Hierarchy and Spatial Contagion: Population in American Cities Between 1990 and 2010

Elizabeth A. Dobis*[,a], Michael S. Delgado[b], Raymond J.G.M. Florax, and Peter Mulder[c]

[a] *Northeast Regional Center for Rural Development, Penn State University, University Park, PA, United States*
[b] *Department of Agricultural Economics, Purdue University, West Lafayette, IN, United States*
[c] *Department of Spatial Economics, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands*

April 3, 2018

Abstract

In this paper, we separate the effects of spatial contagion and urban hierarchy, using central place theory as a theoretical basis. Our analysis includes both the hierarchical relationship among cities of differing sizes and the continuous nature of proximity to other cities through the novel use of a spatially-lagged hierarchical model. Our unique dataset of urban areas from 1990 to 2010 includes all but the smallest rural communities. We find that large urban areas are characterized by urban agglomeration and spatial competition for population, while small urban areas are characterized by both spatial complementarity and position in the urban hierarchy.

JEL Codes: R11, R12, R15, R23
Keywords: population, urban hierarchy, spatial lag, hierarchical linear models

*Please address correspondence to Elizabeth A. Dobis, 7F Armsby Building, NERCRD, Penn State University, University Park, PA 16802; Email: edobis@psu.edu.
The authors dedicate this article to the memory of Dr. Raymond J. G. M. Florax, who passed away unexpectedly on March 1, 2017. Raymond was instrumental in the conceptualization of this project and contributed greatly to early versions of this manuscript.

1 Introduction

Urban systems are characterized by the flow of social and economic activity among cities (de Vries, 1990), which in turn reflects spatial contagion and urban hierarchical structure. Spatial contagion is the idea of interaction between urban areas that is the result of social and economic activity among proximate locations, and urban hierarchy is the idea that urban areas of differing sizes can be ranked according to the variety and specialization of goods and services available. The urban hierarchy is defined such that an urban area higher up on the hierarchy has a wider variety of, and more specialized, goods and services available. Typically, such urban areas are relatively larger areas, offering varied and specialized goods and services to a larger market area. Intuitively, the concepts of spatial contagion and urban hierarchy are intertwined but distinct, yet theoretical and empirical analyses do not typically explore these forces as separate phenomena. In this paper, we build an empirical model that separates the effects of spatial contagion and urban hierarchy and apply this model to a unique dataset of urban areas in the United States to better understand how these two forces have shaped the population level in recent decades.

Traditionally, theoretical models of urban systems and location have not paid much attention to spatial contagion (varied examples include Christaller, 1933; Alonso, 1964; Henderson, 1974; Krugman, 1991; Fujita et al., 1999). In the models of Christaller (1933) and Lösch (1943), traditional central place theory models, space is imagined as a flat, featureless plane on which the urban system is characterized by hierarchical tiers of urban areas (or cities) of different sizes. Urban areas are homogeneous within each hierarchical tier, so spatial spillovers (contagion) do not arise. More recently, Fujita et al. (1999) and Tabuchi and Thisse (2011) have developed theoretical general equilibrium models that show how an urban system can evolve away from an evenly distributed population in one-dimensional space as the number of manufactured goods in the economy increases. Empirically, the notion of urban hierarchy has been modeled via the incremental distance from a city to the nearest city that is higher up on the urban hierarchy (Dobkins and Ioannides, 2001; Partridge et al., 2008, 2009b), and these analyses show that urban population is inversely related to position on the urban hierarchy. Non-hierarchical empirical analyses of the urban system have included the relative proximity among cities by indicating neighbors within distance bands (e.g., Bosker and Buringh, 2017) and through spatial econometric weighting structures (e.g., González-Val, 2015).[1]

A related vein of literature focuses on city size distributions and their changes over time. These analyses use kernel densities (e.g., Kim, 2000; Black and Henderson, 2003; Le Gallo and Chasco,

