ARE HIGH-TECH EMPLOYMENT AND NATURAL AMENITIES LINKED?: ANSWERS FROM A SMOOTHED BAYESIAN SPATIAL MODEL

Jeffrey H. Dorfman
Department of Agricultural and Applied Economics
The University of Georgia
Athens, GA 30602-7509

Mark D. Partridge
Department of Agricultural, Environmental, & Development Economics
The Ohio State University
Columbus, OH 43210

and

Hamilton Galloway
Economic Modeling Specialists Inc.
Moscow, ID 83843

April 30, 2008

Selected paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Orlando FL, July 29, 2008.

---

* The authors thank EMSI for their efforts in compiling the data used in this study and for making that data available.

Abstract

We investigate the recently advanced theory that high-technology workers are drawn to high amenity locations and then the high-technology jobs follow the workers. Using a novel data set that tracks high-technology job growth by U.S. county, we estimate spatial parameters of the response of job growth to the level of local natural amenities. We achieve this estimation with a reasonably new class of models, smooth coefficient models. The model is employed in a spatial setting to allow for smooth, but nonparametric response functions to key variables in an otherwise standard regression model. With spatial data this allows for flexible modeling such as a unique place-specific effects to be estimated for each location, and also for the responses to key variables to vary by location. This flexibility is achieved through the non-parametric smoothing rather than by nearest-neighbor type estimators such as in geographically weighted regressions. The resulting model can be estimated in a straightforward application of analytical Bayesian techniques. Our results show that amenities can definitely have a significant effect on high-technology employment growth; however, the effect varies over space and by amenity level.

Keywords: Bayesian econometrics, employment growth, high technology, smooth coefficient models, spatial modeling.

1. Introduction

Many rural areas have long lagged their urban cousins in terms of job growth and income gains (McGranahan and Beale, 2002). However, recent decades have seen a renaissance in many rural communities due to migration to rural locations driven by high amenities such as pleasant landscapes, mountains, lakes, or oceans (McGranahan, 1999; Deller et al., 2001, 2006). This migration pattern appears to extend far beyond the effects of tourism and retirees, and seems to extend to other businesses.

Along with this shifting migration pattern to natural amenities (aside from climate), there have been technological changes that appear to have reduced the costs of distance and remoteness (Cairncross, 1995, Kotkin, 1998). Indeed, persistent trends toward amenity migration may support a surge in advanced technology firms locating in rural communities. Though such stories often seem to be overblown, one can picture a highly-educated knowledge worker being on her computer, attached to a satellite internet hook-up, in a remote bucolic setting. With the potential for higher wages in technology firms along with the opportunity to diversify the local economy, it is easy to imagine why many rural communities would welcome the prospect of such firms locating in their towns.

There are other arguments supporting the claim that rural communities are increasingly able to attract advanced-technology employment. First, combining the facts that amenities are generally normal goods and higher-skilled workers have more income, it seems to be a natural fit that advanced technology firms would increasingly wish to locate in rural communities to follow their employees preferences. Indeed, Partridge et al. (2008b) find evidence that subsequent nonmetropolitan employment growth was higher in places with greater initial shares of college graduates, especially in high-amenity areas. Moreover, using geographically weighted regression techniques, they found that the marginal impacts of both amenities and college attainment spatially varied across the country, suggesting further heterogeneities.

Despite the possibilities raised here, efforts to attract advanced technology firms and thereby replicate Silicon Valley (even on a mini-scale) are not new. The general story has been that it is hard for state and local policy to influence their location (e.g., Partridge, 1993). For example, Saxenian's (1996) description of Silicon Valley and Route 128 near Boston stresses organizational and cultural aspects rather than public policy, while accidental location effects are also apparent when considering Microsoft's location in Seattle. In terms of nonmetropolitan areas, using data from the 1970 and early 1980s, Barkley et al. (1989) point out that rural areas tended to greatly lag in terms of attracting high-tech manufacturing, while the types of high-tech firms they attracted tended to be more mature and labor-intensive. Thus, one could question whether the notion of attracting advanced technology firms to rural communities is even feasible.

Though many rural areas offer spectacular amenities, they also lack basic elements that are traditionally associated with prosperous advanced technology magnets. Foremost, rural communities often lack the localization and urbanization economies that form the core basis of agglomeration and city formation (Rosenthal and Strange, 2001). In particular, rural areas appear to lack the labor market pooling and knowledge spillovers that are the basis of why many of these firms would prefer locating in cities. Specifically, immigrants and young urban professionals who form the core of the creative class may strongly prefer the Bohemian culture found in urban neighborhoods (Florida, 2002). Yet, there are many high-amenity nonmetropolitan areas that have managed to attract creative-class occupations (McGranahan and Wojan, 2007). Indeed, these areas have significantly faster job growth relative to other nonmetro counties. Thus, natural amenities may counteract some of the effects of urban amenities stressed by Florida and his followers.

Overall, Barkley et al.'s (1989) pessimistic appraisal may no longer apply due to the changes in technology and the rising importance of amenities. Rural areas now may be more competitive in attracting advanced technology firms. Thus, this appears to be an opportune moment to reassess rural advanced technology employment in a comprehensive fashion that considers amenities, human capital, access to urban areas, proximity to research universities, and allows for spatial heterogeneity. Following this roadmap, this study investigates the determinants of county-level advanced technology employment growth over the 2000-2006 period.

