{Y} \tag{6}$$

$$Gaussian: w_{ij} = \exp\left[-\frac{1}{2}\left(\frac{d_{ij}}{b}\right)^2\right] \ if\ d_{ij} < b\ and\ w_{ij} = 0\ otherwise \tag{7}$$

$$Exponential: w_{ij} = \exp\left[-\frac{|d_{ij}|}{b}\right] \ if\ d_{ij} < b\ and\ w_{ij} = 0\ otherwise \tag{8}$$

$$Bisquare: w_{ij} = \left(1 - \left(\frac{d_{ij}}{b}\right)^2\right)^2 \ if\ d_{ij} < b\ and\ w_{ij} = 0\ otherwise \tag{9}$$

$$Tricube: w_{ij} = \left(1 - \left(\frac{d_{ij}}{b}\right)^3\right)^3 \ if\ d_{ij} < b\ and\ w_{ij} = 0\ otherwise \tag{10}$$

In order to choose the optimal kernel bandwidth, one can rely on the cross-validation (CV) score (Bowman 1984; Cleveland 1979) or the Akaike Information Criterion (AIC). However, the former option is known for two shortcomings. First, it only considers model prediction accuracy (Lu et al. 2014a); second, several studies show that it returns a too small optimal bandwidth (McMillen 2010) which may lead to extreme coefficients (Farber and Páez 2007). On the other hand, AIC accounts for a trade-off between model complexity and prediction accuracy and it can be used to compare the goodness-of-fit of a GWR model with the one of a global model (Fotheringham et al. 2002). For these reasons, we adopt it and rely on its corrected version (AICc) with respect to the effective degree of freedom in a GWR model (Hurvich et al. 1998). The optimal bandwidth is the one that minimizes AICc as described in Eq. (11) where n is the number of observations, $\hat{\sigma}$ is the estimated standard deviation of the error term and $tr(\boldsymbol{S})$ is the trace of the hat matrix $\boldsymbol{S}$ (Eq. 12). The vector $\boldsymbol{X_i}$ of the hat matrix is the ith row of the explanatory variable matrix $\boldsymbol{X}$.

$$AIC_c = 2 \cdot n \cdot ln(\hat{\sigma}) + n \cdot \ln(2\pi) + n \cdot \left\{\frac{n+tr(\boldsymbol{S})}{n-2-tr(\boldsymbol{S})}\right\} \tag{11}$$

$$\boldsymbol{S} = \begin{bmatrix} \boldsymbol{X_1'}[\boldsymbol{X'W}(u_1,v_1)\boldsymbol{X}]^{-1}\boldsymbol{X'W}(u_1,v_1) \\ \boldsymbol{X_2'}[\boldsymbol{X'W}(u_2,v_2)\boldsymbol{X}]^{-1}\boldsymbol{X'W}(u_2,v_2) \\ \vdots \\ \boldsymbol{X_n'}[\boldsymbol{X'W}(u_n,v_n)\boldsymbol{X}]^{-1}\boldsymbol{X'W}(u_n,v_n) \end{bmatrix} \tag{12}$$

The GWR calibration allows us to explore the non-stationarity of every coefficient. However, some of them may not vary significantly over space. As a result, we use a Monte Carlo random permutation test (Brunsdon et al. 1998) to identify what factors do not vary spatially. This step is necessary before we calibrate a MGWR model (Eq. 13) in which some coefficients are global and the remaining ones are local. We follow the procedure proposed in Fotheringham et al. (2002, pp. 65-68) where $\boldsymbol{X_G}$ denotes the matrix of explanatory variables with stationary coefficients and $\boldsymbol{X_L}$ is its counterpart for non-stationary coefficients: 1) regress each explanatory variable of $\boldsymbol{X_G}$ on $\boldsymbol{X_L}$ using a basic GWR and compute the residuals; 2) regress $\boldsymbol{y}$ on $\boldsymbol{X_L}$ using a basic GWR and compute the residuals; 3) regress the $\boldsymbol{y}$-residuals of step 2 on the $\boldsymbol{X_G}$-residuals of step 1 using OLS. The OLS coefficients of this regression are a vector of global coefficients $\widehat{\boldsymbol{\beta_G}}$; 4) regress $(\boldsymbol{y} - \boldsymbol{X_G}\widehat{\boldsymbol{\beta_G}})$ on $\boldsymbol{X_L}$ using a basic GWR. The GWR coefficients of this regression are the local MGWR coefficients $\widehat{\boldsymbol{\beta_L}}(u_i, v_i)$. Calibrated global and local coefficients are expressed as Eqs. (14) and (15) respectively where $\boldsymbol{S_L}$ is the hat matrix composed of $\boldsymbol{X_L}$ and $\boldsymbol{I}$ is the compatible identity matrix.

$$y_i = \sum_{k=1}^{q} \beta_{k,G} x_{ik,G} + \sum_{k=q+1}^{K} \beta_{k,L}(u_i,v_i) x_{ik,L} + \varepsilon_i \tag{13}$$

$$\widehat{\boldsymbol{\beta_G}} = [\boldsymbol{X_G'}(\boldsymbol{I}-\boldsymbol{S_L})'(\boldsymbol{I}-\boldsymbol{S_L})\boldsymbol{X_G}]^{-1}\boldsymbol{X_G'}(\boldsymbol{I}-\boldsymbol{S_L})(\boldsymbol{I}-\boldsymbol{S_L})\boldsymbol{Y} \tag{14}$$

$$\widehat{\boldsymbol{\beta_L}}(u_i,v_i) = [\boldsymbol{X_L'W}(u_i,v_i)\boldsymbol{X_L}]^{-1}\boldsymbol{X_L'W}(u_i,v_i)\left(\boldsymbol{Y} - \boldsymbol{X_G}\widehat{\boldsymbol{\beta_G}}\right) \tag{15}$$

4 Calibration results

Tables 2 and 3 present the model fit for the kernel functions described above for the MSA and the non-MSA data. Since the Gaussian kernel returns the lowest AICc values of MGWR for both distance cut-offs and MSA/non-MSA data, we will present the calibration results based on the Gaussian kernel throughout the rest of this paper[2]. Table 4 shows the result of the spatial stationarity test for both the MSA/non-MSA and the 50/75 miles distance cut-offs. The large majority of our coefficients vary spatially at the 5% significance level, which confirms the need to adopt a GWR approach.

[Table 2]

[Table 3]

[Table 4]

For the MSA counties, the GWR and MGWR calibration results are reported in Tables 5 and 6. Despite different spatial extents of localized spillovers (50 miles for Table 5, 75 miles for Table 6), the estimates of the models and the chosen optimal bandwidths are very similar to each other. Since the model fit of MGWR with 75 miles distance cut-off is better in terms of AICc, we focus our interpretation of the results on Table 6.

[Table 5]

[Table 6]

Model 4 of Table 6 shows the OLS results for the sample of 853 MSA counties as the baseline global model. In the model, the columns “2.5%” and “97.5%” denote the lower and upper bounds of the 95% confidence intervals for the estimated coefficients. The estimates show that both the private and the academic knowledge stocks have a significant and positive influence on knowledge creation across MSA counties although the former has a magnitude about 3 times greater than the latter. It is clear that among all our variables, human capital plays the greatest role on knowledge creation. Sectoral diversity and the share of large firms do not have a significant impact on knowledge creation while the size of the local labor force does. It is different from the knowledge spillovers where, with the exception of the localized spillovers of academic knowledge, all types show a significant and positive role. The distant interregional spillovers display a greater elasticity than the local ones. This result is consistent with the previous studies emphasizing the importance of distant knowledge spillovers (Asheim and Isaksen 2002; Gertler and Levitte 2005; Trippl et al. 2009).