[1] Additionally, a whole body of theoretical literature exists that seeks to explain growth and transformation patterns across cities from (changes in) the size of transport costs in relation to the (positive and negative) externalities inherent to agglomeration and congestion forces in cities (see, for example, Abdel-Rahman, 1988, 1990; Abdel-Rahman and Fujita, 1993; Anas and Xiong, 2003; Coşar and Fajgelbaum, 2016; Desmet and Rossi-Hansberg, 2014; Duranton and Turner, 2012; Fujita, 1988, 2012; Glaeser and Kohlhase, 2004; Helsley and Strange, 2014; Rossi-Hansberg, 2005). Duranton and Puga (2014) and Redding and Rossi-Hansberg (2017) provide in-depth reviews.

2008), Zipf's Law distributional analyses (e.g., Black and Henderson, 2003; Le Gallo and Chasco, 2008; González-Val, 2010), growth regressions that look for convergence and divergence (e.g., Black and Henderson, 2003; Desmet and Rappaport, 2017), and Markov chains (e.g., Black and Henderson, 2003; Le Gallo and Chasco, 2008; González-Val and Lanaspa, 2016). These studies have not come to a consensus on whether city size growth is random or non-random and whether the city size distribution is Pareto or log-normal, however González-Val (2010) suggests that this lack of consensus may be due to the use of truncated datasets that restrict analysis to only the largest cities as well as differences in the geographic definition of what is a city (e.g., metropolitan statistical area or incorporated place). Yet, one story that emerges is a shift in the city size distribution (of the United States) away from non-random population growth patterns towards random growth patterns, as a result of structural economic change.

Recently, other authors (e.g., Burger and Meijers, 2016; Capello, 2000; McCann and Acs, 2011) have emphasized the relevance of the network and regional embeddedness of cities within the urban system, rather than hierarchy. Their argument finds its support in empirical observations that undermine a straightforward link between city size, agglomeration economies, and (urban) economic growth. Most notably, the urban system in Europe is predominantly a polycentric pattern of small and medium-sized cities, with second-tier cities often outperforming first-tier cities in economic growth rates. This suggests that factors that are often independent of city size, such as urban infrastructure, institutional capacity, and industry composition, may be more conducive to acting as an engine of growth, compared to just a city's size (Camagni et al., 2015; Castells-Quintana, 2017; Frick and Rodríguez-Pose, 2017; McCann and Acs, 2011). This led to a re-appraisal of connectivity in urban networks as a potential substitute for agglomeration benefits and to the re-introduction of Alonso's concept of "borrowed size" (Meijers et al., 2016), which captures the notion that small and medium-sized cities may internalize the agglomeration economies of nearby larger cities, while avoiding their agglomeration costs. Van Meeteren et al. (2016) argue that research assessing the link between urban size and economic performance needs to come to grips with the fact that urban boundaries are inevitably contentious and regionalizing. However, in this recent wave of studies that emphasize the benefits originating from the functional relationships among cities, the separate effects of contagion and hierarchy usually remain implicit.

In this paper, we disentangle the separate effects of hierarchical structure and spatial heterogeneity in the urban system. In contrast to central place theory, we allow for physical features in the spatial plane, which introduces contagion into our model. Heterogeneous natural features, allow for a more realistic, less regular urban structure to emerge: physical location now matters, and location is no longer only relative. Rivers, oceans, and lakes provide transportation diversity and cost advantages, while mountains and escarpments act as barriers to transportation. Cities are located in traditionally advantageous places, which leads to spatial heterogeneity in the layout of the urban system, even in recent years, because of infrastructural inertia and agglomeration economies

(Polèse and Shearmur, 2004). Consequently, this leads to spatial heterogeneity and a less regular, more realistic urban structure.[2]