This paper thus joins a strong and growing literature using a wide variety of statistical techniques to analyze large, spatially-organized data sets in applications such as the modeling of regional employment growth, industry agglomeration effects, population trends, income distributions, and more. One important aspect of empirical modeling of such spatial problems is the forced tradeoff between flexibility of the statistical model and feasibility of computation. In an ideal world, the coefficients of the statistical model could vary freely by location to allow for place-specific effects and more accurate forecasts of future events. Unfortunately, such place-specific models require sufficient data on each location to allow feasible estimation of all the unknown parameters, and that amount of data is rarely available (or new enough to still be relevant). Thus, models are commonly estimated using data for a region (nearest neighbor estimation and geographically weighted regression techniques), thereby increasing the available data, or parameters are restricted across locations, thus decreasing the number of unknown parameters. Some type of tradeoff is required to allow estimation to succeed.

In this paper, we take a new tack, by applying a new estimation approach that allows location-specific parameters to be estimated that fall somewhere in between the two more traditional approaches. We utilize a variation of the smooth coefficient model developed by Koop and Poirier (2004) and Koop and Tobias (2006) that we have adapted to the problems inherent in spatial data. This semi-parametric model does not fully restrict parameters across locations, but instead only requires the location-specific parameters to vary smoothly in some prescribed pattern. Thus, the parameters could be considered partially restricted and because the pattern of smoothing is, in part, based on geography there are some elements of geographically weighted regressions present as well. However, the method is more flexible than either earlier approach as will be explained below.

The remainder of this paper is organized as follows. In section 2 we discuss the smooth coefficient models adapted for analysis of spatial data in regional economics applications. In section 3, we present data for our application. Section 4 presents econometric results and discusses the policy implications of our findings. Conclusions follow in section 5.

2. A Model with Smooth Spatial and Amenity Characteristics

Begin with a simple linear model that has two coefficients that vary with their associated observation:

$$y_i = x_i\beta + f_1(i) + z_i f_2(i, w_i) + \epsilon_i, \tag{2.1}$$

where y_i is the spatially-indexed variable to be modeled (e.g., high-technology employment growth in location i), x_i is a k-vector of location-specific explanatory variables, β is a vector of coefficients to be estimated that do not vary with location, $f_1(i)$ is a nonparametric location-specific effect, z_i is a variable whose effect on the y_i varies by either location i

or by the value of the variable w_i (or both), and ϵ_i is the observation-specific random stochastic term. Observations are indexed by a spatial subscript $i = 1, 2, ...n$.

The two nonparametric parts of the model in equation (2.1) are written as two functions, f_1 and f_2. While these functions can be generalized, we have written them here so that $f_1(i)$ is a location-specific intercept and $f_2(i, w_i)$ is a location-specific coefficient that designates the expected impact of variable z_i on y_i. Denoting f_2 to be a possible function of both w_i and i allows for the possibility that f_2 varies in a spatial pattern, in relation to changes in variable w_i, or both. Importantly, because f_1 and f_2 are nonparametric the location effects and the effect of variable z_i on y_i are not constrained to be linear or even continuous.

Following up on the work of Partridge et al. (2008), it is likely that the determinants of advanced technology employment growth vary spatially across the country. Thus, we utilize our Bayesian semi-parametric methodology to estimate state-specific intercept parameters which will be smoothed to be similar to the two states with the closest state-average amenity index. These intercepts are the f_1 function. We then allow spatial heterogeneity across the sample in a very flexible manner for the parameter on an amenity index. We make f_2 a function of agglomeration effects by allowing the parameter f_2 to vary by county while smoothing the variation in relation to changes in a variable that tracks the population of the nearest MSA (our choice for the w_i variable). This smoothing will dampen the noise in the county-specific amenity parameters while still introducing considerable spatial flexibility. We present the methodology below in somewhat general terms, but the reader can keep in mind the role to be played by the two non-parametric functions within the model.

In order to demonstrate the methodology of smoothing the two nonparametric components of the model, it is easier to work with all the observations stacked into matrices. Thus, rewrite the model in (2.1) as

$$y = X\beta + S\gamma + Z\kappa + \epsilon = W\lambda + \epsilon, \tag{2.2}$$

where y, X, and ϵ are the usual vertical concatenations of the y_i, x_i and ϵ_i, S is a matrix with n rows and a column for each of n_s included states where the (i, j) element equals one if the observation in row i is in state j, Z is a diagonal matrix of the z_i, and γ and κ are column vectors of the n_s and n values of the two respective nonparametric functions, f_1 and f_2. The model as written obviously needs some restrictions as it has n observations and $n + n_s + k$ parameters to be estimated. The smoothing of the nonparametric functions will serve as the necessary reduction in free parameters even though the parameters are not placed exactly on a continuous function (Koop and Tobias, 2006).

To accomplish the smoothing of the nonparametric functions, one must first define what is meant by "smooth." It makes sense to begin with the county-specific amenity effects.

