Inference Based on Alternative Bootstrapping Methods in Spatial Models with an Application to County Income Growth in the United States

Daniel C. Monchuk, Dermot J. Hayes, John A. Miranowski,
and Dayton M. Lambert

Working Paper 10-WP 507
May 2010

Center for Agricultural and Rural Development
Iowa State University
Ames, Iowa 50011-1070
www.card.iastate.edu

Daniel Monchuk is assistant professor in the Department of Economics at the University of Southern Mississippi. Dermot Hayes and John Miranowski are professors in the Department of Economics at Iowa State University, and Dayton Lambert is assistant professor at the Department of Agricultural and Resource Economics, University of Tennessee.

Questions or comments about the contents of this paper should be directed to Daniel Monchuk, #5072-118 College Dr., Department of Economics, University of Southern Mississippi, Hattiesburg, MS 39406; Ph: (202) 730-9365; Fax: (601) 266-4920; E-mail: dmonchuk@gmail.com.

Abstract

This study examines aggregate county income growth across the 48 contiguous states from 1990 to 2005. To control for endogeneity we estimate a two-stage spatial error model and infer parameter significance by implementing a number of spatial bootstrap algorithms. We find that outdoor recreation and natural amenities favor positive growth in rural counties, densely populated rural areas enjoy stronger growth, and property taxes correlate negatively with rural growth. We also compare estimates from the aggregate county income growth model with per capita income growth and find that these two growth processes can be quite different.

Keywords: county income growth, rural development, spatial bootstrapping.

JEL codes: O18, R11, R58

Inference Based on Alternative Bootstrapping Methods in Spatial Models with an Application to County Income Growth in the United States

1. INTRODUCTION

This study is motivated by the map shown in Figure 1. This map shows the growth in total county income (a close proxy for measure of county gross domestic product) for 48 contiguous states measured in standard deviations from the mean value. It is evident that county income growth during the 1990s has some clear spatial trends. We see that growth in the middle United States tended to be lower than in the rest of the country given the fairly prominent stretch of below-average growth in counties running from eastern Montana and North Dakota southwards to the Texas panhandle. Low growth also stretches across the industrial Midwest. Another prominent spatial trend is the above-average growth experienced in the southeastern region of the United States. Growth appears higher in areas where outdoor amenities are plentiful, such as in the Rocky Mountain region and near large cities such as Minneapolis.

This map stimulates some obvious questions. Is the lack of growth in the midsection and industrial Midwest associated with weather, lack of amenities, or dominance of agriculture? Are there policies that can be adopted at the county or state level to encourage growth? How important are large urban areas for stimulating growth, and are the forces that influence growth fundamentally different in urban and rural counties?

The growth patterns just described and several of the hypothesized explanatory variables have been studied by Carlino and Mills (1987); Khan, Orazem, and Otto

(2001); Deller et al. (2001); and Huang, Orazem, and Wohlgemuth (2002). However, these earlier analyses are based on data compiled prior to 1995. They typically focus on regional growth or on non-metro growth and explain changes in of population, employment, or per capita income rather than on the more comprehensive total county income measure used here. There are a number of shortcomings associated with of measuring economic performance with employment and population growth. Migrants that increase the population without generating significant income may free ride on already stretched local services such as education and health care. Likewise, growth in employment may not generate as much additional government revenue when new jobs are secured by out-of-county residents. The relationship between local employment growth and enhancements in locally provided public goods is highlighted by Renkow (2003), who finds that approximately one-third to one-half of new jobs are secured by non-resident commuters. Furthermore, in sparsely populated rural counties, relative measures of economic performance like wage and per capita income growth may have limited impact on local government revenue. Instead, achieving sufficient scale to allow for the provision of public goods and services might take precedence.

Given the potential shortcomings of these economic performance indicators, it is interesting that relatively little attention has been directed to explaining aggregate measures of economic welfare, such as total county income. A few exceptions are Kusmin, Redman, and Sears (1996); Aldrich and Kusmin (1997); Artz, Orazem, and Otto (2007); and Monchuk et al. (2007). In the first two studies, the variable of interest is total county earnings growth, ultimately a combination of wage and employment growth. The articles by Artz, Orazem, and Otto and Monchuk et al. use aggregate county income growth as the dependent variable but consider only southern or midwestern counties.

This study includes many of the proposed explanatory variables included in these earlier studies while considering potential endogeneity among key explanatory variables—local amenities and property taxes – and controls for dependence between spatial units as well as allowing for heteroskedastic data since, as is clear in Figure 1, the size of the spatial units is not uniform. A number of studies have considered endogenous covariates and heteroskedasticity in spatial process models (e.g., Conley, 1999; Anselin and Lozano-Gracia, 2008; Kelejian and Prucha, 2010; Pinske, Slade, and Brett, 2002; Fingleton and LeGallo, 2008). Most of these studies use asymptotic results to make inferences about relationships between covariates and response variables. Unlike traditional analytic methods using asymptotic results to approximate the sampling distribution, bootstrapping is a method using computer brute force to estimate the sampling distribution of the model parameters. In this paper, we use a bootstrap approach to achieve this objective. Bootstrap procedures are not new in the spatial econometric literature. One of the earliest examples was provided by Anselin (1988). More recent developments have been considered by Fingleton (2008) and Fingleton and Le Gallo (2008), both studying problems arising from endogenous covariates. However, to our knowledge, this is the first time inference based on alternative bootstrap methods and those based on asymptotics for a spatial model with endogenous regressors are compared in an empirical setting.

We outline the steps for conducting inference using the alternative bootstrap methods, including a routine for heteroskedastic data, when faced with a spatial error model with endogenous regressors and apply these procedures in an empirical application examining aggregate income growth in U.S. counties. We also allow for different responses among rural and non-rural counties to key variables like amenities and property taxes, which are amenable to policy. Comparing inference based on bootstrapping with inferences based on asymptotics (i.e., Kelejian and Prucha, 2010), we find that the outcomes are very similar, supporting the idea that bootstrapping can be used effectively when analyzing spatial data. Robust to the alternative methods for conducting inference, we find that outdoor recreation and natural amenities stimulate growth in *rural* counties, more densely populated rural areas enjoy stronger growth, and property taxes correlate negatively with rural growth.

Finally, because total income is not commonly used to gauge economic growth, we compare the results with a set of regressions using per capita income as the dependent variable. We find that while there are many similarities between the aggregate and per capita income growth models, there are a number of salient differences. In particular, whereas rural outdoor and recreational amenities are associated with positive aggregate county income growth, these types of amenities are *negatively* associated with per capita income growth.

2. CONCEPTUAL FRAMEWORK

Total county income (*TCI*) is the product of population and per capita income; total county income growth between the current period (t) and the next (t+1) is ln $[TCI_{t+1} / TCI_t]$. By using total county income growth, we consider the combined effects of population and per capita income growth. In our economic growth model, total county income growth is a function of a number of initial economic, social, and demographic conditions, region-specific characteristics, and government fiscal variables. Each of these variables and their relationship to (regional) county income growth is discussed next in greater detail (see Appendix for data sources and summary statistics).