Model 5 presents the calibration results of GWR. The chosen optimal adaptive bandwidth of the model is 61 nearest neighboring counties. With respect to the adjusted R^2 and the Residual Sums of Squares (RSS), the GWR results show better goodness-of-fits than the OLS model. In addition, the AICc of GWR is smaller and the difference in AICc between Models 4 and 5 are much greater than 3, the rule of thumb brought to the fore by Fotheringham et al. (2002, p.96) for model selection. Although the GWR model fits the MSA sample better than the OLS model, the median values of the GWR coefficients are generally close to the OLS estimates, especially for the significant ones. In addition, the 95% confidence intervals of the OLS estimates overlap the intervals of the first and third quartiles of the GWR coefficients. However, we note that the ranges of the minimum and maximum coefficients of the GWR are much wider than the OLS confidence intervals. Even though all the GWR coefficients show spatially varying patterns across MSA counties, the Monte Carlo test based on 499 random permutations (Brunsdon et al. 1998) does not reject the hypothesis of spatially stationary coefficients at the 5% significance level for the following variables: private knowledge stock, human capital level, intra-regional spillovers, local and distant spillovers of academic knowledge.

As such we turn to the MGWR coefficients which are reported in Model 6. The chosen optimal bandwidth in the first and second steps of the MGWR procedure (see the previous section) is 61 and the one for the fourth step is 54. Although the difference in the AICc values between MGWR and GWR is less than 3, the former has a smaller AICc value (in the case of 50 miles distance cut-off, the difference in the AICc is about 6 which confirms that MGWR is a better model, see Table 5). In order to test the significance level of the global explanatory variables, we adopt the F-approximation suggested by Mei et al. (2006). We find that such coefficients are significant at the 1% level and generally closer to the corresponding GWR median values than to the OLS estimates. In addition, the local spillovers of academic knowledge which are not significant in the OLS model show a significant and positive global impact on knowledge creation now. Considering that many previous studies support the significant role of such spillovers (Acs et al. 1994; Audretsch and Feldman 1996; Jaffe 1989; Ó hUallacháin and Leslie 2007), our MGWR results suggest that the role of significant factors can be concealed when spatial heterogeneity is not sufficiently controlled for in a global model.

For the local MGWR coefficients, we find that their median, first and third quartiles are generally close to those of the GWR model. When we focus on the intervals of the first and third quartiles, most local coefficients present a consistent sign. The exceptions are the diversity index and the share of large firms. This result is not necessarily surprising considering that the literature has not reached a consensus on the relative importance of specialization vs. diversity (Fung and Chow 2002; Glaeser et al. 1992; Henderson 2003; Jaffe 1986) and on the role of small vs. large firms (Acs and Armington 2004; Anselin et al. 1997) on knowledge creation. It could be that the relative role of diversity and small firms is subject to heterogeneous regional conditions such as institutional infrastructures and social assets so that their spatially varying elasticity is to be expected.

Our results also indicate a non-stationary elasticity associated to both local and distant private knowledge spillovers. Since a region's absorptive capacity to utilize external knowledge sources is determined by the combination of the local agents' innovation ability, regional knowledge stocks (Mukherji and Silberman 2013; Verspagen and Schoenmakers 2004), learning cost (Cohen and Levinthal 1990) and region-specific conditions such as institutions, industrial structure and historical similarity (Agrawal et al. 2010; Döring and Schnellenbach 2006), each MSA county is expected to display a different level of regional absorptive capacity. Figs. 1 and 2 display the geographical distribution of the MGWR coefficients of local and distant private knowledge spillovers respectively. Because statistical inference on the local MGWR coefficients is based on 853 multiple hypotheses, we calculate the Bonferroni style adjusted t-value following the Fotheringham-Byrne procedure (Byrne et al. 2009) and report only the coefficients significant at 5% in our figures. As shown in Fig. 1, we find that the significant impact of localized private knowledge spillovers is concentrated in the central and eastern MSAs of the country with the largest values found in Texas and several Southeastern states. Fig. 2 displays the MSA counties' elasticity of distant spillovers of private knowledge. Their largest value appears in the Western and Northeastern states. Many Texas counties that display a large impact of local spillovers do not show any significant role for the distant ones. The opposite holds true for California, Washington, Massachusetts, New Jersey and Connecticut.

The relative importance of the local buzz and global pipeline on regional innovation may explain this distinct geographical distribution of the spillover coefficients (Feldman and Kogler 2010; Moreno and Miguélez 2012). As many empirical studies point out (Asheim and Isaksen 2002; Gertler and Levitte 2005; Trippl et al. 2009), firms in one location can benefit from distant but highly relevant external knowledge sources as well as from local knowledge pools. We believe it is the case of many central MSA counties that display significant local coefficients for both localized and distant private knowledge spillovers. At the same time, some very innovative regions play a critical role in local knowledge creation but do not benefit much from their direct neighbors. Instead, they rely on distant collaborators connected with through global pipelines. Many counties in California, Washington, Massachusetts and New Jersey correspond to the previous case: as Fig. 3 indicates, they are the most innovative MSA counties, yet their MGWR coefficients of local private knowledge spillovers are not significant. Instead, the role of distant spillovers for these counties is not only significant but also greater than anywhere else. On the other hand, in many highly innovative counties of Texas (see Fig. 3) the significant spillovers are local only. It might be related to the composition of the innovative industry in this region: Texas is specialized in Computers and Semiconductors (Audretsch and Feldman 1996) whereas the aforementioned innovative states are leading regions across various innovative industries. For instance, the average of relative diversity index (***Diversity***) of the Texas MSA counties is 4.79. It is 4.83, 4.98, 5.03, 5.41 and 5.65 in the MSA counties of Massachusetts, California, New Jersey, Washington and Connecticut respectively. In addition, Anselin et al. (2000) show that Machinery (SIC 35) and Electronics (SIC 36) sectors which include Computers and Semiconductors experience significant local knowledge spillovers but their estimates are not measured with the presence of long-distance spillovers. Hence a sectoral analysis combining both types of spillovers would shed new light into the dichotomy found here. This is left for future research.

[Fig. 1]

[Fig. 2]

[Fig. 3]

Tables 7 and 8 show the results for the sample of the 2,256 non-MSA counties with the 50 and 75 miles distance cut-offs respectively. The OLS results with 50 miles cut-off (Model 7) indicate that the private and academic knowledge stocks show a significant and positive role on knowledge creation. Human capital is the knowledge input variable with the largest elasticity. The index of diversity is still not significant while the share of large firms is significantly detrimental to the innovation output among non-MSA counties. All types of knowledge spillovers, including the localized academic spillovers, have a significant and positive impact now. The GWR results (Model 8) show a better goodness-of-fit than the OLS model with respect to the adjusted R^2, RSS and AICc. As in the MSA case, the median values of GWR coefficients are generally close to the OLS estimates so that the 95% confidence interval of the OLS estimates overlaps with the values of the first and third quartiles of the GWR coefficients. The Monte Carlo test based on 499 random permutations indicates that the hypothesis of stationary coefficients is not rejected at the 5% level for the following variables: academic knowledge stock, diversity and all types of intra- and interregional spillovers (see Table 4). As a result, the MGWR coefficients calibrated for the non-MSA counties are reported in Model 9. The MGWR shows a better model goodness-of-fit than the GWR in terms of adjusted R^2, RSS and AICc. All the global coefficients are significant at the 1% significant level based on the F-approximation and their magnitudes are closer to the corresponding GWR median values than to the OLS estimates. For the local MGWR coefficients, their median, first and third quartiles are generally close to those of the GWR model and most of the local coefficients in the intervals of the first and third quartiles display a consistent sign.