The function f_2 is made smooth in the sense that the effect of z_i changes "smoothly" as the variable w_i increases. To make this concrete, order the observations so that w_i is increasing from first to last observation. Then the necessary smoothing matrix is

$$D_2 = \begin{bmatrix} 1 & -2 & 1 & 0 & \dots & & 0 \\ 0 & 1 & -2 & 1 & 0 & \dots & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & \dots & & \dots & 1 & -2 & 1 \end{bmatrix}, \tag{2.3}$$

The reader should note that D_2 is $((n-2)\times n)$, not square, due to the inability to smooth the first two data points of the nonparametric function with this approach. That is, the initial conditions are left free. This matrix D_2 will allow a transformation of the county-specific amenity-effects into a vector of the smoothed differences which will actually be estimated along with the first two county-specific effects. Then the differencing transformation can be reversed to recover the remainder of the county-specific effects by use of the simple formula $f_2(i) = f_2^*(i) + 2f_2(i-1) - f_2(i-2)$ where $f_2^*(i)$ are the smoothed differences.

Next, the function f_1 will be made smooth in the sense that the state-specific intercepts (or effects) that it represents will be restricted to be similar to those of the two states with the adjoining ranked state-average amenity index. This concept is made mathematical through the application of a smoothing matrix, denoted D_1. Similar to D_2, D_1 will be $((n_s-2)\times n_s)$ where n_s is the number of states. If the observations were arranged so that states appeared in order of increasing state-average amenity ranks, the smoothing matrix would be identical in form to D_2. However, because the observations will be ordered to facilitate the modeling of the amenity effect, the pattern of the state-specific intercepts will be non-standard and cannot be written out in general. Each row of D_1 will have a -2 on the main diagonal, but the two elements that equal 1 will not be on each side as they are in D_2. Instead the 1's appear in the columns for the adjacent states in the ordering of the state-average amenity index. All other elements of the row will equal 0. Thus, the matrix D_1 defines smooth as the state-specific effect changing in a nearly linear manner from low amenity states to high amenity states.

The definition of a full-model smoothing matrix R,

$$R = \begin{bmatrix} 0 & D_1 & 0 \\ 0 & 0 & D_2 \end{bmatrix}, \tag{2.4}$$

allows for smoothness of the two nonparametric functions to be imposed by the linear approximate restriction

$$R\lambda \approx 0, \tag{2.5}$$

which imposes $n + n_s - 4$ approximate restrictions on the $n + n_s + k$ parameters. If the restrictions in (2.5) were imposed exactly, the state effects would be linear in state-average amenities and the effect of the amenity variable z_i on y would be linearly increasing throughout the range of the data (in the order of variable w_i). By imposing the restrictions embodied in (2.5) through a Bayesian prior with a nonzero prior variance, we will allow these two effects to vary over space, but not to be completely unfettered. Thus, the model will have state-specific effects that vary gradually over space and the effect of the variable z_i can vary as w_i increases, but in a gradual, more continuous way than without the smoothness prior.

2.1 Some Helpful Definitions and Notation

To simplify the derivation of the posterior distribution of the parameters of interest, it is useful to define a few more subsets of parameters to be estimated and smoothing matrices. Let

$$\lambda^* = \begin{pmatrix} \lambda_1 \\ R_2\lambda_2 \end{pmatrix}, \tag{2.6}$$

where $\lambda_1 = (\beta', \gamma_1, \gamma_2, \kappa_1, \kappa_2)', \lambda_2 = (\tilde{\gamma}', \tilde{\kappa}'), \tilde{\gamma} = (\gamma_3, ..., \gamma_s)'$, and $\tilde{\kappa} = (\kappa_3, ..., \kappa_n)'$. Note that $R_2\lambda_2$ is the vector of smoothed effects, both the state-specific effects and the spatially-varying amenity parameter. Further, partition R to define R_2 from (2.6) as

$$R_2 = \begin{bmatrix} \tilde{D_1} & 0 \\ 0 & \tilde{D_2} \end{bmatrix}, \tag{2.7}$$

where $\tilde{D_1}$ and $\tilde{D_2}$ are respectively D_1 and D_2 minus the first two columns on the left, and

$$R_1 = \begin{bmatrix} 0 & D_1^* & 0 \\ 0 & 0 & D_2^* \end{bmatrix}. \tag{2.8}$$

The definition of R_1 includes the first two columns of D_1 and D_2 that were removed to free up the initial conditions of the smoothing functions. Finally, define the data matrix partition $W = [W_1 \; W_2]$ such that

$$S = [\, s_1 \quad s_2 \quad S_2 \,]\,,\; Z = [\, z_1 \quad z_2 \quad Z_2 \,]\,,$$

$$W_1 = [\, X \quad s_1 \quad s_2 \quad z_1 \quad z_2 \,]\,,\; W_2 = [\, S_2 \quad Z_2 \,]\,, \tag{2.9}$$

where s_1, s_2, z_1, and z_2 are the first two columns of the respective matrices S and Z.

With these additional matrices defined, a transformed model can be defined that is easier to work with. This model has the transformed data matrices

$$W_1^* = W_1 - W_2 R_2^{-1} R_1 \text{ and } W_2^* = W_2 R_2^{-1}. \tag{2.10}$$

Using these transformed data matrices, we can rewrite the model as

$$y = W\lambda + \epsilon = W_1^* \lambda_1^* + W_2^* \lambda_2^* + \epsilon = W^* \lambda^* + \epsilon. \tag{2.11}$$

2.2 The Prior

To analyze this model within a Bayesian framework we need a prior distribution for all the unknown random parameters. In particular, we need prior distributions for λ^* and for σ_ϵ^2. If we employ the natural conjugate prior, this model can actually be examined analytically, avoiding the need for numerical approximation methods. Thus, unless one has a strong reason for choosing another shape for one's prior beliefs about the model parameters, using the normal-Gamma prior seems a good choice. Therefore, we assume a prior distribution of the form

$$p(\lambda^*, \sigma_\epsilon^{-2}) \sim NG(m_o, V_o, s_o^{-2}, \nu_o). \tag{2.12}$$