Population Density, Per Capita Income, Demographics, and Entrepreneurs

Initial *population density* and *per capita income* variables allow us to control for conditional convergence. Which counties are getting richer: those with wealthy residents or the more densely populated ones? Since population densities vary in our cross-section of counties, considering initial population density as an explanatory variable allows us to assess the impact of population concentration on economic growth while holding the extent to which economies grow based on economic well-being of residents constant and vice versa.

The role of human capital is a key variable in many growth models, and counties with high levels of human capital may potentially attract more firms, thereby increasing the demand for labor, which in turn raises wages and county incomes. That human capital has a positive effect on labor demand is documented by Wu and Gopinath (2008), who examine variation in economic development across U.S. counties. However, high levels of human capital in rural counties can lead to a “brain drain,” in which highly educated and skilled rural residents migrate to urban areas where the returns to human capital investment are higher, as documented in the study by Huang, Orazem, and Wohlgemuth (2002). To control for the level of human capital in a county, we use the *share of the population age 25 or older having a college degree or higher* as an initial condition. To build a more complete picture of the effects of human capital, we further include the share completing *some college* as well as those with a *high school degree* only.

Structural changes in agriculture and related agribusinesses (e.g., agricultural mechanization accompanied by large-scale operations and declining labor opportunities and outmigration of the young) have left many rural counties with aging populations and the question of who will maintain the county income base. To examine the effect of initial demographic distributions on county income growth, we include the *percentage of the population age 65 and over,* and to control for "the next generation," we include the *share of the population under age 20,* that rely on local public funding for education and burden local public services without contributing to county revenue.

An issue that often receives considerable attention in policy circles is the role of entrepreneurship in economic growth (see Carree et al., 2002; Low, Henderson, and Weiler, 2005; Wojan and McGranahan, 2007; Wojan, Lambert, and McGranahan, 2007). A problem that arises when analyzing entrepreneurship impact on growth is how to quantify entrepreneurship. Following Acs and Armington (2004), who used a similar measure when studying the relationship between entrepreneurial activity and employment growth in cities in the early 1990s, we use *proprietors per capita*.

Location Characteristics

The role of spatial location and spillovers in the economic growth process has received much attention. Khan, Orazem, and Otto (2001) found that wage growth in neighboring counties complemented population growth in the home county. At the same time, agglomeration diseconomies arising from past manufacturing activity in urban areas (e.g., congestion, higher land values, pollution, higher labor costs) are a reason rural manufacturing experienced significant employment growth in the Midwest in the 1970s and 1980s (Haynes and Machunda, 1987). Wu and Gopinath (2008) report that "remoteness" is a significant factor in explaining variation in economic development across U.S. counties. For rural counties, recent research also considers proximity to different tiers within the urban hierarchy (Partridge et al., 2008, 2009). Given the importance of rural county location relative to market centers, and to control for metropolitan influences (e.g., market access and physical proximity to large metro markets), we include *adjacency to a metro area* and *distance to a metro area*[1] and allow for a non-linear relationship of the latter. We also include a *micropolitan* variable, coded one if the county had a city population greater than 10,000 but also had a total county population of less than 50,000, and zero otherwise, to control for those rural counties that would lack an urban designation but at the same time are not among the most rural of counties. One reason for including this designation is to determine whether medium-sized cities in sparsely populated areas attracted retail business from nearby small towns, as retailers may have consolidated during the study period.

The literature on agglomeration economies and economic spillovers also suggests that the county location and access to major markets play an important role in the growth process (especially in rural areas). To further control for these initial location-specific characteristics, we include the *percentage of the county population that commutes 30 minutes or more* to work. In a study of U.S. cities during the 1990s, Glaeser and Shapiro (2003) found that regions with high levels of commuting by automobile (as opposed to public transport) experienced greater levels of economic growth. Growth enjoyed by commuter counties is one example of a spatial externality. The hypothesis that areas with high levels of commuting activity enjoy additional growth, holding everything else constant, is consistent with Renkow's (2003) findings that as much as half of new jobs created locally are filled by non-resident commuters.

Amenity Index

Previous studies have indicated that amenities and quality of life play an important role in county-level economic growth (McGranahan, 2008; McGranahan, Wojan, and Lambert, 2010). Quality of life is a multi-dimensional concept. Surveys focusing on quality of life attributes find that recreational amenities are important to location decisions, especially for high-technology and information-intensive firms relying on skilled workers. Other studies suggest that positive amenities are capitalized into wages and higher housing values (Roback, 1982; 1988) or land values (Cheshire and Sheppard, 1995), while negative factors such as pollution have adverse impacts on labor market growth (Pagoulatos et al., 2004). To control for outdoor and recreational amenities, we use the same dataset as Deller et al. (2001) and compute an *outdoor recreation and natural amenity index*, which combines a variety of amenities (trails, park characteristics, recreational land and water areas, etc.) from the home plus neighboring counties (see Monchuk et al., 2007). To control for potential Sunbelt effects in southern regions, we also include the average number of *January sun hours*.

Local Government Fiscal Activity

An important decision facing local policymakers is the amount of revenue to collect through county taxes and fees. Local fiscal policy can provide both incentives and disincentives for economic growth. In general, policies designed to induce growth (i.e., better government services) may be offset by taxes (i.e., property taxes) required to pay for those services. Huang, Orazem, and Wohlgemuth (2002) find that local government expenditures on public welfare and highways contribute positively to rural population growth in the Midwest and South. However, they also suggest that the net effect of local fiscal expenditure and county taxation is neutral or even slightly negative on rural working-age populations. To control for the local tax burden, we use initial *property tax revenues per capita*, the predominant source of discretionary local government revenue in rural areas.

Agricultural Influence

Since agriculture has traditionally held the greatest influence in many rural counties, we examine the impact of agriculture's income share within the county on economic growth to address the question: Is dependence on common agriculture good or bad for economic growth? To see how counties with a strong primary agricultural sector have fared, we compute the *share of county income from farming*, which is defined as farm income net of farm employer contributions for government social security divided by total county income.

In the next section, we discuss the estimation details for this type of two-stage, instrumental variable model with spatially correlated errors and the alternative bootstrap methods of conducting inference.

3. EMPIRICAL MODEL

In addition to specifying a typical spatial error model, we also need to consider potential endogeneity issues that arise based on our selection of explanatory variables in our growth model. One method commonly used to control for such simultaneity is through a two-step process in which an instrumental variable, correlated with the endogenous explanatory variable but not the model residuals, is used in a first-step regression to obtain predicted or fitted values. In the second stage, these fitted values are included as an explanatory variable in the regression on the dependent variable, here, county income growth. To conduct inference we adopt bootstrap procedures since they provide a suitable alternative to conducting inference based on asymptotic results provided the data generating process has been specified (Efron and Tibshirani, 1986). Applications of bootstrapping for hypothesis testing and computing confidence intervals are found in Brownstone and Valletta (2001), Efron and Tibshirani (1993), and English (2000), among others. Anselin (1988) and Fingelton (2008) provide the general steps for bootstrap resampling of spatial process models. Our comparison focuses on the spatial error process model, but the algorithms are easily extended to more general spatial process models, including lag, lag-error, or moving average specifications.