[Table 7]

[Table 8]

When the spatial extent of localized spillovers is extended up to 75 miles, the results of OLS (Model 10) and GWR (Model 11) are generally consistent with those based on 50 miles. However, the estimates associated to local academic spillovers and its relative role with respect to local private spillovers are greater in both the OLS and the GWR. We believe that since more than 90% of the non-MSA counties have zero values for the local academic spillovers (see Table 1), an increase in the spatial extent of spillovers leads to a sensitive change in the estimates than in the MSA case. Similarly, while the MGWR coefficients are generally not sensitive to a change in distance cut-off, we find some differences. For instance, in Model 12 local private spillovers and distant academic spillovers become spatially non-stationary compared to Model 9. Furthermore, the elasticity of local academic spillovers becomes greater than the one of the median local private spillovers in Model 12. Although the MGWR based on 75 miles shows a better model fit than when based on 50 miles, we are cautious about the results since the model may be over-fitted. Indeed, in Fig. 4 that presents the geographical distribution of the MGWR coefficients of local private spillovers, one can observe that few are significant at the Bonferroni adjusted p-value of 5%. In addition, their calibrated elasticity is greater than 1 in several locations. We conjecture that the smaller optimal number of nearest neighbors based on 75 miles (31 neighbors) than in the case of 50 miles (38 nearest neighbors) is the reason why we find these extreme coefficients (Farber and Páez 2007).

[Fig. 4]

5 Conclusion

Since the contribution of Jaffe (1989), an increasing number of studies has relied on geographically aggregated units to understand the spatial dynamics of knowledge creation. This literature has built on theoretical and empirical advances in the geography of innovation (Anselin et al. 1997; Audretsch and Feldman 2004; Feldman and Kogler 2010; Jaffe 1989) and has recognized the role of region-specific conditions such as the degree of diversity of its economy, human capital of the local workforce and the intra-regional and interregional knowledge spillovers. Since these regional conditions are heterogeneous over space, one should expect the regional mechanisms of knowledge creation to vary spatially as well. However, most empirical studies in the literature have relied on regression models that implicitly assume the elasticity of the knowledge inputs does not vary over space (Anselin et al. 1997; Jaffe 1989; Ó hUallacháin and Leslie 2007; Ponds et al. 2010). Such elasticity corresponds to the average impact of an input across the entire sample, which may mask a positive impact in some areas and a less positive or even negative one in other areas. As such it may lead to locally biased misspecifications (McMillen and Redfearn 2010). In order to explore the spatial heterogeneity in the regional knowledge production functions of the US counties, this paper adopts the nonparametric local modeling approaches of GWR and MGWR. Yet another contribution compared to the existing literature consists in investigating the individual effect of private vs. academic knowledge stocks and of localized spillovers vs. distant knowledge spillovers. The latter are measured through a unique matrix of interregional patent creation-citation flows.

Based on two distinct samples of MSA and non-MSA counties, the calibration results show that the role of most knowledge input variables varies over space but some display a spatially homogenous elasticity on regional knowledge creation. For the MSA sample, the MGWR calibrations uncover the significant but spatially homogenous role of local knowledge spillovers emanating from universities, a factor that is not statistically significant in the OLS model. Considering that the GWR approach partly accounts for county fixed effects by using the spatially varying constant coefficients (Partridge et al. 2008), these results support the idea that a significant factor can be concealed when spatial heterogeneity is not sufficiently controlled for in a global model. The local MGWR coefficients may show both positive and negative elasticities for the same variable, thus reflecting the spatial heterogeneity present in our data and model. It is the case for the index of industrial diversity, which does not come as a surprise considering that the relative importance of diversity (Jacobian externalities) vs. specialization (Marshall-Arrow-Romer externalities) is still debated in the literature in general (Fung and Chow 2002; Glaeser et al. 1992; Henderson 2003; Jaffe 1986) and is subject to the heterogeneous characteristics of our spatial observations in particular. We also find that the spillovers of private knowledge vary spatially especially in MSA counties (both locally and over long distances). It confirms our expectations about the dominance of and more complex innovation capacity of the MSA counties in the nation’s knowledge production. At the same time, localized academic spillovers are found to display spatially homogenous elasticities in both MSA and non-MSA counties. However, the spatial stationarity of the distant academic spillovers is sensitive to the distance cut-off (50 vs. 75 miles) in the case of the non-MSA.

Our MGWR results suggest two policy implications. First, policy makers should pay more attention to the spatial heterogeneity present in the regional knowledge production process, especially across metropolitan regions. Recently Kang and Dall’erba (2015) finds that there are significant structural differences in the mechanisms of regional knowledge creation between MSA and non-MSA counties. The MGWR results provided here go one step further by highlighting the significant presence of heterogeneity at the county level. In particular, we find that the impact of sectoral diversity is not spatially homogenous across MSA counties. As such it would be fruitful to modify on-going debates about regional specialization vs. regional diversity (Fung and Chow 2002; Glaeser et al. 1992;

Henderson 2003; Jaffe 1986) to questions that embrace the naturally large differences in the regions' characteristics such as "how can a region capitalize on its specific industrial and institutional structures to innovate and grow?" In addition, the returns of local and distant interregional spillovers of private knowledge show significant spatial variations, but the relative role of these two types of spillovers does not show similar patterns across MSA counties. Some innovative regions generate knowledge spillovers locally but they do not learn much from their geographical neighbors in return as their main knowledge partners are remotely located. On the other hand, other innovative regions depend on their local neighbors by a much greater extent. As such, regional policy makers need to identify which type of spillovers is the most beneficial to their local innovation process before they can figure out how to facilitate knowledge spillovers.

Second, policy makers need to support academic knowledge spillovers at the national level. Our results show that the localized spillovers of academic knowledge generate spatially homogenous contributions to knowledge creation across both MSA and non-MSA counties. When we use a 50 miles distance cut-off, the distant spillovers of academic knowledge display spatially homogenous returns across both MSA and non-MSA counties as well. Considering that academic knowledge has the characteristics of a public good (Foray and Lissoni 2010), regions should have a relatively high capacity to absorb it even if they are short in human capital stock and cannot afford high learning cost (Liu 2013; Verspagen and Schoenmakers 2004). Therefore, the cost-efficiency of academic knowledge spillovers should be recognized and be supported for greater regional innovation.

Although spatial heterogeneity is definitely an intrinsic component of knowledge creation, each industrial sector has distinctive innovation characteristics as reported in previous studies (Anselin et al. 2000; Mansfield 1995). We also cannot disregard the fact that the temporal aspect of innovation (LeSage and Sheng 2014; Parent 2012), especially in the form of path dependence, is a critical dimension to consider (David 1994). As such, our future research proposes to consider simultaneously the sectoral, spatial and temporal heterogeneities present in our data to deepen our understanding of the dynamics of knowledge creation and innovation across US counties.

---

[1] We thank an anonymous reviewer for suggesting us to try out various kernel functions and distance cut-offs in the definition of the spatial extent of localized spillovers in order to find the best model specification.

[2] The GWR and MGWR calibration results with the other kernel functions are available upon request.