The prior mean of the regression model parameters, m_o, would commonly be set to a vector of zeros unless the researcher possessed specific information on the parameters. In the smooth parameter model developed here, choosing zeroes for the prior means of the parameters within λ_2^* is critical as that is what imposes the smoothness on the nonparametric functions f_1 and f_2, so the final $n + n_s - 4$ elements of the prior mean should always be set to zero. The variance of the prior on λ^*, V_o, controls how near to m_o one believes the elements of λ^* to be, as well as whether one believes the parameters to be independent or correlated in some way. However, because some of the parameters are part of the smoothed non-parametric functions, it is likely best to specify this matrix in four parts,

$$V_o = \begin{bmatrix} \tau_1 I_k & 0 & 0 & 0 \\ 0 & \tau_2 I_4 & 0 & 0 \\ 0 & 0 & \tau_3 I_{s-2} & 0 \\ 0 & 0 & 0 & \tau_4 I_{n-2} \end{bmatrix}. \tag{2.13}$$

This partion of the prior variance allows for the researcher to place a loose prior on the structural parameters in λ_1 by setting τ_1 to a relatively large scalar. The parameter τ_2 allows a prior variance on the initial conditions for the smoothed parameters; this is also likely to be set to a fairly large scalar in most applications. In turn, τ_3 and τ_4 control how smooth the individual spatial effects and the effect of z_i on the dependent variable are to be; smaller values of (τ_3, τ_4) lead to smoother non-parametric functions. Conversely, as τ_3, τ_4 go to infinity, f_1, f_2 become fully flexible non-parametric functions and the model cannot be estimated due to underidentification.

The Gamma prior on the error variance term is a standard one. Common choices of values for s_0^{-2} are on the order of 0.1 or 0.01. The degree of freedom hyperparameter ν_o in the Gamma prior is typically set to a small, positive integer representative of the size of an imaginary sample of data used to measure the amount of prior information held about the variance.

2.3 The Posterior Distributions

If one assumes that the ϵ_i are *i.i.d.* as normal random variables with zero mean and constant variance σ_ϵ^2, that is equivalent to specifying the standard normal-gamma likelihood function for the observations on y_i. With such a likelihood function and the prior described in the previous subsection, Bayes' Theorem leads one to a posterior distribution in the normal-Gamma form:

$$p(\lambda^*, \sigma_\epsilon^{-2}) \sim NG(m_p, V_p, s_p^{-2}, \nu_p), \tag{2.14}$$

where

$$V_p = (V_o^{-1} + W^{*\prime} W^*)^{-1}, \tag{2.15}$$

$$\nu_p = \nu_o + n, \tag{2.16}$$

$$m_p = V_p \left(V_o^{-1} m_o + W^{*\prime} y \right), \tag{2.17}$$

and

$$s_p^2 = \nu_p^{-1} \left(\nu_o s_o^2 + (y - W^* m_p)'(y - W^* m_P) + ((m_o - m_p)' V_o^{-1} (m_o - m_p) \right). \tag{2.18}$$

Because the conditional posterior distribution of λ^* is normal and the transformation from λ to λ^* was a linear one, it is simple to recover the posterior estimates of the elements of λ and those original, structural parameters will also have conditional posterior distributions that are normal. Also, note that if one chooses to work with the marginal distribution of λ^*, integrating out σ_ϵ^2 will yield a t-distribution for λ^*. Either the conditional or marginal distribution makes it easy to construct a variety of probability statements about elements of λ^* or any linear function of these parameters, say $A\lambda^*$.

For example, to recover the location-specific estimates of f_1 and f_2 that were denoted in stacked form by λ_2 from the marginal distribution of λ_2^*, one simply inverts the transformation in (2.6). If we define m_{2p} and V_{2p} as the relevant subvector and submatrix of m_p and V_p, respectively, that would give a marginal distribution for λ_2 of

$$p(\lambda_2|s_p^2) \sim t_{\nu_p-k-2}(R_2^{-1}m_{2p}, s_p^2R_2^{-1}V_{2p}(R_2^{-1})'). \tag{2.19}$$

This marginal posterior for the nonparametric function estimates is a t-distribution, making it straightforward for the researcher to make statistical inferences about the location-specific effects and the location-specific impact of amenities.

3. Data for Modeling County High-Tech Employment Growth

To assess the role of amenities in drawing high-technology employment to specific locations, we take advantage of an innovative dataset provided by EMSI. This data set provides accurate county sectoral employment to the four-digit level, allowing us to much more closely measure advanced technology employment than past studies that often were forced to rely on two or three digit data. In order to capture its complete picture of industry employment, EMSI combines covered employment data from Quarterly Census of Employment and Wages (QCEW) produced by the Department of Labor with total employment data in Regional Economic Information System (REIS) published by the Bureau of Economic Analysis (BEA), modified with County/ZIP Business Patterns (CBP) and Non-employer Statistics (NES), both published by the U.S. Census Bureau.