Consider a model in which the dependent variable, county income growth, is an nx1 vector of cross-sectional growth rates represented by y, and $\mathbf{X}$ represents an nxk matrix of explanatory variables. Further, suppose there exist potentially unobservable factors that may be correlated across space and are captured in the model error (u), an nx1 vector that contains both a spatial and random error component (ε).The intensity of the unobserved spatial relationship is determined by the parameter λ, and the nature of the spatial relationship is determined by the spatial weights matrix, **W**, an *n*x*n* matrix with zeros along the main diagonal and whose non-zero off-diagonal elements, with row sums equal to unity, represent spatial neighbors. The model may be represented as follows:

$$y = \mathbf{X}\beta + u$$

$$u = \lambda \mathbf{W} u + \varepsilon \tag{1}$$

$$\varepsilon \sim N(0, \sigma^2)$$

If one or more of the variables in the **X** matrix is endogenous, a common method to deal with this involves a two-stage procedure, which effectively purges the endogenous variable's correlation with the residuals. When asymptotic properties are not known or a particular asymptotic result not yet widely accepted, the bootstrap may be used to assess parameter significance. The procedures outlined below describe how each of the bootstrap methods—non-parametric residual, parametric residual, and the paired bootstrap—might be applied to a spatial error model such as that specified in equation (1). In the case of the non-parametric residual bootstrap, the procedure involves sampling with replacement from the residuals of the estimated equation.

Algorithm 1—Non-parametric Residual Bootstrap

Step 1 – Use the selected instruments and the other exogenous variables to predict the endogenous variables and use as an explanatory variable. Obtain an estimate of the parameters $\hat{\beta}$ and $\hat{\lambda}$ using maximum likelihood.

Step 2 – Retrieve the residuals, $\hat{\varepsilon} = \left[I - \hat{\lambda}\mathbf{W}\right] y - \left[I - \hat{\lambda}\mathbf{W}\right] \mathbf{X}\hat{\beta}$.

Step 3 – Loop over the next three steps (3.1–3.3) L times to obtain bootstrap estimates of the model parameters $\{\beta_b, \lambda_b\}_{b=1}^{L}$:

3.1 – Using the vector of residuals from step 2, sample with replacement to construct a vector of bootstrap residuals ε_b.

3.2 – Using the bootstrap vector of residuals from step 3.1, next is computed a vector of pseudo-dependent variables: $y_b = \mathbf{X}\hat{\beta} + \left[I - \hat{\lambda}\mathbf{W}\right]^{-1} \varepsilon_b$.

3.3 – With this new vector of dependent variables, y_b, estimate the following equation with maximum likelihood to obtain bootstrap parameter estimates: $y_b = \mathbf{X}\beta_b + u$ where $u = \lambda_b \mathbf{W}u + \varepsilon$, and $\varepsilon \sim N\left(0, \sigma^2\right)$. Collect and store the estimates β_b and λ_b.

Steps 3.1–3.3 are repeated L times to create an empirical sampling distribution for each parameter. Creating a histogram using the sequence of bootstrap values for each parameter reveals an approximation of its distribution, which need not be symmetric, and can be used to determine whether or not a particular parameter was significantly different from zero at a given level of significance. A (1-α)*100% confidence interval for a particular parameter β_q is found by ordering the L bootstrap estimates from lowest to highest and then removing the lowest (α/2)*L observations from both the lower and upper end of the sequence. Denoting the lowest value in the remaining sample by $\beta^{l}_{q,\frac{\alpha}{2}}$ and the largest remaining value by $\beta^{h}_{q,\frac{\alpha}{2}}$, it follows that a (1-α)*100% confidence interval for β_q is given by $\left[\beta^{l}_{q,\frac{\alpha}{2}}, \beta^{h}_{q,\frac{\alpha}{2}}\right]$. For a particular level of significance α, if this interval does not include zero we would reject the null hypothesis that the parameter β_q is equal to zero.

Algorithm 2—Parametric Residual Bootstrap

Unlike algorithm 1, which does not impose a particular structure on the residuals, in algorithm 2 assumes the residuals follow some known distribution, and the bootstrap routine involves sampling from that particular distribution.

Step 1 – Use the selected instruments and the other exogenous variables to predict the endogenous variables and use as an explanatory variable. Obtain maximum likelihood estimates of the model parameters $\hat{\beta}$, $\hat{\lambda}$, and $\hat{\sigma}^2$.

Step 2 – Loop over the next three steps (2.1–2.3) L times to obtain bootstrap estimates of the model parameters $\{\beta_b, \lambda_b\}_{b=1}^{L}$:

2.1 – Draw randomly from the normal distribution with mean zero and variance $\hat{\sigma}^2$ and create a vector of residuals ε_b.

2.2 – Using the vector of residuals from step 2.1, next compute a vector of pseudo-dependent variables: $y_b = \mathbf{X}\hat{\beta} + \left[I - \hat{\lambda}\mathbf{W}\right]\varepsilon_b$.

2.3 – With this new vector of dependent variables, y_b, estimate the following equation to obtain bootstrap parameter estimates: $y_b = \mathbf{X}\beta_b + u$ where $u = \lambda_b \mathbf{W}u + \varepsilon$ and $\varepsilon \sim N\left(0, \sigma^2\right)$. Collect and store the estimates β_b, and λ_b.

Determining variable significance and inference proceeds in the same manner as indicated in algorithm 1.

Algorithm 3—Paired Bootstrapping

The most general method is the paired bootstrap, which involves sampling with replacement from the data itself rather than the residuals (parametric or non-parametric). Of the three methods, only the paired bootstrap provides consistent estimates if the true model errors are heteroskedastic (Brownstone and Valletta, 2001). In our application we might expect the data to be heteroskedastic since the size of the spatial units is not uniform as is easily verified from Figure 1. In addition, when the spatial weights matrix is asymmetric, the model residuals will always be heteroskedastic (Anselin, 1988). However, the application of the paired method to spatial models requires a modified method that involves transforming the data to "remove" the spatial component by applying a Cochrane-Orcutt type of transformation.

Step 1 – Use the selected instruments and the other exogenous variables to predict the endogenous variables and use as an explanatory variable. Obtain maximum likelihood estimates of the parameters $\hat{\beta}$ and $\hat{\lambda}$.

Step 2 – Using the estimate of the spatial autoregressive term, $\hat{\lambda}$, apply a Cochrane-Orcutt type of transformation to spatially filter the original data to obtain $\tilde{y} = [I - \hat{\lambda}\boldsymbol{W}]y$ and $\tilde{\boldsymbol{X}} = [I - \hat{\lambda}\boldsymbol{W}]\boldsymbol{X}$ $\tilde{\boldsymbol{X}} = [I - \hat{\lambda}\boldsymbol{W}]\boldsymbol{X}$, and use this to create the *n*x*(1+k)* matrix $\tilde{\boldsymbol{P}} = \begin{bmatrix} \tilde{y} & \tilde{\boldsymbol{X}} \end{bmatrix}$.