We use our approach to analyze the urban system of the United States between 1990 and 2010, focusing on population levels in each urban area as our dependent variable of interest. We first construct a unique dataset of United States urban areas, which capture both agglomerated economic activity and the built extent of urban locations, that allows for the delineation between urban and rural areas without relying on political boundaries, for all but the smallest communities. Our use of urban areas is quite novel in the literature. Perhaps the most common unit of analysis for urban systems is metropolitan and micropolitan statistical areas (e.g., Dobkins and Ioannides, 2001; Neal, 2011), while others have used incorporated places (legal government units) and their boundaries (e.g., González-Val, 2010). The advantage of urban areas over these other approaches is that urban areas are constructed via population requirements based at the census block group and census tract levels, providing a more nuanced delineation between urban and rural areas. Further, the urban area delineation does not rely on legal boundaries, the definition of which can vary across states (US Census Bureau, Geography, 1994a).

We then model the hierarchy of the urban system by classifying the urban areas in our dataset into three tiers following the urban tier classification of Overman and Ioannides (2001) and Dobkins and Ioannides (2001).[3] Our urban hierarchy results in a top tier of central place urban areas that contains world regional and national nodal centers, a middle tier of regional urban areas that contains regional and subregional nodal centers (as in Knox and McCarthy, 2005), and a bottom tier of urban areas that contains all other areas. We refer to our highest tier areas as central place nodes, our middle tier areas as regional nodes, and our lowest tier areas as urban areas. We then combine this tier-classification with a spatial lag hierarchical linear regression model that allows us to capture spatial contagion through the spatial lag structure in which proximate neighbors influence each other and urban hierarchy through the hierarchical model structure (specifically, covariates constructed at different hierarchical levels and random error components from each hierarchical tier). We estimate our model using the full sample of urban areas, as well as a subset of urban areas with a minimum population of 50,000, and explore both the cross-sectional and panel structure of the data.

We find that contagious forces have heterogeneous effects throughout the urban system. There is a dispersive effect on metropolitan areas due to competition and an agglomerative effect on urban

[2]Note that spatial contagion need not manifest as a complementary force between cities, and may be a result of competition. Bosker and Buringh (2017) show that early European sites located next to (0–20km away) other existing cities are less likely to have evolved into an urban area, and that, since the seventeenth century, sites a moderate distance (20–100km) away from existing cities have expanded due to co-location benefits and decreasing transportation costs.

[3]Classifying cities into a hierarchy using population size has precedence in the literature, e.g., Borchert (1967) and Partridge et al. (2008, 2009a), who use population as a proxy for the centrality of a city's markets. Others, like Lorenzen and Andersen (2009) and Maliszewski and Ó hUallacháin (2012), completely avoid tier classifications by using the rank-size distribution of cities to represent the urban hierarchy.

areas due to complementarity. Contagion is also not static; it varies over time. Further, explicitly including hierarchy in the model structure influences the magnitude, direction, and significance of explanatory regional market area variables. Utilizing a dataset that truncates the urban system by excluding small urban areas has a similar effect. This is because the large number of urban clusters (they comprise nearly 86 percent of the urban system) overwhelms the effect of metropolitan areas in the urban system. In addition, we find four general trends in the urban system related to contagion, hierarchy, and whether the urban system is truncated. First, urban areas with a larger initial population experience higher population growth, for both metropolitan and urban areas. Second, the mix of goods and services available only affects urban areas. In keeping with the assumptions of central place theory, the goods and services available are not linked to population size for metropolitan areas. Third, there is only weak evidence that natural amenities that may attract residents affect the population level of metropolitan and urban areas. Finally, the region's share of employment in manufacturing has a negative effect on urban areas, as one would expect after the structural transformation, but a positive effect on metropolitan areas because of inertia in population. Together, these results emphasize that explicitly including contagion and hierarchy in the model structure matters when studying an urban system.

The structure of the paper is as follows. Section 2 describes our spatially-lagged hierarchical model and our empirical specification. Section 3 discusses the data and provides descriptive statistics, while our results and main discussion are in Section 4. Section 5 provides our concluding remarks.