The data integration process seeks to unsuppress QCEW to the 6-digit industry level for all counties in the United States and to combine the dataset with various other sources to fill in existing holes. The first step involves combining QCEW with CBP to remove

QCEWs internally placed suppressions and arrive at industry specific county level data. QCEW is the best, highly detailed single data source available but this dataset is corroborated with REIS, an equally reliable but less detailed source, to ensure its accuracy. To achieve this, modified QCEW and NES data sets are used as seed numbers to disaggregate 2 and 3-digit REIS numbers among 3 and 4-digit industries. At this point, proprietors and non-covered employment data from NES and REIS are also extracted. This is necessary because QCEW does not include these sections of the workforce. After these three steps, mid-level industry employment data specificity is achieved. The last step focuses on inputting the mid-level unsuppressed REIS numbers back into the unsuppressed 6-digit QCEW data and adjusting the QCEW numbers accordingly. Once the REIS matrix has been incorporated back into QCEW the industry employment data set is complete. In this project, only 4-digit level employment specificity was used.

High technology employment is defined for the purposes of this study as employment in 14 4-digit NAICS industries, sometimes called level-I high technology industries because they have the highest percentage of employees that fit the definition of high technology jobs. Based on Hecker (2005), these 14 industries have at least 5 times the average proportion of high technology workers and constitute 24.7% of all high technology employment. The list of these 4-digit industries is provided in table 1. We sum the employment in all 14 industies using the unsuppressed data from EMSI for two years: 2000 and 2006. Our dependent variable is the difference in employment between these two years; that is, 2006 employment minus 2000 employment.

We also check the county-level high-technology employment data by ensuring that all counties have at least minimal employment estimates. This is to minimize the chance that our unsuppression methods do not invent jobs and to make any errors in generating the unsuppressed data less important in magnitude (since at very low numbers, any error would be magnified). Thus, all counties that did not have at least 10 high technology employees in both 2000 and 2006 were eliminated. The data set is left with 2,937 counties for our regression.

For explanatory variables in our model, we used a set of seven variables in addition to our state-specific effects and the amenity variable that is of particular interest. Those seven variables capture the impact on high-tech employment growth of local population, nearby research universities, urban amenities, housing prices, and education levels. We use 1990 values for time dependent variables as "deep lags" to minimize endogeneity concerns. Details of these variables are as follows.

Our first variable is the 1990 log of the county's high-tech employment to account for localization or Marshallian economies in which greater agglomeration economies take place due to factors such as labor-market pooling, better access to inputs, knowledge spillovers, etc (Glaeser et al., 1992; Rosenthal and Strange, 2001; Desmet and Fafchamps, 2005, Partridge, 1993; Partridge et al., 2008). Though increased concentration of industry employment may be associated with localization economies or "clusters," the general evidence is that concentrations of sectoral employment in a given location are inversely associated with subsequent local growth in that industry–i.e., a reversion to the mean effect.

A key input for knowledge-oriented businesses is highly skilled labor. We proxy for this variable by accounting for the 1990 share of the population over 25 years old that has a four-year college degree (using 1990 Census of Population data). We anticipate greater shares of college graduates to be positively related to subsequent growth in high-tech employment (Partridge, 1993). Because we are modeling high-technology employment, rates of achievement for lower levels of education were not included as a control variable.

Another advantage of rural locations is lower costs for land and for housing, which would be attractive for both the firm and for the workforce. We account for this by including the log of the 1990 median home value from the 1990 Census of Population.

A rural life style in high-amenity locations is appealing to many individuals. Yet, these same individuals may want access to higher-ordered consumer services only found in successively larger cities (Partridge et al., 2007, 2008). Likewise, advanced technology firms need access to urban markets as well as need to be closely proximate to specialized inputs. For example, using data from the 1970s and early 1980s, Smith and Barkley (1991) find that metropolitan adjacency is a key factor supporting rural high-tech manufacturers. We account for proximity to urban areas with a series of distance and population variables used by Partridge (2008) in their study of the urban hierarchy. The first measure is distance to the nearest metropolitan area (or urban center) of any size, measured as the distance in kilometers from the population-weighted centroid of the rural county to the population-weighted centroid of the metro area. Beyond the nearest metro area of any size, we also include the incremental distances in kms to larger, higher-tiered urban centers to reflect added "penalties" such as costs for households and businesses in acquiring higher-order services and access to customer markets. These distances reflect the incremental or marginal costs to reach each higher-tiered (larger) urban center. Specifically, the model includes the incremental distance in kms from the county to reach a metro area of at least 250,000, at least 500,000, and at least 1.5 million people. The largest category generally corresponds to national and top-tier regional centers, with the 500,000-1.5 million category reflecting sub-regional tiers.

Because the closest metropolitan area may be most important to households and firms for urban services and markets, we also include the 1990 population of the nearest metropolitan or micropolitan area in the model. Though the effects of this population variable are partially captured by the distance variables, we expect that a larger nearest urban area will be positively related to rural advanced technology employment growth.

The last control variable is the one to measure proximity to a major university. Using a list of major land-grant and research universities, a map was developed containing the geographic locations of the institutions. In addition, an additional map was developed using county-level geographically weighted population center data provided by the U.S. Census. The weighted population center map was superimposed upon the university map. From here, a 50 mile buffer was placed around the universities, allowing the geographically weighted population centers contained within each of the buffer zones to be identified. The result is a variable that equals one for all counties within the 50 mile buffer surrounding at least one of these research universities.