Step 3 – Loop over the next three steps (3.1–3.2) L times to obtain bootstrap estimates of the model parameters $\{\beta_b, \lambda_b\}_{b=1}^{L}$:

3.1 – Sample with replacement from the matrix $\tilde{\boldsymbol{P}}$ to create a pseudo-dataset, $\tilde{\boldsymbol{P}}_b = [\tilde{y}_b \quad \tilde{\boldsymbol{X}}_b]$, and use this to create a vector of dependent variables and explanatory variables $y_b = \left[I - \hat{\lambda}\mathbf{W}\right]^{-1} \tilde{y}_b$ and $\mathbf{X}_b = \left[I - \hat{\lambda}\mathbf{W}\right]^{-1} \tilde{\mathbf{X}}_b$, respectively.

3.2 – With this new vector of dependent variables, y_b, and explanatory variables $\mathbf{X}_b$, estimate the following equation to obtain bootstrap parameter estimates: $y_b = \boldsymbol{X}_b\beta_b + u$ where $u = \lambda_b\boldsymbol{W}u + \varepsilon$ and $\varepsilon \sim N\left(0, \sigma^2\right)$. Collect and store the estimates β_b and λ_b.

Determining variable significance and inference proceeds in the same manner as indicated in algorithm 1.

The results obtained using the above bootstrapping algorithms are compared with inference based on the following covariance structures: (1) a heteroskedastic-robust 2SLS covariance estimator (Greene, 2003, p. 685), (2) the General Moment heteroskedastic-robust spatial error estimator (SEM-GM-IV) of Kelejian and Prucha (2010), and (3) the heteroskedastic robust spatial error maximum likelihood covariance estimator (SEM-ML-ASY).

The 2SLS robust standard errors are estimated as

$$Cov(\beta_{2SLS}) = \frac{n}{n-k}\left(\hat{Z}'\hat{Z}\right)^{-1} \hat{Z}'\Omega\hat{Z}\left(\hat{Z}'\hat{Z}\right)^{-1}$$

where $\hat{Z}$ contains the set of exogenous variables and the predicted variables of the endogenous covariates, $\Omega = diag(\varepsilon_i^2)$, and $\varepsilon = y - Z\beta_{2SLS}$, noting that the matrix Z contains the actual values of the endogenous covariates. The multiplier $n/(n-k)$ is a degrees of freedom adjustment.

Kelejian and Prucha (2010) and Arraiz et al. (2008) provide details of the steps to estimate the heteroskedastic GM spatial error model (SEM-GM). The covariance matrix of this model is estimated similarly as above, but with the matrix including the exogenous regressors and the predicted values of the endogenous covariates and the residual vector spatially filtered as (respectively) $\hat{Z}^* = (I - \lambda_{GM}W)\hat{Z}$ and $\varepsilon^* = (I - \lambda_{GM}W)(y - Z\beta_{2SLS}^{GM-ERROR})$, which together estimate the covariance of the SEM-GM model with instrumented variables (SEM–GM–IV).

The maximum likelihood spatial error estimator (SEM–ML) with heteroskedastic robust covariance adjusted for the first-stage regression (SEM-ML-ASY) is derived in a similar fashion as the robust SEM–GM–IV, except that the coefficients are estimated using maximum likelihood (Anselin, 1988).

4. RESULTS

Although we use a 15-year growth rate and our explanatory variables are initial conditions, some of our key explanatory variables may suffer from endogeneity in a forward-looking environment. In particular, are amenities endogenous in the sense that places expected to grow have more *a priori* development? Likewise property taxes could also be endogenous in a forward-looking environment. In our application, these variables are also interacted with a rural dummy variable, which adds a degree of difficulty when conducting routine tests for endogeneity, suitability of instruments, and identification. To avail ourselves to a greater variety of tools for models with endogenous variables, diagnostic tests are conducted using the 2SLS estimator and we use the implications of these tests to determine the selection of instruments when estimating the spatial models.

When constructing the instruments, we follow Kelejian and Prucha (1998). Here we consider two possible spatial weights matrices, W and M (both $n \times n$), which are based on the four nearest neighbors and within a fifty-mile-radius contiguity rule, respectively. Representing the potentially endogenous variables by matrix X ($k \times n$), the candidate instruments we consider are (WX, W^2X, MX, M^2X, WMX). When considering the potential endogeneity of the amenity variable, for example, this leaves us with a total of 10 instruments (five for the amenity variable itself and five for the rural interacted amenity term). Using this set of instruments, we proceeded to conduct both the Durbin and Wu-Hausman tests for endogeneity for (i) amenity and rural interacted amenity term; and (ii) property taxes and rural interacted property tax term. Based on the Durbin and Wu-Hausman[2] test results, we conclude amenities are endogenous while property taxes per capita are exogenous.

Having found the amenity index potentially endogenous, the next step determines the appropriateness and strength of our chosen instruments and to ensure overidentification restrictions are satisfied. Since we have two endogenous regressors (i.e., amenity index and rural interacted amenity term), we use a measure based on Shea (1997) to determine the strength of the instruments. Whether overidentifying restrictions are satisfied is based on the Sargan and Basmann tests. When including all 10 spatial lag variables as discussed above (five for each amenity variable) we generally find that these are relatively strong instruments but that the overidentification restrictions were not satisfied. In the end, and after some sensitivity analysis, we settled on a set of five instruments MX_a, WX_{ra}, W^2X_a, M^2X_{ra}, and WMX_{ra} where X_a and X_{ra} represent the amenity and amenity by rural county interactions, respectively, which provide reasonably strong instruments as well as satisfying overidentification requirements (Table 1).

The results in Table 2 include a full complement of state controls. The first set of estimates correspond to the non-spatial, 2SLS specification in which approximately 60% of the variation in total county income growth over 1990-2005 compared to slightly less than 48% when state controls are omitted (results not reported to conserve space). When estimating the spatial models, SEM-ML and SEM-GM-IV, the spatial weights matrix is constructed using the four nearest neighbors. In the second column, the parameter estimates correspond to the SEM-ML model, which explains nearly 66% of the variation in county income growth. Under the SEM-ML heading, the columns ASY, NR, PR, and PAIR correspond to significance under the asymptotic heteroskedastic robust, non-parametric residual, parametric residual, and paired bootstrap methods, respectively. The third column of estimates corresponds to the SEM-GM-IV model, which explains slightly less than 60% of variation in income growth. To conserve space when inferring parameter significance, rather than reporting the bootstrap confidence intervals, significance for the "*," "**," and "***" represent the *smallest* of the 10%, 5%, and 1% significance levels, corresponding to 90%, 95%, and 99% confidence intervals. A similar coding is used for reporting parameter significance for the 2SLS, SEM-ML-ASY, and SEM-GM-IV models.