Table 1 Descriptive statistics

		MSA county (853)						Non-MSA county (2,256)					
Variable	Explanation	Mean	Med.	S.D.	Min	Max	Zeros	Mean	Med.	S.D.	Min	Max	Zeros
Patent	Total patents (fractional count)	150.9	22.4	494.3	0.0	9359.7	5	3.4	0.9	9.2	0.0	162.4	433
Private	Private R&D (million dollars)	821.2	0.0	4909.4	0.0	99161.9	423	2.6	0.0	51.1	0.0	2097.0	2100
Univ	Academic R&D (million dollars)	139.8	0.0	587.9	0.0	11826.9	569	3.8	0.0	52.1	0.0	1815.9	2154
Graduate	Share of graduate degree holders (%)	5.1	4.4	2.8	1.2	25.2	0	3.1	2.7	1.4	0.5	14.7	0
Diversity	Level of sectoral employment diversity	4.5	4.4	1.5	1.1	13.0	0	2.8	2.7	0.9	0.8	6.5	0
Large	Share of large firms (employees > 499) (%)	8.4	8.0	3.1	1.1	40.0	0	7.7	7.6	3.4	0.0	26.2	35
Intra	Share of intra-county citation (%)	11.9	10.5	11.7	0.0	100.0	164	5.3	0.0	14.5	0.0	100.0	1731
Local.Private50	Local private knowledge spillovers within 50 miles	82.1	0.2	440.7	0.0	7682.2	347	0.2	0.0	1.9	0.0	33.0	2045
Local.Univ50	Local academic knowledge spillovers within 50 miles	10.4	0.0	80.6	0.0	2226.7	422	0.1	0.0	1.0	0.0	20.6	2111
Distant.Private50	Distant private knowledge spillovers over 50 miles	554.3	95.4	1694.1	0.0	32228.3	27	13.4	0.8	38.3	0.0	594.3	911
Distant.Univ50	Distant academic knowledge spillovers over 50 miles	108.5	22.8	318.6	0.0	6093.9	35	3.2	0.1	9.6	0.0	163.5	987
Local.Private75	Local private knowledge spillovers within 75 miles	95.2	0.6	461.7	0.0	7791.7	297	0.5	0.0	3.9	0.0	110.2	1973
Local.Univ75	Local academic knowledge spillovers within 75 miles	11.9	0.1	81.8	0.0	2245.3	351	0.2	0.0	1.3	0.0	21.4	2037
Distant.Private75	Distant private knowledge spillovers over 75 miles	541.2	92.5	1671.7	0.0	31832.8	27	13.1	0.8	37.0	0.0	586.4	915
Distant.Univ75	Distant academic knowledge spillovers over 75 miles	107.0	22.6	317.3	0.0	6075.4	35	3.2	0.1	9.5	0.0	163.5	989
Size	Total employees (thousand employees)	123.7	58.2	230.3	2.3	3953.4	0	10.4	7.3	10.0	0.0	83.7	1

Note: The column Zeros means the number of counties of having zero values for each variable.

Table 2 Model fit by kernel function (MSA)

		Distance cut-off: 50 miles				Distance cut-off: 75 miles			
Model	Model Fit statistic	Gaussian	Exponential	Bisquare	Tricube	Gaussian	Exponential	Bisquare	Tricube
GWR	Bandwidth	61	56	305	305	61	59	297	297
	Adjusted R^2	0.934	0.932	0.934	0.933	0.935	0.933	0.935	0.934
	RSS	146.9	141.8	149.4	152.2	144.8	141.6	146.6	149.4
	AICc	1120.0	1131.0	1123.6	1124.8	1107.9	1119.1	1112.2	1113.4
MGWR	Bandwidth (in 1st and 2nd steps)	61	56	305	305	61	59	297	297
	Bandwidth (in 4th step)	54	36	234	258	54	38	233	240
	Adjusted R^2	0.933	0.930	0.933	0.933	0.933	0.932	0.933	0.933
	RSS	156.3	152.3	154.5	157.0	155.0	149.3	156.9	157.2
	AICc	1113.8	1142.9	1114.9	1120.2	1106.9	1116.0	1110.9	1115.3

Note: An adaptive bandwidth is used in the GWR and mixed GWR models. The bandwidths of the two models are chosen by minimizing the value of the corrected Akaike Information Criteria (AICc). RSS means Residual Sums of Squares.

Table 3 Model fit by kernel function (non-MSA)

		Distance cut-off: 50 miles				Distance cut-off: 75 miles			
Model	Model Fit statistic	Gaussian	Exponential	Bisquare	Tricube	Gaussian	Exponential	Bisquare	Tricube
GWR	Bandwidth	86	38	575	660	79	38	575	577
	Adjusted R2	0.729	0.742	0.727	0.724	0.736	0.747	0.731	0.730
	RSS	451.1	377.1	464.7	476.9	436.8	368.6	456.2	462.4
	AICc	3084.3	3040.2	3097.6	3105.2	3039.5	2996.2	3057.0	3065.9
MGWR	Bandwidth (in 1st and 2nd steps)	86	38	575	660	79	38	575	577
	Bandwidth (in 4th step)	38	31	471	527	31	24	227	517
	Adjusted R2	0.732	0.727	0.725	0.724	0.747	0.741	0.735	0.728
	RSS	447.7	444.9	476.7	484.4	398.6	395.9	442.2	476.3
	AICc	3057.9	3068.7	3083.9	3089.3	2987.1	2995.9	3042.6	3055.2

Note: An adaptive bandwidth is used in the GWR and mixed GWR models. The bandwidths of the two models are chosen by minimizing the value of the corrected Akaike Information Criteria (AICc). RSS means Residual Sums of Squares.

Table 4 Non-stationarity Test (H0: spatially stationary coefficient, kernel function: Gaussian)

	MSA				Non-MSA			
	Distance cut-off: 50 miles P-value		Distance cut-off: 75 miles P-value		Distance cut-off: 50 miles P-value		Distance cut-off: 75 miles P-value	
Intercept	0.018	L	0.040	L	0.014	L	0.036	L
ln Private	0.186		0.196		0.018	L	0.020	L
ln Univ	0.032	L	0.014	L	0.176		0.140	
ln Graduate	0.100		0.142		0.006	L	0.002	L
ln Diversity	0.000	L	0.000	L	0.072		0.112	
ln Large	0.016	L	0.024	L	0.004	L	0.002	L
ln Intra	0.226		0.136		0.976		0.978	
ln Local.Private	0.018	L	0.002	L	0.066		0.006	L
ln Local.Univ	0.784		0.206		0.520		0.354	
ln Distant.Private	0.004	L	0.008	L	0.328		0.492	
ln Distant.Univ	0.414		0.410		0.168		0.040	L
ln Size	0.000	L	0.000	L	0.000	L	0.000	L

Note: Non-stationary local GWR coefficients are tested using the Monte Carlo test based on 499 random permutations. L indicates that the local GWR coefficient is significant at the 5% level.

Table 5 Calibration results (MSA, distance cut-off: 50 miles)

Dep. Variable:	Model 1			Model 2					Model 3				
ln Patent	OLS			GWR					Mixed GWR				
	2.5%	Estimate	97.5%	Min.	Q1	Median	Q3	Max.	Min.	Q1	Median	Q3	Max.
Intercept	-1.409	-1.001 ***	-0.593	-3.168	-1.466	-1.139	-0.668	0.247	-2.472	-1.553	-1.174	-0.830	-0.215
ln Private	0.041	0.061 ***	0.080	0.005	0.039	0.055	0.066	0.110			0.044 ***		
ln Univ	-0.002	0.017 *	0.036	-0.027	-0.002	0.016	0.034	0.081	-0.028	0.001	0.017	0.032	0.083
ln Graduate	0.449	0.569 ***	0.688	0.299	0.577	0.694	0.781	1.060			0.662 ***		
ln Diversity	-0.167	-0.036	0.094	-0.574	-0.288	-0.109	0.107	0.415	-0.523	-0.301	-0.083	0.140	0.351
ln Large	-0.169	-0.054	0.061	-0.410	-0.135	-0.049	0.070	0.444	-0.426	-0.112	-0.051	0.066	0.407
ln Intra	0.003	0.038 **	0.072	-0.083	0.011	0.028	0.055	0.145			0.028 ***		
ln Local.Private50	0.086	0.115 ***	0.145	0.034	0.103	0.130	0.154	0.232	-0.004	0.100	0.142	0.167	0.240
ln Local.Univ50	-0.048	-0.007	0.035	-0.028	0.001	0.030	0.051	0.089			0.029 ***		
ln Distnat.Private50	0.197	0.246 ***	0.296	0.119	0.157	0.203	0.293	0.428	0.094	0.165	0.216	0.269	0.505
ln Distnat.Univ50	0.087	0.145 ***	0.203	0.022	0.060	0.087	0.138	0.269			0.107 ***		
ln Size	0.338	0.400 ***	0.462	0.138	0.294	0.472	0.556	0.705	0.206	0.298	0.496	0.571	0.703
Observations		853		853					853				
Adjusted R^2		0.918		0.934					0.933				
RSS		210.17		146.89					156.26				
AICc		1252.23		1119.99					1113.76				
Kernel function				Gaussian					Gaussian				
Bandwidth				61					61 (in 1st and 2nd steps), 54 (in 4th step)				