Because the role of natural amenities on high technology employment and migration is of particular interest in this study, the choice of variable to measure amenities is very important. The amenity variable used is a 1 to 7 amenity ranking produced by the Economic Research Service of the U.S. Department of Agriculture, with higher numbers reflecting greater amenities. This measure is a composite of various natural amenities that have been shown to influence migration such as climate, topography, and access to water area. As already noted, natural amenities are increasingly key factors driving rural and urban migration flows (McGranahan, 1999; Deller et al., 2001, 2006; Ferguson et al, 2007). As in the case of urban amenities, natural amenities are generally normal goods, which would be especially attractive to knowledge workers with high incomes. Indeed, the notion that information technology workers could live anywhere–including high amenity locales–is a driving factor behind claims such as the "death of distance" and "Forty Acres and a Modem" (Cairncross, 1995; Kotkin, 1998). Despite these romantic notions, many observers have stated that personal contact and access to urban areas are still key features of business location including for advanced technology firms (Kolko, 1999; Smith and Barkley, 1991; Partridge et al., 2008).

4. Econometric Results and Policy Implications

In estimating our model of high-technology employment growth and amenities using the county data set just described, we have to specifiy all the prior parameters in the model as shown in (2.12). We do so by setting these parameters to $m_o = \mathbf{0}$, $\nu_o = 10$, and $s_o^2 = 0.1$. The specification of V_o according to (2.13) is made by setting $\tau_1 = \tau_2 = 10$, $\tau_3 = 0.001$, and $\tau_4 = 0.0001$.

The estimation results of the smoothed coefficient model for the structural parameters in β are presented in table 2, using posterior means as point estimates for the regression parameters. We find by comparing posterior means to posterior standard deviations that satisfactory statistical precision has been achieved for the variables measuring lagged high technology employment, distance to the nearest metropolitan area with a population of at least 250,000, the lagged share of college graduates, and the lagged population of the nearest metropolitan area. Also, for a relatively simple model on a huge cross-section data set with enormous variability, the R^2 is quite impressive at 0.6481.

The parameters of these significant variables are quite interesting. First, one notices that the impact of lagged high-tech employment is negative; the more high-technology employment there was in 1990, the less growth in high-tech employment one should expect in that county from 2000 to 2006. This is slightly surprising and supports the reversion to the mean effect in local high technology employment. The negative effect may also be impacted by the fact that the composition of high technology jobs has changed quite a bit from the 1980s to the 2000s. Distance to the nearest metropolitan area with a population of at least 250,000 also has a negative effect which makes sense due to the increasing difficulty in obtaining urban amenities and agglomeration effects. A higher proportion of college graduates in 1990 leads to greater high-technology job growth 10-16 years later. This result agrees with the conventional wisdom that employers in need of highly educated workers are having to follow the workers rather than *vice versa*. Interestingly, higher population in the nearest city has a small negative effect, implying that being near a big city is not a universal positive. It is also worth noting that the presence of a research or land grant university within 50 miles has no significant impact on high technology employment growth.

Moving to consideration of the smoothed state-specific effects, table 3 shows a few summary statistics for the estimated parameters (again using posterior means for point estimates). The range of state-specific effects is quite large. Figure 1 shows that the evolution of the state-specific effects is, indeed, smooth. Since the state-specific intercepts are still being applied to county-level data, these magnitudes suggest that counties in the most favorable state have an expected growth in high technology employment of about 350 more jobs in a 6 year period relative to the least favorable state. That is almost 60 jobs per year, which in non-urban counties is a significant economic development success.

Finally, the county-specific amenity effects are also summarized in table 3. The range of these smoothed effects is enormous, from a minimum of -211 to a maximum of 5,180. Recall that these are parameters on an index with a range of 1-7, so a value of 100 would imply an additional 100 expected new jobs in the county over the study period for each unit increase in the amenity index. Clearly, the magnitude of the amenity parameters says that natural amenities can have a significant role in high technology job growth. Unfortunately, almost all of the amenity parameters have highest posterior density regions that include 0 (the Bayesian equivalent of a confidence interval that covers 0), so we cannot find much evidence for a statistical impact of natural amenities on high technology employment growth. Still, we have uncovered significant spatial heterogeneity in the estimates of amenity effects that cover a wide range of both positive and negative values. There is certainly evidence of a role for amenities in high technology job creation and it is one with an important spatial component.

It is somewhat interesting to note that the median value amenity parameter is -30, suggesting that in over half the counties more natural amenities are (weakly) correlated with a loss in high technology jobs. That certainly is not the effect we expected to find, although in some counties the amenity effect does has the expected positive sign. The conundrum and low statistical precision suggests several possibilities. One is that the amenity index is not sufficient to capture the effect of natural amenities, individual component variables are needed. Perhaps natural beauty, water, and recreation opportunities matter but not temperatures, or different components in the index matter in different regions. A second possibility is that the impact of natural amenities is only felt in certain circumstances such as in the absence of urban amenities or agglomeration effects.

Finally, for comparison purposes, the same model was estimated by maximum likelihood without state-specific effects and with a single, constant amenity parameter. The parameters on the structural variables (the β vector) have very similar estimates in both magnitude and precision, with the same variables being significant. However, the amenity parameter is estimated to be 59.3 and is statistically significant with a t-value of 2.03. This result, in conjunction with the positive effect of lagged share of college graduates, would support our hypothesis of jobs following high-skill workers to locations with high natural amenities. However, the difference between the constant coefficient estimate and the median of the county-specific parameters (-30.0) suggests that the restriction to a single, linear effect introduces an aggregation bias that causes a possibly faulty conclusion.