2 Empirical Setup and Estimation

2.1 Spatial Hierarchical Model

We use a spatial hierarchical model to account for both spatial contagion and the hierarchical relationship among cities of differing sizes. The spatial hierarchical model we use is an extension of the traditional spatial lag model into the three-level random intercepts hierarchical model:

$$
\begin{aligned}
y_{tkji} &= \beta_{tkj0} + \lambda_t \sum_{k=1}^{N} \sum_{j=1}^{M_k} \sum_{f=1}^{L_{kj}} w_{kjf} y_{tkjf} + \sum_{p=1}^{P} \beta_{00p} x_{p,tkji} + e_{tkji} \\
\beta_{tkj0} &= \pi_{tk00} + \sum_{q=1}^{Q} \pi_{0q} z_{q,tkj} + \mu_{tkj} \\
\pi_{tk00} &= \gamma_{t000} + \alpha_{tk}
\end{aligned}
\tag{1}
$$

where the location indices are defined such that $i, f = 1, \ldots, L_{kj}$ represents an urban area, $j = 1, \ldots, M_k$ represents a regional (small) market area, $k = 1, \ldots, N$ represents a central place (large) market area, and $t = 1, \ldots, T$ represents the time period. There are $L = \sum_{k=1}^{N} \sum_{j=1}^{M_k} L_{kj}$ urban areas, $L_k = \sum_{j=1}^{M_k} L_{kj}$ urban areas in central place k, and L_{kj} urban areas in region j of central place k. Similarly, there are $M = \sum_{k=1}^{N} M_k$ regions and M_k regions in central place k. The total sample size, defined across the smallest spatial unit (urban areas) and time, is H.

The dependent variable, y_{tkji}, is the population level in urban area i of regional market area j of central place market area k at time t. Similarly, the p urban area independent variables, $x_{p,tkji}$, and q regional market area independent variables, $z_{q,tkj}$, are for urban area i of regional market area j of central place market area k at time t. There are $p = 1, \ldots, P$ independent variables in level one (urban area) and $q = 1, \ldots, Q$ in level two (regional market area), resulting in $P + Q$ independent variables in the model. Each level has a time-varying error that is assumed to be mean zero, *i.i.d.* normal: $e_{tijk} \sim \mathcal{N}(0, \sigma_e^2)$, $\mu_{tkj} \sim \mathcal{N}(0, \sigma_\mu^2)$, and $\alpha_{tk} \sim \mathcal{N}(0, \sigma_\alpha^2)$.

We estimate both panel and cross-sectional versions of this model, noting that the cross-sectional model is a version of the panel specification such that the time dimension is fixed at one. To construct these models, the data are first stacked according to the spatial hierarchy; the stacked spatial hierarchies are then stacked over time to create the panel structure. This data structure facilitates our modeling of cross-sectional spatial contagion through a spatial proximity matrix (the neighbor matrix) that, in our models, is time invariant. For added flexibility, we allow the spatial contagion parameter to vary across time, denoted λ_t.[4]

2.2 Maximum Likelihood Estimation

2.2.1 The Model Parameters

We use a maximum likelihood estimator for our three-level spatial lag hierarchical model (Baltagi et al., 2015).[5] The maximum likelihood estimator allows for spatially unbalanced data (i.e., differing numbers of individuals in each group) and generalizes to an arbitrary number of hierarchical levels. To obtain the estimator, first rewrite the model by nesting the higher-level equations within the base-level urban area equation to generate a single linear equation that includes the random effects:

$$y_{tkji} = \gamma_t + \lambda_t \tilde{y}_{tkji} + \mathbf{x}_{tkji}\boldsymbol{\beta} + \mathbf{z}_{tkj}\boldsymbol{\pi} + u_{tkji} \tag{2}$$