Note: * P-value < 10%, ** P-value < 5%, *** P-value < 1%. The columns of 2.5% and 97.5% mean the lower and upper limits of the 95% confidence intervals for the OLS estimates. The columns of Q1 and Q3 mean the first and third quartiles of the calibrated coefficients associated to the explanatory variables. An adaptive bandwidth is used in the GWR and mixed GWR models. The bandwidths of the two models are chosen by minimizing the value of the corrected Akaike Information Criteria (AICc). Mixed GWR is calibrated based on the procedure described in Fotheringham et al. (2002, pp.65-68). Calibration is implemented using the R package "GWmodel" (Gollini et al. 2015; Lu et al. 2014b).

Table 6 Calibration results (MSA, distance cut-off: 75 miles)

Dep. Variable: ln Patent	Model 4 OLS			Model 5 GWR					Model 6 Mixed GWR				
	2.5%	Estimate	97.5%	Min.	Q1	Median	Q3	Max.	Min.	Q1	Median	Q3	Max.
Intercept	-1.402	-0.997 ***	-0.591	-2.936	-1.427	-1.136	-0.639	0.156	-2.432	-1.523	-1.166	-0.848	-0.182
ln Private	0.039	0.059 ***	0.078	0.010	0.041	0.056	0.067	0.113			0.044 ***		
ln Univ	0.000	0.019 **	0.038	-0.034	-0.001	0.016	0.034	0.082	-0.034	0.004	0.017	0.031	0.085
ln Graduate	0.445	0.564 ***	0.683	0.297	0.575	0.689	0.777	1.079			0.664 ***		
ln Diversity	-0.161	-0.032	0.098	-0.565	-0.269	-0.110	0.114	0.409	-0.487	-0.277	-0.074	0.130	0.334
ln Large	-0.154	-0.040	0.075	-0.348	-0.161	-0.060	0.066	0.420	-0.382	-0.132	-0.065	0.057	0.368
ln Intra	0.003	0.038 **	0.072	-0.093	0.009	0.028	0.053	0.152			0.028 ***		
ln Local.Private75	0.090	0.118 ***	0.147	0.041	0.097	0.128	0.152	0.208	-0.001	0.098	0.141	0.166	0.235
ln Local.Univ75	-0.031	0.010	0.051	-0.059	0.027	0.053	0.072	0.154			0.041 ***		
ln Distant.Private75	0.190	0.239 ***	0.287	0.108	0.153	0.195	0.269	0.433	0.086	0.157	0.210	0.255	0.488
ln Distant.Univ75	0.087	0.144 ***	0.202	0.018	0.059	0.087	0.143	0.269			0.108 ***		
ln Size	0.330	0.391 ***	0.453	0.136	0.280	0.461	0.552	0.713	0.201	0.301	0.486	0.573	0.713
Observations		853		853					853				
Adjusted R^2		0.919		0.935					0.933				
RSS		207.2		144.8					155.0				
AICc		1240.2		1107.9					1106.9				
Kernel function				Gaussian					Gaussian				
Bandwidth				61					61 (in 1st and 2nd steps), 54 (in 4th step)				

Note: * P-value < 10%, ** P-value < 5%, *** P-value < 1%. The columns of 2.5% and 97.5% mean the lower and upper limits of the 95% confidence intervals for the OLS estimates. The columns of Q1 and Q3 mean the first and third quartiles of the calibrated coefficients associated to the explanatory variables. An adaptive bandwidth is used in the GWR and mixed GWR models. The bandwidths of the two models are chosen by minimizing the value of the corrected Akaike Information Criteria (AICc). Mixed GWR is calibrated based on the procedure described in Fotheringham et al. (2002, pp.65-68). Calibration is implemented using the R package "GWmodel" (Gollini et al. 2015; Lu et al. 2014b).

Table 7 Calibration results (non-MSA, distance cut-off: 50 miles)

Dep. Variable: ln Patent	Model 7 OLS			Model 8 GWR					Model 9 Mixed GWR				
	2.5%	Estimate	97.5%	Min.	Q1	Median	Q3	Max.	Min.	Q1	Median	Q3	Max.
Intercept	-0.828	-0.662 ***	-0.496	-1.431	-1.113	-0.860	-0.540	-0.225	-1.722	-1.161	-0.886	-0.556	0.026
ln Private	0.069	0.107 ***	0.144	-0.067	0.032	0.097	0.148	0.286	-0.111	0.026	0.102	0.174	0.586
ln Univ	0.049	0.084 ***	0.119	-0.077	0.052	0.095	0.130	0.203			0.100 ***		
ln Graduate	0.373	0.455 ***	0.536	-0.232	0.261	0.351	0.516	0.756	-0.611	0.211	0.342	0.473	0.914
ln Diversity	-0.138	-0.022	0.094	-0.461	-0.128	0.069	0.186	0.323			0.055 ***		
ln Large	-0.232	-0.185 ***	-0.137	-0.325	-0.143	-0.078	-0.026	0.138	-0.567	-0.171	-0.071	-0.007	0.227
ln Intra	0.046	0.065 ***	0.084	0.004	0.047	0.057	0.070	0.104			0.052 ***		
ln Local.Private50	0.098	0.165 ***	0.233	-0.332	0.108	0.182	0.270	0.509			0.149 ***		
ln Local.Univ50	0.038	0.128 ***	0.217	-0.169	0.027	0.137	0.250	0.610			0.142 ***		
ln Distnat.Private50	0.073	0.096 ***	0.119	-0.004	0.059	0.085	0.105	0.173			0.081 ***		
ln Distnat.Univ50	0.105	0.138 ***	0.172	0.033	0.088	0.120	0.166	0.336			0.113 ***		
ln Size	0.460	0.501 ***	0.541	0.225	0.437	0.505	0.591	0.796	0.252	0.435	0.522	0.622	0.906
Observations		2256		2256					2256				
Adjusted R2		0.701		0.729					0.732				
RSS		544.7		451.1					447.7				
AICc		3222.4		3084.3					3057.9				
Kernel function				Gaussian					Gaussian				
Bandwidth				86					86 (in 1st and 2nd steps), 38 (in 4th step)				

Note: * P-value < 10%, ** P-value < 5%, *** P-value < 1%. The columns of 2.5% and 97.5% mean the lower and upper limits of the 95% confidence intervals for the OLS estimates. The columns of Q1 and Q3 mean the first and third quartiles of the calibrated coefficients associated to the explanatory variables. An adaptive bandwidth is used in the GWR and mixed GWR models. The bandwidths of the two models are chosen by minimizing the value of the corrected Akaike Information Criteria (AICc). Mixed GWR is calibrated based on the procedure described in Fotheringham et al. (2002, pp.65-68). Calibration is implemented using the R package "GWmodel" (Gollini et al. 2015; Lu et al. 2014b).