5. Summary and Conclusions

The question of whether policies exist that can help counties to increase high technology employment is an important one. This paper has tried to address a narrower part of that question: whether natural amenities can be used to entice high-skill workers to a location where the high technology employers will follow. This paper reveals a few insights into both the broad and more narrow question, but much more work is still needed to settle these issues.

On the broader question, it appears that some strategies do not pay off, notably having a major research or land grant university. This suggests that all the locations trying to duplicate Research Triangle (NC) by pushing high technology parks near their major universities are unlikely to be successful. Distance to large metropolitan areas was a negative, when measured to the nearest metro population of at least 250,000. However, we found no effect to the distance to the nearest 500,000 or 1,500,000 population center, suggesting that any urban amenities or services that do matter can be captured by relatively small metro areas. Further, we find a negative effect to increases in the nearest MSA's population. In conjunction, these two results suggest that we want some urban amenities, but not too close and not at the expense of major city inconveniences. We also find empirical support for the reversion to the mean effect in local high technology employment, but do not see how this helps policy makers other than to encourage those who have not succeeded so far.

The clearest prescription we find for policy makers is in the results for our variable on the proportion of college graduates in 1990. Having a population with an extra 1 percent of college graduates is expected to result in about 53 more high technology jobs in our six year study period. Clearly, this has the potential to produce sizable results for small to medium counties if they can find strategies for attracting or producing more college graduates.

On the issue of natural amenities, unfortunately, our results are inconclusive. A fixed coefficient model found significant results in favor of amenities leading to more high technology job growth. However, our spatial model of amenities produced county-specific amenity parameters that were mostly statistical insignificant and just as likely to be negative as positive. However, the state-level effects were important and varying significantly across states and the county-specific amenity effects also have sizable magnitudes from which further refining of the statistical modeling may help to draw statistical inferences on the important and spatially-variable role of natural amenities in high technology employment growth.

References

Adamson, Dwight, David Clark and Mark Partridge, "Do Urban Agglomeration Economies and Amenities Have a Skill Bias?" *Journal of Regional Science* 44 (2004):201-224.

Barkley, David L., John Keith, and Steve M. Smith. *The Potential for High Technology Manufacturing in Nonmetropolitan Areas.* (Western Rural Development Center, Corvallis, Oregon, 1989).

Cairncross, Frances, "The Death of Distance," *The Economist* 336(Sept. 30 1995):5-6.

Deller, Steven, David W. Marcoullier, Donald B.K. English and Victor Lledo. "Regional Economic Growth with a Focus on Amenities," in *Amenities and Rural Development: Theory, Methods and Public Policy*, eds. Green, G., Deller, S., and Marcouiller, D. (Northhampton, MA: Edward Elgar, 1996).

Deller, S.C., T. Tsung-Hsiu, D.W. Marcouiller, and D.B.K. English. "The Role of Amenities and Quality of Life in Rural Economic Growth," *American Journal of Agricultural Economics* 83(2001):352-365.

Desmet, Klaus, and Marcel Fafchamps, "Changes in the Spatial Concentration of Employment across US Counties: A Sectoral Analysis 1972-2000," *Journal of Economic Geography* 5(2005):261-84.

Ferguson, Mark, Kamar Ali, M. Rose Olfert, Mark D. Partridge. "Voting with their Feet: Jobs Versus Amenities," *Growth and Change* 38(2007):77-110.

Florida, Richard, "Bohemia and Economic Geography," *Economic Geography* 2(2002):55-71.

Glaeser, Edward L., Heidi D. Kallal, Jos A. Scheinkman and Andrei Shleifer, "Growth in Cites," *Journal of Political Economy* 100(1992):1126-52.

Glaeser, E.L., J. Kolko, and A. Saiz, "Consumer City," *Journal of Economic Geography* 1(2001):27-50.

Goldstein, Harvey and Joshua Drucker, "The Economic Development Impacts of Universities on Regions: Do Size and Distance Matter?" *Economic Development Quarterly* 20(2006):22-43.

Hecker, D.E., "High Technology Employment: A NAICS Based Update," *Monthly Labor Review Online* (June 2005). Available online at http://www.bls.gov/opub/mlr/2005/07/art6full.pdf.

Kolko, Jed. "The High-Tech Rural Renaissance?: Information Technology, Firm Size and Rural Employment Growth." (1996) Harvard University Working Paper.

Koop, G. and J. L. Tobias, "Semiparametric Bayesian Inference in Smooth Coefficient Models," *Journal of Econometrics* 134(2006):283-315.

Koop, G., and D. J. Poirier, "Bayesian Variants of Some Classical Semiparametric Regression Techniques," *Journal of Econometrics* 123(2004):259-282.

Kotkin, Joel, "Why Wall Street is Losing Out to Forty Acres and a Modem?" *New York Times* (December 27, 1998).

McGranahan, D. "Natural Amenities Drive Rural Population Change." Economic Research Service, United States Department of Agriculture. Agricultural Economic Report No. AER781 (1999).

McGranahan, D. and C.L. Beale, "Understanding Rural Population Loss," *Rural America* 17(2002):2-11.

McGranahan, D., and T. Wojan, "The Creative Class: A Key to Rural Growth," *Amber Waves* (April, 2007). Available at: http://www.ers.usda.gov/AmberWaves/April07/Features/Creative.htm.