where $u_{tkji} = \alpha_{tk} + \mu_{tkj} + e_{tkji}$, $\mathbf{x}$ is the $(1 \times P)$ vector of urban area independent variables, $\boldsymbol{\beta}$ is the $(P \times 1)$ vector of urban area coefficients, $\mathbf{z}$ is the $(1 \times Q)$ vector of regional market area independent variables, $\boldsymbol{\pi}$ is the $(Q \times 1)$ vector of regional market coefficients, and the spatial lag variable is:

$$\tilde{y}_{tkji} = \sum_{k=1}^{N} \sum_{j=1}^{M_k} \sum_{f=1}^{L_{kj}} w_{kjf} y_{tkjf}. \tag{3}$$

[4]We require the constraint $1/\epsilon_{min} < \lambda_t < 1$, for all t to ensure spatial stationarity.

[5]The Baltagi et al. (2015) maximum likelihood estimator we use is based on the Antweiler (2001) double-nested unbalanced panel estimator. See also Baltagi et al. (2014) for an alternative, nested spatial two-stage least squares estimator based on Kelejian and Prucha (1998).

Following Baltagi et al. (2015), the log-likelihood function of pooled observations is:

$$\ln \mathscr{L} = -\frac{1}{2}H\ln(2\pi) - \frac{1}{2}\ln|\mathbf{\Omega}| + \ln|\mathbf{A}| - \frac{1}{2}\mathbf{u}'\mathbf{\Omega}\mathbf{u} \tag{4}$$

where $\mathbf{u} = \mathbf{A}\mathbf{y} - \boldsymbol{\gamma} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\pi}$, $\mathbf{A} = \mathbf{I}_H - (\boldsymbol{\lambda} \otimes \mathbf{I}_L)\mathbf{W}$, $\mathbf{I}_L$ is an $(L \times L)$ identity matrix, and $\mathbf{I}_H$ is an $(H \times H)$ identity matrix. The resulting neighbor matrix, $\mathbf{W} \equiv bdiag(\mathbf{W}_1, \mathbf{W}_2, ..., \mathbf{W}_T,)$, is a block diagonal matrix of the $(L \times L)$ neighbor matrices $\mathbf{W}_1, \ldots, \mathbf{W}_T$, and the time-varying contagion parameters $\lambda \equiv bdiag(\lambda_1, \lambda_2, ..., \lambda_T)$. The structure of the variance-covariance matrix, $\mathbf{\Omega}$, follows from Antweiler (2001). Letting $\mathbf{R}_\mu$ and $\mathbf{R}_\alpha$ denote $(H \times M)$ and $(H \times N)$ regional and central place membership matrices, and defining $\mathbf{J}_\mu = \mathbf{R}_\mu\mathbf{R}_\mu'$ and $\mathbf{J}_\alpha = \mathbf{R}_\alpha\mathbf{R}_\alpha'$, the variance-covariance matrix can be written $\mathbf{\Omega} = E[\mathbf{u}\mathbf{u}'] = \sigma_e^2[\mathbf{I}_H + \rho_\mu\mathbf{J}_\mu + \rho_\alpha\mathbf{J}_\alpha]$ where $\rho_\mu = \sigma_\mu^2/\sigma_e^2$ and $\rho_\alpha = \sigma_\alpha^2/\sigma_e^2$.[6] We maximize the log-likelihood function numerically making use of the analytic gradients of the model, imposing the constraints $|\lambda_t| < 1$, $\sigma_e^2 > 0$, and $\rho_\mu, \rho_\alpha \geq 0$.