Table 8 Calibration results (non-MSA, distance cut-off: 75 miles)

Dep. Variable:	Model 10			Model 11					Model 12				
ln Patent	OLS			GWR					Mixed GWR				
	2.5%	Estimate	97.5%	Min.	Q1	Median	Q3	Max.	Min.	Q1	Median	Q3	Max.
Intercept	-0.824	-0.659 ***	-0.494	-1.457	-1.138	-0.852	-0.573	-0.263	-1.823	-1.191	-0.900	-0.485	0.081
ln Private	0.066	0.103 ***	0.140	-0.099	0.040	0.098	0.148	0.307	-0.163	0.007	0.090	0.162	0.755
ln Univ	0.053	0.088 ***	0.122	-0.079	0.056	0.105	0.138	0.235			0.103 ***		
ln Graduate	0.371	0.452 ***	0.534	-0.248	0.247	0.341	0.510	0.761	-0.733	0.206	0.329	0.480	0.922
ln Diversity	-0.132	-0.016	0.099	-0.472	-0.110	0.078	0.192	0.361			0.053 ***		
ln Large	-0.231	-0.183 ***	-0.136	-0.346	-0.136	-0.069	-0.016	0.147	-0.611	-0.181	-0.060	0.012	0.334
ln Intra	0.047	0.066 ***	0.084	0.005	0.046	0.058	0.071	0.103			0.053 ***		
ln Local.Private75	0.080	0.131 ***	0.181	-0.192	0.100	0.155	0.275	0.662	-0.797	0.050	0.156	0.304	1.695
ln Local.Univ75	0.124	0.200 ***	0.276	-0.138	0.088	0.197	0.344	0.648			0.195 ***		
ln Distant.Private75	0.071	0.095 ***	0.118	-0.005	0.056	0.085	0.104	0.167			0.072 ***		
ln Distant.Univ75	0.090	0.124 ***	0.158	-0.023	0.073	0.105	0.161	0.342	-0.089	0.061	0.112	0.172	0.656
ln Size	0.457	0.497 ***	0.537	0.204	0.434	0.506	0.592	0.788	0.128	0.404	0.513	0.639	0.982
Observations		2256		2256					2256				
Adjusted R2		0.704		0.736					0.747				
RSS		538.9		436.8					398.6				
AICc		3198.4		3039.5					2987.1				
Kernel function				Gaussian					Gaussian				
Bandwidth				79					79 (in 1st and 2nd steps), 31 (in 4th step)				

Note: * P-value < 10%, ** P-value < 5%, *** P-value < 1%. The columns of 2.5% and 97.5% mean the lower and upper limits of the 95% confidence intervals for the OLS estimates. The columns of Q1 and Q3 mean the first and third quartiles of the calibrated coefficients associated to the explanatory variables. An adaptive bandwidth is used in the GWR and mixed GWR models. The bandwidths of the two models are chosen by minimizing the value of the corrected Akaike Information Criteria (AICc). Mixed GWR is calibrated based on the procedure described in Fotheringham et al. (2002, pp.65-68). Calibration is implemented using the R package "GWmodel" (Gollini et al. 2015; Lu et al. 2014b).

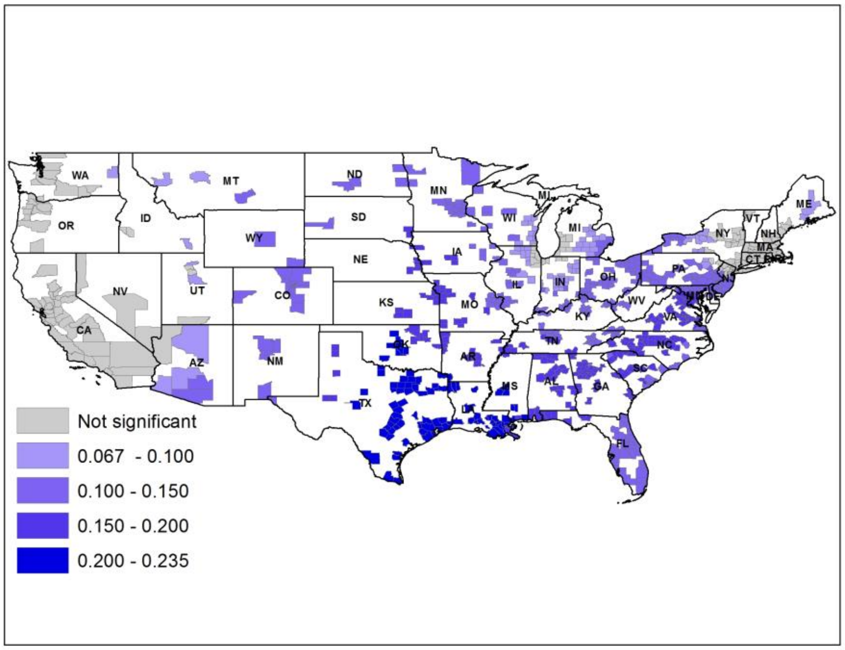

Fig. 1 MGWR coefficients of local private knowledge spillovers (MSA, distance cut-off: 75 miles)

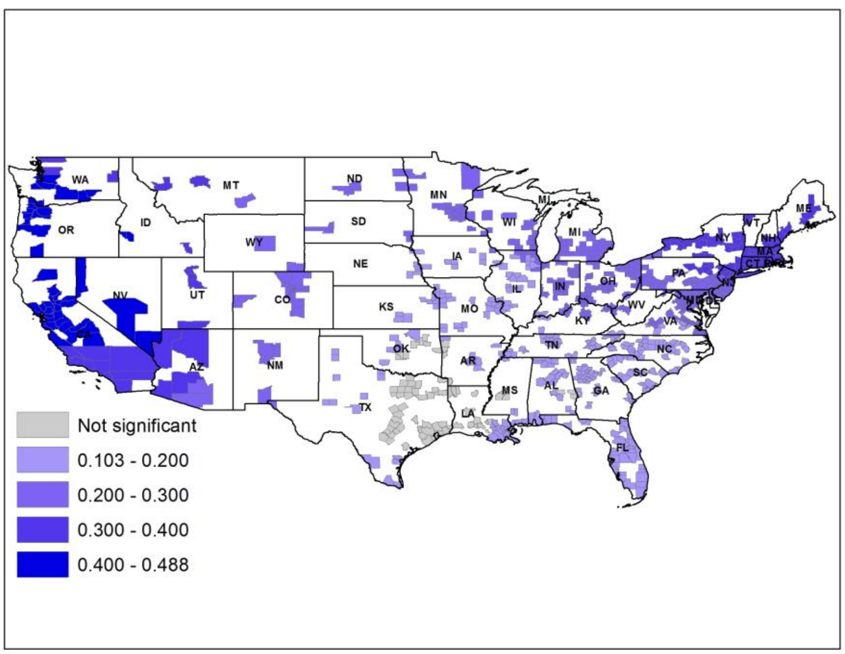

Fig. 2 MGWR coefficients of distant private knowledge spillovers (MSA, distance cut-off: 75 miles)

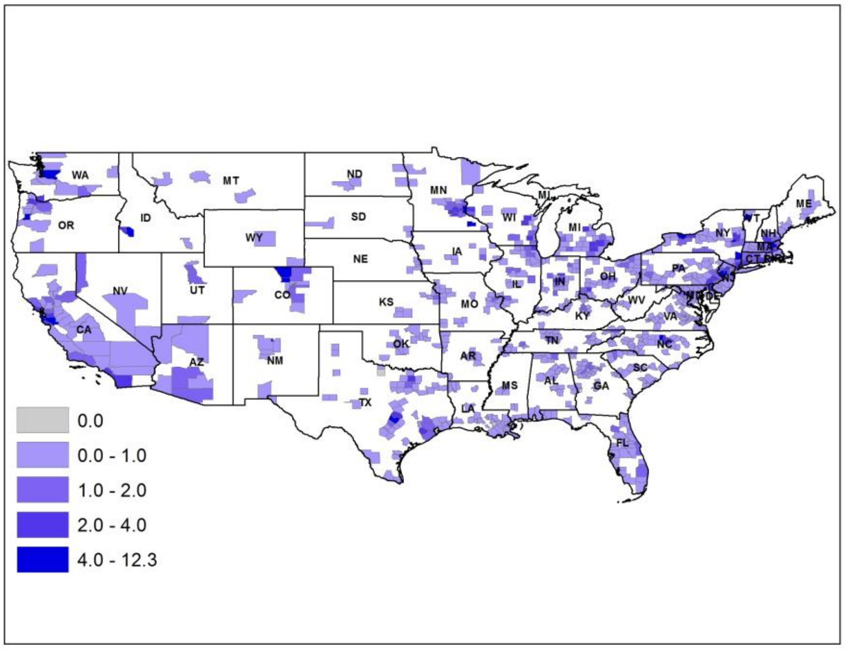

Fig. 3 Total patent applications per employee (MSA)