Partridge, Mark D, "High-Tech Employment and State Economic Development Policies," *The Review of Regional Studies* 23(Winter 1993):287-306.

Partridge, Mark D., Ray Bollman, M. Rose Olfert and Alessandro Alasia, "Riding the Wave of Urban Growth in the Countryside: Spread, Backwash, or Stagnation," *Land Economics* 83(May 2007):128-152.

Partridge, Mark D., Dan S. Rickman, Kamar Ali and M. Rose Olfert, "Employment Growth in the American Urban Hierarchy: Long Live Distance," *Berkeley Journal of Macroeconomics–Contributions* 8(2008a) in press. Available at: http://www.bepress.com/bejm/vol8/iss1/art10.

Partridge, Mark D., Dan S. Rickman, Kamar Ali and M. Rose Olfert, "The Geographic Diversity of U.S. Nonmetropolitan Growth Dynamics: A Geographically Weighted Regression Approach," *Land Economics* (2008b) in press.

Rosenthal, Stuart S. and William C. Strange, "The Determinants of Agglomeration," *Journal of Urban Economics* 50(2001):191-229.

Saxenian, AnnaLee, *Regional Advantage: Culture and Competition in Silicon Valley and Route 128* (Cambridge, MA: Harvard University Press, 1996).

Zellner, A., *An Introduction to Bayesian Inference in Econometrics* (J. Wiley and Sons, Inc., New York, 1971).

Table 1: High Technology Industries

Industry	4-digit NAICS
Pharmaceutical and medicine manufacturing	3254
Computer and peripheral equipment manufacturing	3341
Communications equipment manufacturing	3342
Seminconductor and other electronic component manufacturing	3344
Navigational, measuring, electromedical, and control instruments manufacturing	3345
Aerospace product and parts manufacturing	3364
Software publishers	5112
Internet publishing and broadcasting	5161
Other telecommunications	5179
Internet service providers and Web search portals	5181
Data processing, hosting, and related services	5182
Architectural, engineering,and related services	5413
Computer systems design and related services	5415
Scientific research and development services	5417

Table 2: Parameter Estimates for Basic Variables

Variable	post. mean	post. std. dev.	ratio
1990 HT jobs	-0.1907	0.0032	-59.33
univ. 50 miles	-28.8034	71.3826	-0.40
km metro 250	-0.6965	0.3695	-1.89
km metro 500	-0.1211	0.5137	-0.24
km metro 1500	-0.2988	0.2796	-1.07
1990 ln(home p)	-3.8561	105.3852	-0.04
1990 college grads	52.8410	5.9487	8.88
1990 near MSA pop	-0.0010	0.0001	-12.75

Table 3: Smoothed Parameter Estimate Summaries

Measure	Value
State effect minimum	-282.26
State effect maximum	74.20
State effect median	-109.35
Amenity effect minimum	-211.73
Amenity effect maximum	5179.7
Amenity effect median	-30.00

Figure 1, State Specific Effects

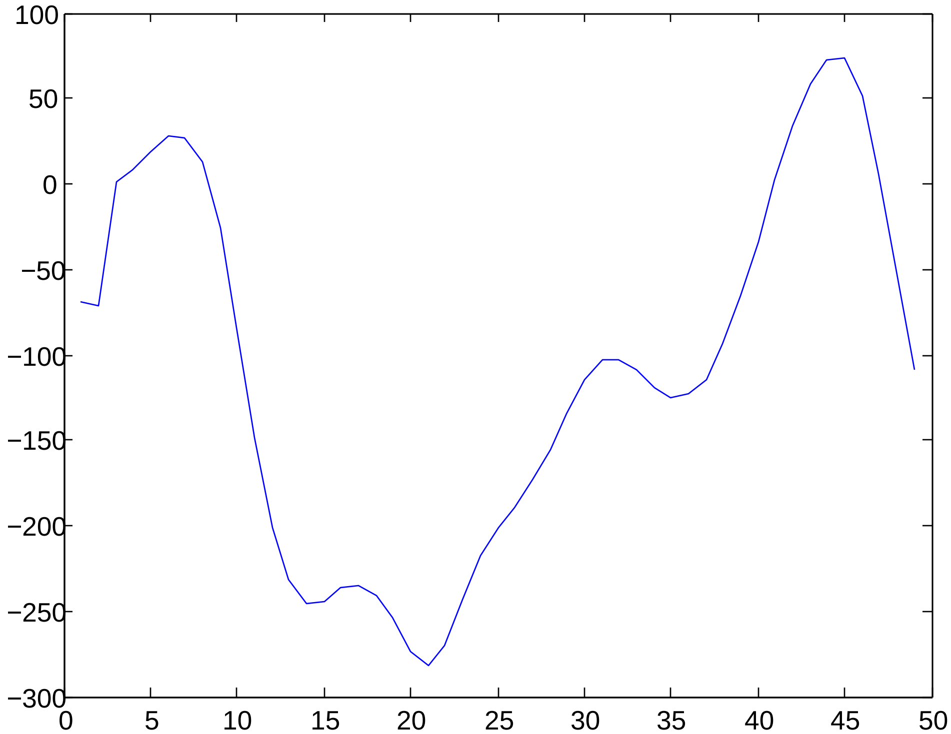

Figure 2. County-Specific Amenity Parameters