A quick glance at the results in Table 2 indicates how similar the methods of inferring significance under the SEM-ML model are, at least from a statistical significance point of view. There are, however, some minor discrepancies. For example, the amenity index-by-rural county interaction is significant at the 1% level under both types of residual bootstrap, but it is significant at the 5% level under the paired bootstrap and ASY estimates. The coefficient on January sun hours is significant at the 10% level under each of the bootstrap algorithms, but it is not statistically significant under the robust covariance structure (ASY). Comparing with the SEM-GM-IV results, the January sun hours coefficient is significant at the 5% level. To contrast further with the SEM-GM-IV results, the marginal effect for January sun hours and the rural interacted amenity term tend to be slightly larger than those of the SEM-ML although most of the other variables are of similar magnitude and significance.

In addition to demonstrating the viability of the bootstrap in spatial models, comparing outcomes from the alternative models and inference methods gives us greater confidence in the discussion to follow. In what might be loosely considered a standard convergence result, we find that counties with a high per capita income and high population density in 1990 experienced lower growth in total county income than those that did not have these attributes. However, a high population density in rural counties was associated with *greater* growth, suggesting that the growth dynamic for population-dense rural counties is such that growth is faster than what would be otherwise expected.

Counties with a large proportion of older individuals in 1990 and those with a high percentage of young people in 1990 grew more slowly than would otherwise have been the case. Given outside opportunities for young people to move away from counties with stagnant local economies, it makes sense that those left with a larger proportion of older people do poorly. However, this does not explain why counties with young people did not fare well. In part, people younger than 20 may not be a predictor of growth because they are not yet a productive (income earning) part of the community, but rather an indicator of a higher public education burden coupled with a migration risk to urban areas for college and employment opportunities. Remember, the age group 20-65 (the excluded group) is associated with growth because they are both the main income earning and taxpaying component of rural communities. Further, if these young people remain in the county and do not pursue higher education and/or become proprietors, they apparently earn below-average incomes.

As mentioned earlier, population-dense counties did poorly, and contrary to what may have been expected, counties adjacent to a metropolitan county did not fare significantly better and distance was not found to be an important factor. However, those counties with a high proportion of commuters did grow at a faster rate, ceteris paribus, over the period3. Our admittedly crude measure of entrepreneurship, the number of proprietors per capita, was also associated with higher growth. Although farmers are typically classified as proprietors, the model was able to separate the generally negative influence of the agriculture sector from the positive influence of this entrepreneurial variable.

January sunshine was correlated with county income growth consistent with more individuals moving or retiring to the Sunbelt. The country-wide measure of outdoor recreation and natural amenity index did not appear to contribute to growth. However, when this term is interacted with a rural indicator variable, it is a positive and significant variable suggesting the marginal impact of amenities in rural counties is one of enhancing growth[4]. Likewise, rural counties with relatively higher population densities did well, especially in contrast to non-rural, high-density counties. Possibly, metro counties with a high level of amenities had already exhausted these features by 1990. Among those non-Sunbelt counties that remained rural in 1990, those endowed with amenities appeared to have generated growth. For rural counties, a viable policy could be adding, expanding, and improving existing recreational amenities to attract population thereby generating increases in aggregate income.

Several measures of the size and relative importance of local government were available to us. These included relative salaries of local government workers, total county tax burden, and intergovernmental transfers. We report only one of these variables, per capita property taxes, because these variables are highly correlated (especially with rural interaction terms), and all provide essentially the same result but with the linkage between property taxes and residents and businesses being more tangible. When applied to the entire dataset, the impact of per capita property taxes is positive and significant at the 1% level. However, when applied only to rural counties, the property tax variable is negative and statistically significant[5]. One hypothesis is that in order for rural counties with a declining population base to cover large fixed costs associated with education, roads, and judicial systems they have increased local taxes with an unintended consequence of making retaining and attracting residents more difficult. Non-rural counties typically have much greater population density and higher property values. They can more easily spread the fixed costs of running the county across many more individuals, and as such they can offer more public goods to residents. There is clearly a minimum population level that is required to effectively fund the fixed costs associated with running a county, and some rural counties now appear to be below that critical level.

In rural areas we find a negative relationship between property taxes and aggregate income growth. To reduce the negative effect of local property taxes, one alternative for rural counties is to shift the burden to another tax base or revenue source (e.g., from buildings to land if permitted by state, from county revenue to state cost-sharing revenue) that does not deter in-migration or outside investment.

A reviewer pointed out that because our data ends in 2005 we may have missed an inversion in the relative rankings of counties. Areas of the country that attracted people also appear to have suffered more from the mortgage crisis. Areas that had been dependant on grain-based agriculture may have benefitted from the high energy prices through biofuel expansion and associated grain price increases. Energy-producing counties will also have benefited from renewed interest in domestic energy while counties that depend on long-distance commutes have probably fared less well. Analysis of these trends will require data that is much more current than presently available and require additional time to determine the permanence of the mortgage crisis and energy price increases.

To highlight the similarities and differences in the growth process for *aggregate* vs. *per capita* income growth, we repeat the analysis above but with per capita income as the dependent variable (Table 3). Comparing the results in Table 3 with those in Table 2, we see there are similarities such as the positive and significant impact of share of college degree and the negative impact of initial per capita income (convergence). However, there are some notable differences, in particular, those concerning amenities. We find that amenities are negatively associated with per capita income growth, and this contrasts with the positive impact of amenities in rural counties when considering aggregate income growth. At first glance these results may seem contradictory but are upon reflection, consistent with theory. The negative impact of amenities on per capita income growth in Table 3 appears to us to be a standard Roback-type result whereby people will be willing to accept a lower income in exchange for more amenities. The positive effect of rural amenities is consistent with amenities attracting people where aggregate county growth is driven by population growth. This simple comparison highlights that the two growth processes need to be interpreted differently, and that goals and objectives need to be clearly defined to ensure effective formulation and application of policy.

5. CONCLUSIONS

This study updates and expands on earlier studies explaining the forces driving economic activity at the county level. Our study is unique in that it considers aggregate county income growth, seldom the focus of empirical research but a highly relevant for county planners, which captures both income and population changes in a way that mimics county gross domestic product. Using counties in the lower 48 states we explain total county income growth from 1990-2005 using an extensive list of explanatory variables, including amenities, agricultural dependence, rural/non-rural comparisons, and rural county proximity and distance to metro centers, and property taxes. Because of our focus on rural economic growth we did not include additional detail on industrial structure.

Our diagnostic analysis shows a potential endogenous relationship between aggregate county income growth and our amenity variable and leads us to implement a two-stage instrumental variable approach. As an alternative to inference based on asymptotics, we implement three common bootstrapping methods, including one for heteroskedastistic data, in the context of a spatial error model with endogenous regressors. Comparing inference based on bootstrapping with those based on asymptotics, we find that the results are quite similar. Finding such similarities leads us to further advance bootstrapping can be a viable alternative to conducting routine tests of hypothesis in spatial models so that practitioners need not be held up in their analysis in situations where asymptotic results are computationally complex, not well established, or are altogether nonexistent.

The results suggest that growth in total county income in the United States was lower in counties that had the following: larger per capita income in 1990, a higher population density in 1990, a higher proportion of older individuals, and a higher proportion of population under 20 years of age. Counties with a heavy dependence on agriculture grew more slowly in general. Counties that grew at a faster rate had: a high proportion with a college degree, close to a metropolitan area, a high proportion of commuters, and relatively more sunshine in January.