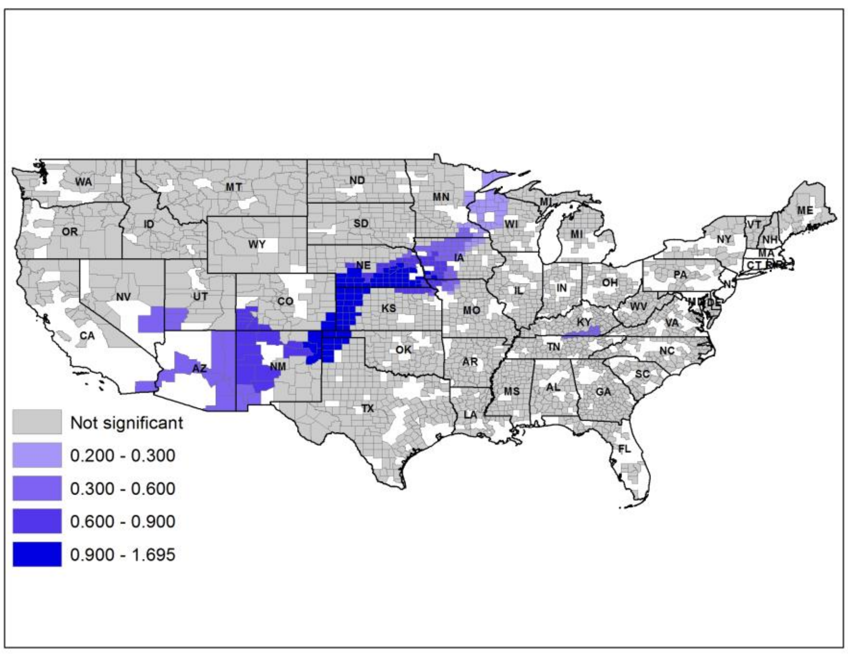

Fig. 4 MGWR coefficients of local private knowledge spillovers (non-MSA, distance cut-off: 75 miles)

References

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

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

Acs ZJ, Audretsch DB, Feldman MP (1994) R & D spillovers and recipient firm size The Review of Economics and Statistics 76:336-340

Agrawal A, Cockburn I, Rosell C (2010) Not invented here? Innovation in company towns Journal of Urban Economics 67:78-89

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

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

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

Asheim B, Isaksen A (2002) Regional innovation systems: The integration of local 'sticky' and global 'ubiquitous' knowledge The Journal of Technology Transfer 27:77-86

Asheim BrT (1996) Industrial districts as 'learning regions': A condition for prosperity European Planning Studies 4:379-400

Asheim BT, Gertler MS (2006) The geography of innovation: Regional innovation systems. In: Fagerberg J, Mowery DC (eds) The oxford handbook of innovation. Oxford University Press, Oxford, pp 291-317.

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

Audretsch DB, Feldman MP (2004) Chapter 61 knowledge spillovers and the geography of innovation. In: Henderson JV, Jacques-François T (eds) Handbook of regional and urban economics, vol Volume 4. Elsevier, San Diego, CA, pp 2713-2739.

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

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

Bárcena MJ, Menéndez P, Palacios MB, Tusell F (2014) Alleviating the effect of collinearity in geographically weighted regression J Geogr Syst 16:441-466

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

Bowman AW (1984) An alternative method of cross-validation for the smoothing of density estimates Biometrika 71:353-360

Brunsdon C, Fotheringham AS, Charlton ME (1996) Geographically weighted regression: A method for exploring spatial nonstationarity Geographical Analysis 28:281-298

Brunsdon C, Fotheringham S, Charlton M (1998) Geographically weighted regression-modelling spatial non-stationarity Journal of the Royal Statistical Society Series D (The Statistician) 47:431-443

Byrne G, Charlton M, Fotheringham S Multiple dependent hypothesis tests in geographically weighted regression. In: Lees BG, Laffan SW (eds) the 10th International Conference on GeoComputation, Sydney, Australia, 30 November - 02 December 2009 2009.

Camagni R (1991) Local 'milieu', uncertainty and innovation networks: Towards a new dynamic theory of economic space. In: Camagni R (ed) Innovation networks: Spatial perspectives. Belhaven, London, pp 121-142.

Capello R, Lenzi C (2014) Spatial heterogeneity in knowledge, innovation, and economic growth nexus: Conceptual reflections and empirical evidence Journal of Regional Science 54:186-214

Cleveland WS (1979) Robust locally weighted regression and smoothing scatterplots Journal of the American Statistical Association 74:829-836
Cohen WM, Levinthal DA (1990) Absorptive capacity: A new perspective on learning and innovation Administrative Science Quarterly 35:128-152
Cooke P, Gomez Uranga M, Etxebarria G (1997) Regional innovation systems: Institutional and organisational dimensions Research Policy 26:475-491
David PA (1994) Why are institutions the 'carriers of history'?: Path dependence and the evolution of conventions, organizations and institutions Structural Change and Economic Dynamics 5:205-220
Döring T, Schnellenbach J (2006) What do we know about geographical knowledge spillovers and regional growth?: A survey of the literature Regional Studies 40:375-395
Duranton G, Puga D (2000) Diversity and specialisation in cities: Why, where and when does it matter? Urban Studies 37:533
Farber S, Páez A (2007) A systematic investigation of cross-validation in gwr model estimation: Empirical analysis and monte carlo simulations J Geogr Syst 9:371-396
Feldman MP, Audretsch DB (1999) Innovation in cities: Science-based diversity, specialization and localized competition European Economic Review 43:409-429
Feldman MP, Kogler DF (2010) Chapter 8 - stylized facts in the geography of innovation. In: Bronwyn HH, Nathan R (eds) Handbook of the economics of innovation, vol Volume 1. North-Holland, pp 381-410.
Fischer MM, Revilla-Diez J, Snickars F (2001) 1. Systems of innovation: An attractive conceptual framework for comparative innovation research. In: Fischer MM, Revilla-Diez J, Snickars F (eds) Metropolitan innovation systems. Advances in spatial science. Springer, Berlin Heidelberg, pp 1-21.
Fischer MM, Scherngell T, Jansenberger E (2006) The geography of knowledge spillovers between high-technology firms in europe: Evidence from a spatial interaction modeling perspective Geographical Analysis 38:288-309
Florida R (1995) Toward the learning region Futures 27:527-536
Foray D, Lissoni F (2010) Chapter 6 - university research and public–private interaction. In: Bronwyn HH, Nathan R (eds) Handbook of the economics of innovation, vol Volume 1. North-Holland, pp 275-314.
Fotheringham AS, Brunsdon C, Charlton M (2002) Geographically weighted regression: The analysis of spatially varying relationships. Wiley, Hoboken, NJ
Fritsch M, Stephan A (2005) Regionalization of innovation policy—introduction to the special issue Research Policy 34:1123-1127
Fung MK, Chow WW (2002) Measuring the intensity of knowledge flow with patent statistics Economics Letters 74:353-358
Gertler MS, Levitte YM (2005) Local nodes in global networks: The geography of knowledge flows in biotechnology innovation Industry and Innovation 12:487-507
Glaeser EL, Kallal HD, Scheinkman JA, Shleifer A (1992) Growth in cities Journal of Political Economy 100:1126-1152
Gollini I, Lu B, Charlton M, Brunsdon C, Harris P (2015) Gwmodel: An r package for exploring spatial heterogeneity using geographically weighted models Journal of Statistical Software 63:50
Griliches Z (1979) Issues in assessing the contribution of research and development to productivity growth The Bell Journal of Economics 10:92-116
Griliches Z (1992) The search for R&D spillovers The Scandinavian Journal of Economics 94:29-47
Hall BH, Jaffe AB, Trajtenberg M (2001) The nber patent citations data file: Lessons, insights and methodological tools. National Bureau of Economic Research, Cambridge, MA
Harris R (2011) Models of regional growth: Past, present and future Journal of Economic Surveys 25:913-951
Henderson JV (2003) Marshall's scale economies Journal of Urban Economics 53:1-28