In light of our results, it might be reasonable to expect that adding, expanding, and improving existing recreational amenities in rural counties can generate increases in aggregate income through a combination of attracting employment or population. In rural areas we find a negative relationship between property taxes and aggregate income growth. To reduce the negative effect of high local property taxes, county government officials might explore alternative revenue sources (e.g., shifting property tax base, using other taxes, cost-sharing arrangement with state and federal governments) that do not deter in-migration or outside investment.

Finally, we also compare estimates from the aggregate county income growth model with per capita income growth and find that these two growth processes can be quite different. In particular, whereas rural outdoor and recreational amenities are found to be associated with positive aggregate county income growth, these types of amenities are negatively associated with per capita income growth.

ENDNOTES

[1] Since the distance to a metro county for a metro county is zero, we use the following to allow for taking of logs:

$$ldist^* = \begin{cases} \ln(dist) \text{ if } dist > 0 \\ 1 \text{ if } dist=0 \end{cases} \quad \text{and} \quad dm = \begin{cases} 0 \text{ if } dist > 0 \\ 1 \text{ if } dist=0 \end{cases}$$

where *dist* is the distance in miles to a metropolitan county and is equal to zero if a county is classified as metro.

[2] These endogeneity tests were conducted using the non-spatial form of the model, allowing us to draw upon a larger set of diagnostic tools for assessing the appropriateness of our instruments. For the amenity variables (*amenity* and *amenity*rural*) the value of the Durbin test statistic was 24.214 (p-val<0.00) and the Wu-Hausman test was 11.92 (p-val<0.00). For the property tax variables (*property taxes* and *property taxes*rural*) the value of the Durbin test statistic was 0.47 (p-val=0.79) and the Wu-Hausman test was 0.23 (p-val=0.79). In each of the tests the null hypothesis is that variables are exogenous.

[3] Although the results are not reported here, it is interesting to note that when share of commuters is excluded we find that distance from a metro is negatively correlated with growth suggesting much of the distance effect is captured by commuters.

[4] Notice that we can only conclude that the marginal effect of amenities in rural counties is positive in relation to metro counties since estimating the net effect would require jointly considering the sum of the parameter estimates for amenity index and the rural interaction. In the 2SLS version of the model the combined amenity impact is

approximately 0.09 (= -0.36 + 0.1269) and found to be significant at the 1% level where inference is based on the Delta method.

[5] In the 2SLS version of the model the net effect of property taxes is -0.034 (= 0.054 - 0.088) and found to be significant at the 1% level where inference is based on the Delta method.

6. REFERENCES

Acs, Zoltan J. and Catherine Armington. 2004. “Employment Growth and Entrepreneurial Activity in Cities,” *Regional Studies,* 38, 911–927.

Aldrich, Lorna and Lorin Kusmin. 1997. *Rural Economic Development. What Makes Rural Communities Grow?* Washington, D.C.: U.S. Department of Agriculture, ERS, Agr. Info. Bull. No. 737, September.

Anselin, Luc. 1988. *Spatial Econometrics: Methods and Models.* Dordrecht: Kluwer.

Anselin, L. and N. Lozano-Gracia. 2008. “Error in Variables and Spatial Effects in Hedonic House Price Models of Ambient Air Quality,” *Empirical Economics* 34, 5–34.

Arraiz, I., D.M. Drukker, H.H. Kelejian and I.R. Prucha. 2008. “A Spatial Cliff-Ord Ty Model with Heteroskedastic Innovations: Small and Large Sample Results,” CESifo Working Paper No. 2485.

Artz, Georgeanne M., Peter F. Orazem and Daniel M. Otto. 2007. “Measuring the Impact of Meat Packing and Processing Facilities in Nonmetropolitan Counties: A Difference-in-Differences Approach,” *American Journal of Agricultural Economics,* 89, 557–570.

Brownstone, David and Robert Valletta. 2001. “The Bootstrap and Multiple Imputations: Harnessing Increased Computing Power for Improved Statistical Tests,” *Journal of Economic Perspectives,* 15, 129–141.

Carlino, Gerald A. and Edwin S. Mills. 1987. “The Determinants of County Growth,” *Journal of Regional Science,* 27, 39–54.

Carree, Martin, Andre van Stel, Roy Thurik and Sander Wennekers. 2002. “Economic Development and Business Ownership: An Analysis Using Data of 23 OECD Countries in the Period 1976–1996,” *Small Business Economics,* 19, 271–290.

Cheshire, Paul, and Stephen Sheppard. 1995. “On the Price of Land and the Value of Amenities,” *Economica,* 62, 247–267.

Conley, T.G. 1999. “GMM Estimation with Cross-sectional Dependence,” *Journal of Econometrics* 92: 1–45.

Deller, Steve C., Tsung-Hsiu Tsai, David W. Marcouiller and Donald B.K. English. 2001. “The Role of Amenities and Quality of Life in Rural Economic Growth,” *American Journal of Agricultural Economics,* 83, 352–365.

Efron, Bradley and Robert Tibshirani. 1986. “Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy,” *Statistical Science,* 1, 54–75.

———. 1993. *An Introduction to the Bootstrap.* New York: Chapman and Hall.

English, D.B.K. 2000. “Calculating Confidence Intervals for Regional Economic Impacts of Recreation by Bootstrapping Visitor Expenditures,” *Journal of Regional Science,* 40, 523-529.

Fingleton, B. 2008. "A Generalized Method of Moments Estimator for a Spatial Model with Moving Average Errors, with Application to Real Estate Prices," *Empirical Economics* 34, 1, 35–57.

Fingleton, B. and J. Le Gallo. 2008. "Estimating Spatial Models with Endogenous Variables, a Spatial Lag and Spatially Dependent Disturbances: Finite Sample Properties." *Papers in Regional Science* 87, 3, 319–339.

Glaeser, Edward L. and Jesse M. Shapiro. 2003. "Urban Growth in the 1990s: Is City Living Back?" *Journal of Regional Science,* 43, 139–165.

Haynes, Kingsley E. and Zachary B. Machunda. 1987. "Spatial Restructuring of Manufacturing and Employment Growth in the Rural Midwest: An Analysis for Indiana," *Economic Geography,* 63, 319–333.

Huang, Tzu-Ling, Peter F. Orazem and Darin Wohlgemuth. 2002. "Rural Population Growth, 1950-1990: The Roles of Human Capital, Industry Structure, and Government Policy," *American Journal of Agricultural Economics,* 84, 615–627.

Kelejian, H.H. and I.R. Prucha. 1998. "A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances," *Journal of Real Estate Finance and Economics,* 17, 99-171.

———. 2010. "Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances," forthcoming in *Journal of Econometrics*.

Khan, Romana, Peter F. Orazem and Daniel M. Otto. 2001. "Deriving Empirical Definitions of Spatial Labor Markets: The Roles of Competing versus Complementary Growth," *Journal of Regional Science,* 41, 735–756.

Kusmin, Lorin D., John M. Redman and David W. Sears. 1996. *Factors Associated with Rural Economic Growth. Lessons from the 1980's.* Washington, DC: U.S. Department of Agriculture, Technical Bulletin No. 1850, September.

Low, S., J. Henderson and S. Weiler. 2005. "Gauging a Region's Entrepreneurial Potential." *Economic Review*, 3rd Quarter, 61–88.

McGranahan, D.A. 2008. "Landscape Influence on Recent Rural Migration in the U.S." *Landscape and Urban Planning* 85, 2, 228–240.

McGranahan, D.A., T.R. Wojan and D.M. Lambert. 2010. "The Rural Trifecta: Outdoor Amenities, Creative Class, and Entrepreneurial Context." Forthcoming in *Journal of Economic Geography*.

Monchuk, Daniel C., John A. Miranowski, Dermot J. Hayes and Bruce A. Babcock. 2007. "An Analysis of Regional Economic Growth in the U.S. Midwest," *Review of Agricultural Economics*, 29, 17–39.

Pagoulatos, Angelos, Stephan J. Goetz., David I. Debertin and Tulin Johannson. 2004. "Interactions between Economic Growth and Environmental Quality in U.S. Counties," *Growth and Change,* 35, 90–108.

Partridge, M., D. Rickman, K. Ali and M. Olfert. 2008. "Lost in Space: Population Growth in the American Hinterlands and Small Cities," *Journal of Economic Geography* 8, 727–757.

———. 2009. "Agglomeration Spillovers and Wage and Housing Cost Gradients across the Urban Hierarchy," *Journal of International Economics* 78, 126–140.

Pinske, J., M.E. Slade and C. Brett. 2002. "Spatial Price Competition: A Semiparametric Approach," *Econometrica* 70: 1111–1153.

Renkow, Mitch. 2003. "Employment Growth, Worker Mobility, and Rural Economic Development," *American Journal of Agricultural Economics,* 85, 503–513.

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life," *Journal of Political Economy,* 90, 1257–1278.

———. 1988. "Wages, Rents, and Amenities: Differences among Workers and Regions." *Economic Inquiry* 26, January, 23–41.

Shea, J. 1997. "Instrument Relevance in Multivariate Linear Models: A Simple Measure," *Review of Economics and Statistics*, 26, 393–415.

Wojan, T., R, D. Lambert and D. A. McGranahan. 2007. "Emoting with Their Feet: Bohemian Attraction to Creative Milieu," *Journal of Economic Geography* 7, 711–736.

Wojan, T.R. and D.A. McGranahan. 2007. "Ambient Returns: Creative Capital's Contribution to Local Manufacturing Competitiveness," *Agricultural and Resource Economics Review* 36, 1, 133–148.

Wu, Junjie and Munisamy Gopinath. 2008. "What Causes Spatial Variations in Economic Development in the United States?" *American Journal of Agricultural Economics,* 90, 392–408.

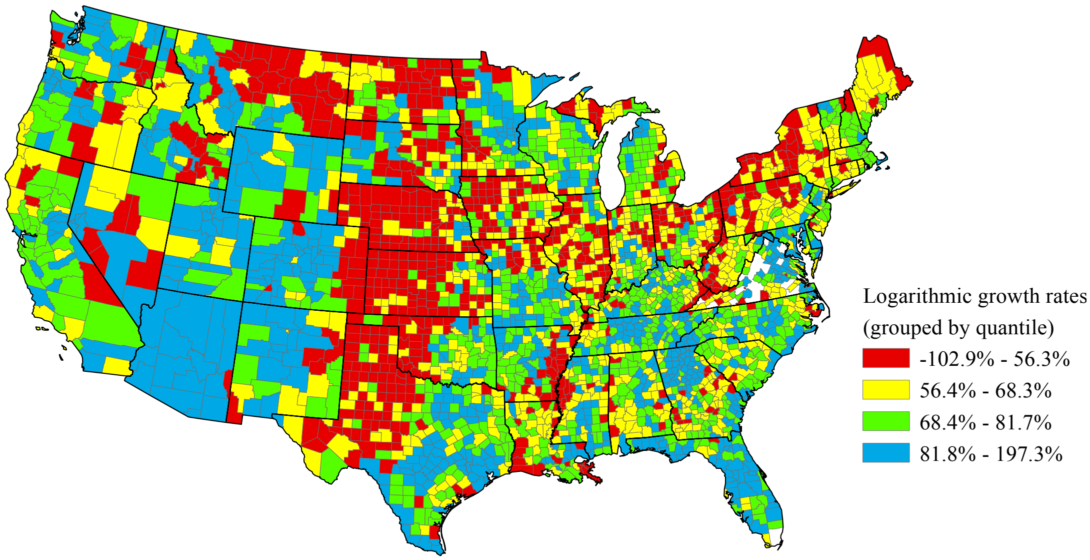

FIGURE 1: Aggregate County Income Growth, 1990–2005.

TABLE 1: Summary of Diagnostic Tests

Tests of Endogeneity		Strength of instruments		Tests of overidentifying restrictions	
Durbin	16.175 (p-val=0.00)	Variable	Shea's partial R-square	Sargan:	1.843 (p-val=0.61)
Wu-Hausman	7.940 (p-val=0.00)	Amenity	0.5777	Basmann:	1.800 (p-val=0.61)
		Amenity*rural	0.5986		

TABLE 2: Regression Results: County Income Growth 1990–2005, State Effects Included

Variable	2SLS		SEM-ML	Method of Inference				SEM-GM-IV	
				ASY	NPR	PR	Pair		
Instrumented									
Normalized Combined Amenity Index	-0.0359		0.0048					-0.0203	
Rural Normalized Combined Amenity Index	0.1269	***	0.0656	**	***	***	**	0.1012	***
(ln) Per Capita Income 1990	-0.2783	***	-0.2307	***	***	***	***	-0.2541	***
(ln) Population per Square Mile 1990	-0.0287	***	-0.0339	***	***	***	***	-0.0311	***
Percent of Pop. 65+ 1990	-0.0088	***	-0.0095	***	***	***	***	-0.0091	***
Percent of Pop. under Age 20 1990	-0.0063	***	-0.0061	***	***	***	***	-0.0062	***
Percent of Pop. 25+ with College Degree 1990	0.0052	***	0.0045	***	***	***	***	0.0048	***
Percent of Pop. 25+ with High School 1990	-0.0015		-0.0014					-0.0016	
Percent of Pop. 25+ with Some College 1990	0.0070	***	0.0065	***	***	***	***	0.0067	***
Micropolitan Variable (city 10-50K and total pop<50K)	-0.0297	***	-0.0228	***	**	**	***	-0.0268	***
Adjacent to a Metropolitan Area (=1) 1993	0.0043		0.0006					0.0024	
(ln) Distance to Nearest Metro 1990	-0.0030		-0.0005					0.0029	
Square [(ln) Distance to Nearest Metro 1990]	0.0009		-0.0012					-0.0007	
Metropolitan County (1990)	-0.0889		-0.0901					-0.0795	
Percent of Pop. Commuting 30+ Mins. 1990	0.0104	***	0.0097	***	***	***	***	0.0102	***
(ln) Proprietors per Capita 1990	0.1047	***	0.1028	***	***	***	***	0.1033	***
(ln) January Sun Hours	0.0945	***	0.0558		*	*	*	0.0795	**
(ln) Property Taxes Per Capita 1992	0.0544	***	0.0443	***	***	***	***	0.0499	***
Share of County Income from Farming 1990	-0.6550	***	-0.7021	***	***	***	***	-0.6836	***
Rural (ln) Population per Square Mile 1990	0.0929	***	0.0841	***	***	***	***	0.0887	***
Rural (ln) Property Taxes Per Capita 1992	-0.0879	***	-0.0761	***	***	***	***	-0.0824	***
Constant	1.2393	***	1.4028	***	***	***	***	1.2886	***
Spatial Error Interaction (Lambda)			0.3800	***	***	***	***	0.3496	***
R-Square	0.6023		0.6555					0.5926	