Hurvich CM, Simonoff JS, Tsai C-L (1998) Smoothing parameter selection in nonparametric regression using an improved akaike information criterion Journal of the Royal Statistical Society: Series B (Statistical Methodology) 60:271-293
Jacobs J (1969) The economy of cities. Random House, New York
Jaffe AB (1986) Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value The American Economic Review 76:984-1001
Jaffe AB (1989) Real effects of academic research The American Economic Review 79:957-970
Jaffe AB, Trajtenberg M, Henderson R (1993) Geographic localization of knowledge spillovers as evidenced by patent citations The Quarterly Journal of Economics 108:577-598
Keeble D, Lawson C, Moore B, Wilkinson F (1999) Collective learning processes, networking and ‘institutional thickness' in the cambridge region Regional Studies 33:319-332
Lazaric N, Lorenz E (1998) Trust and organisational learning during inter-firm cooperation. In: Lazaric N, Lorenz E (eds) Trust and economic learning. Edward Elgar, Cheltenham, pp 209-226.
LeSage J, Sheng Y (2014) A spatial econometric panel data examination of endogenous versus exogenous interaction in chinese province-level patenting J Geogr Syst 16:233-262
Liu W-H (2013) The role of proximity to universities for corporate patenting: Provincial evidence from china The Annals of Regional Science 51:273-308
Lorenz EH (1992) Trust, community and co-operation: Toward a theory of industrial districts. In: Scott AJ, Storper M (eds) Pathways to industrialization and regional development. Routledge, London, pp 175-182.
Lu B, Charlton M, Harris P, Fotheringham AS (2014a) Geographically weighted regression with a non-euclidean distance metric: A case study using hedonic house price data International Journal of Geographical Information Science 28:660-681
Lu B, Harris P, Charlton M, Brunsdon C (2014b) The gwmodel r package: Further topics for exploring spatial heterogeneity using geographically weighted models Geo-spatial Information Science 17:85-101
Maggioni MA, Uberti TE, Usai S (2010) Treating patents as relational data: Knowledge transfers and spillovers across italian provinces Industry and Innovation 18:39-67
Mansfield E (1995) Academic research underlying industrial innovations: Sources, characteristics, and financing The Review of Economics and Statistics 77:55-65
Marshall A (1920) Principles of economics. 8 edn. Macmillan, London
Maurseth PB, Verspagen B (2002) Knowledge spillovers in europe: A patent citations analysis The Scandinavian Journal of Economics 104:531-545
McMillen DP (2010) Issues in spatial data analysis Journal of Regional Science 50:119-141
McMillen DP, Redfearn CL (2010) Estimation and hypothesis testing for nonparametric hedonic house price functions Journal of Regional Science 50:712-733
Mei CL, Wang N, Zhang WX (2006) Testing the importance of the explanatory variables in a mixed geographically weighted regression model Environment and Planning A 38:587-598
Moreno R, Miguélez E (2012) A relational approach to the gegogrpahy of innovation: A typology of regions Journal of Economic Surveys 26:492-516
Morgan K (1997) The learning region: Institutions, innovation and regional renewal Regional Studies 31:491-503
Mukherji N, Silberman J (2013) Absorptive capacity, knowledge flows, and innovation in u.S. Metropolitan areas Journal of Regional Science 53:392-417
Guide for public use data files national science foundation’s higher education research and development survey: Fiscal year 2011 (2013) NSF National Center for Science and Engineering Statistics. http://www.nsf.gov/statistics/herd/pub_data.cfm.
Ó hUallacháin B, Leslie TF (2007) Rethinking the regional knowledge production function Journal of Economic Geography 7:737-752

Oughton C, Landabaso M, Morgan K (2002) The regional innovation paradox: Innovation policy and industrial policy The Journal of Technology Transfer 27:97-110
Paez A, Farber S, Wheeler D (2011) A simulation-based study of geographically weighted regression as a method for investigating spatially varying relationships Environment and Planning A 43:2992-3010
Parent O (2012) A space-time analysis of knowledge production J Geogr Syst 14:49-73
Parent O, LeSage JP (2008) Using the variance structure of the conditional autoregressive spatial specification to model knowledge spillovers Journal of Applied Econometrics 23:235-256
Partridge MD, Rickman DS, Ali K, Olfert MR (2008) The geographic diversity of u.S. Nonmetropolitan growth dynamics: A geographically weighted regression approach Land Economics 84:241-266
Peri G (2005) Determinants of knowledge flows and their effect on innovation The Review of Economics and Statistics 87:308-322
Ponds R, Oort Fv, Frenken K (2010) Innovation, spillovers and university–industry collaboration: An extended knowledge production function approach Journal of Economic Geography 10:231-255
Romer PM (1986) Increasing returns and long-run growth Journal of Political Economy 94:1002-1037
Rosenkopf L, Almeida P (2003) Overcoming local search through alliances and mobility Management Science 49:751-766
Simmie J (2011) Learning regions. In: Cooke P, Asheim B, Boschma R, Martin R, Schwartz D, Tödtling F (eds) Handbook of regional growth and innovation. Edward Elgar, Cheltenham, pp 547-555.
Smallen D (2004) 3.3 million americans are stretch commuters traveling at least 50 miles one-way to work. United States Department of Transportation. https://www.rita.dot.gov/bts/sites/default/files/rita_archives/bts_press_releases/2004/bts010_04/html/bts010_04.html. Accessed November 16 2014
Standard & Poor’s (2011) Standard & poor’s compustat® xpressfeed understanding the data
Storper M (1993) Regional “worlds” of production: Learning and innovation in the technology districts of france, italy and the USA Regional Studies 27:433-455
Storper M (1997) The regional world: Territorial development in a global economy. Guilford Press,
Stough RR (2003) Strategic management of places and policy The Annals of Regional Science 37:179-201
Trippl M, Tödtling F, Lengauer L (2009) Knowledge sourcing beyond buzz and pipelines: Evidence from the vienna software sector Economic Geography 85:443-462
USPTO (2010) Patents bib: Selected bibliographic information from us patents issued 1969 to present. Alexandria, VA
Verspagen B, Schoenmakers W (2004) The spatial dimension of patenting by multinational firms in europe Journal of Economic Geography 4:23-42
Wei C-H, Qi F (2012) On the estimation and testing of mixed geographically weighted regression models Economic Modelling 29:2615-2620
Wheeler D, Tiefelsdorf M (2005) Multicollinearity and correlation among local regression coefficients in geographically weighted regression J Geogr Syst 7:161-187
Zysman J (1994) How institutions create historically rooted trajectories of growth Industrial and Corporate Change 3:243-283