The level of significance at the 1%, 5%, and 10% levels are represented by ***,**, and ** respectively.

Notes: (i) inference in 2SLS based on heteroskedastic robust standard errors; (ii) see text for details on the GMM-SEM-IV model; (iii) inference in SEM-ML-ASY model based on heteroskedastic robust asymptotic, NPR (non-parametric residual bootstrap), PR (parametric residual bootstrap), Pair (paired bootstrap). See text for more details.

TABLE 3: Regression Results: Per Capita Income Growth 1990–2005, State Effects Included

Variable	2SLS		SEM-ML	ASY	NPR	PR	Pair	SEM-GM-IV	
				Method of Inference					
Instrumented									
Normalized Combined Amenity Index	-0.0614	**	-0.0362		***	***	*	-0.0491	*
Rural Normalized Combined Amenity Index	0.0372		0.0070					0.0232	
(ln) Per Capita Income 1990	-0.2285	***	-0.2283	***	***	***	***	-0.2263	***
(ln) Population per Square Mile 1990	0.0037		0.0076					0.0050	
Percent of Pop. 65+ 1990	-0.0016		-0.0023	*	**	**	*	-0.0019	
Percent of Pop. under Age 20 1990	-0.0025	**	-0.0026	**	**	**	**	-0.0027	**
Percent of Pop. 25+ with College Degree 1990	0.0036	***	0.0033	***	***	***	***	0.0034	***
Percent of Pop. 25+ with High School 1990	-0.0006		-0.0006					-0.0007	
Percent of Pop. 25+ with Some College 1990	0.0005		0.0009					0.0006	
Micropolitan Variable (city 10-50K and total pop<50K)	-0.0152	***	-0.0149	***	**	**	**	-0.0159	***
Adjacent to a Metropolitan Area (=1) 1993	-0.0176	***	-0.0126	**	**	**	**	-0.0151	***
(ln) Distance to Nearest Metro 1990	-0.0232		-0.0249					-0.0201	
Square [(ln) Distance to Nearest Metro 1990]	0.0024		0.0028					0.0021	
Metropolitan County (1990)	-0.0568		-0.0624					-0.0526	
Percent of Pop. Commuting 30+ Mins. 1990	0.0022	***	0.0017	***	***	***	***	0.0020	***
(ln) Proprietors per Capita 1990	0.0250	**	0.0419	***	***	***	***	0.0317	***
(ln) January Sun Hours	0.0260		0.0075					0.0199	
(ln) Property Taxes Per Capita 1992	-0.0001		-0.0040					-0.0023	
Share of County Income from Farming 1990	-0.4569	***	-0.4525	***	***	***	***	-0.4557	***
Rural (ln) Population per Square Mile 1990	0.0082		0.0050					0.0069	
Rural (ln) Property Taxes Per Capita 1992	-0.0051		-0.0019					-0.0038	
Constant	1.2366	***	1.3805	***	***	***	***	1.2839	***
Spatial Error Interaction (Lambda)			0.2820	***	***	***	***	0.2452	***
R-Square	0.3318		0.3813					0.3361	

The level of significance at the 1%, 5%, and 10% levels are represented by ***,**, and ** respectively.

Notes: (i) inference in 2SLS based on heteroskedastic robust standard errors; (ii) see text for details on the GMM-SEM-IV model; (iii) inference in SEM-ML model based on heteroskedastic robust asymptotic, NPR (non-parametric residual bootstrap), PR (parametric residual bootstrap), Pair (paired bootstrap). See text for more details.

APPENDIX TABLE A.1: Summary Statistics

Variable	Source	Mean	Std. Dev.
Dependent Variable (n=3014)			
(ln) Total County Income Growth 1990-2005	REIS	0.70	0.22
(ln) Per Capita Income Growth 1990-2005	REIS	0.58	0.12
Independent Variables			
Per Capita Income 1990 ($000's)	REIS	15.27	3.50
Population per Square Mile 1990	REIS, Authors' est.	167.08	1323.60
Percent of Pop. 65+ 1990	1990 Census	14.99	4.33
Percent of Pop. under age 20 1990	1990 Census	29.86	3.50
Percent of Pop. 25+ with College Degree 1990	1990 Census	13.39	6.38
Percent of Pop. 25+ Completed High School 1990	1990 Census	16.44	4.53
Percent of Pop. 25+ with Some College 1990	1990 Census	34.42	6.09
Micropolitan Variable (city 10-50K and total pop<50K)	ERS, USDA	0.10	0.30
Adjacent to a Metropolitan Area 1993	ERS, USDA	0.32	0.47
Distance to a Metro (2555 counties)	1990 Census, Authors' est.	56.21	37.81
Metropolitan County 1990	1990 Census	0.15	0.36
Percent of Pop. Commuting 30+ Mins. 1990	1990 Census	16.38	6.39
Proprietors per Capita 1990	REIS	0.12	0.06
Share of County Income from Farming 1990	REIS	0.05	0.08
January Sun Hours	ERS	151.75	33.28
Amenity Index	NORSIS, Authors' est.	0.43	0.38
Property Taxes Per Capita 1992	1992 CofG, REIS	552.45	438.61
Rural Interacted Terms (2226 rural counties)			
(ln) Population per Square Mile 1990	REIS, Authors' est.	35.62	41.87
Amenity Index	NORSIS, Authors' est.	0.39	0.36
(ln) Property Taxes Per Capita 1992	1992 CofG, REIS	548.72	463.92

Notes: REIS = Bureau of Economic Analysis's Regional Economic Information Systems; NORSIS = National Outdoor Recreation Supply Information Survey (NORSIS); ERS, USDA = Economic Research Service, US Department of Agriculture. As documented in Deller et al. (2001), NORSIS dataset documents outdoor recreational amenities as of 1998 but does not indicate the year in which the amenities were established. Adjacency to a metro is determined by using 1993 ERS rural-urban continuum codes 4, 6, and 8. Micropolitan variable is based on 1993 ERS urban influence codes 3, 5, and 7.CofG=US Census of Governments.