2.2.2 The Marginal Effects

It is well known that in spatial lag models, independent variables have both direct and indirect effects, with the indirect effects occurring through the spatial contagion process. To recover these effects—through which we empirically recover the effects of spatial contagion on urban population—we look to the marginal effects of the regressors on y. For the p^{th} urban area level independent variable, the marginal effects are given by:

$$\begin{aligned}\frac{\partial y}{\partial x_p} &= \hat{\beta}_p(\mathbf{I}_H - \hat{\boldsymbol{\lambda}}\mathbf{W})^{-1} \\ &= \hat{\beta}_p \begin{bmatrix} (\mathbf{I}_L - \hat{\lambda}_1\mathbf{W}_1)^{-1} & & & \\ & (\mathbf{I}_L - \hat{\lambda}_2\mathbf{W}_2)^{-1} & & \\ & & \ddots & \\ & & & (\mathbf{I}_L - \hat{\lambda}_T\mathbf{W}_T)^{-1} \end{bmatrix}.\end{aligned} \tag{5}$$

from which the direct effects are computed as the average of the diagonal elements of this matrix and the total effects are computed as the average across all elements in the matrix. The indirect effect is recovered as the total effect minus the direct effect. We bootstrap these marginal effects to conduct inference: we draw 1,000 sets of random parameters from a multivariate normal distribution that takes the estimated parameters as its mean and estimated variance-covariance matrix as its

[6] Putting the pieces together, the log-likelihood function becomes:

$$\ln \mathscr{L} = -\frac{1}{2}\left[H\ln(2\pi\sigma_e^2) + \sum_{t=1}^{T}\left\{\ln|\mathbf{I}_t - \lambda_t\mathbf{W}_t| + \sum_{k=}^{N_t}\left\{\ln\theta_{tk} + \sum_{j=1}^{M_{tk}}\left\{\ln\theta_{tkj} + \frac{V_{tkj}}{\sigma_e^2} - \frac{\rho_\mu}{\theta_{tkj}}\frac{U_{tkj}^2}{\sigma_e^2}\right\} - \frac{\rho_\alpha}{\theta_{tk}}\frac{U_{tk}^2}{\sigma_e^2}\right\}\right\}\right],$$

where $V_{ikj} = \sum_{i=1}^{L_{tkj}} u_{tkji}^2$, $\theta_{tkj} = 1 + \rho_\mu L_{tkj}$, $U_{tkj} = \sum_{i=1}^{L_{tkj}} u_{tkji}$, $\theta_{tk} = 1 + \rho_\alpha\phi_{tk}$, $U_{tk} = \sum_{j=1}^{M_{tk}} \frac{U_{tkj}}{\theta_{tkj}}$, $\phi_{tk} = \sum_{j=1}^{M_{tk}} \frac{L_{tkj}}{\theta_{tkj}}$, and the residual is $u_{tkji} = y_{tkji} - \lambda_t\tilde{y}_{tkji} - \gamma_t - \mathbf{x}_{tkji}\boldsymbol{\beta} - \mathbf{z}_{tkj}\boldsymbol{\pi}$.

variance-covariance matrix and, for each set, compute the marginal effects. We conduct inference based on this empirical distribution of marginal effects.

2.3 Empirical Specification

The dependent variable is the natural logarithm of the population of an urban area, measured in decadal intervals. In many studies, urban growth and urbanization are often explained by an assortment of factors related to physical geography, natural amenities, market structure and/or economic factors, demographic variables, and first and second nature geographies (Nzaku and Bukenya, 2005; Partridge et al., 2008; Bosker and Buringh, 2017; Olfert et al., 2012). We develop our empirical specification based on key tenets of central place theory—economic distance, market-based city attributes, and the notion of a goods hierarchy—so that we can directly measure how hierarchy and spatial contagion affect the urban population. As described, spatial contagion is modeled via the econometric structure (the neighborhood matrix and spatial lag setup), and hierarchy is incorporated through the nested data structure. All independent variables are lagged one time period (i.e., one decadal unit).

Specifically, the empirical specification is:

Urban Area:

$$X = \begin{cases} Pop_{t-1} \\ GoodsIndex_{t-1} \\ WaterProximity_{t-1} \\ Ruggedness_{t-1} \\ TemperateClimate_{t-1} \end{cases}$$

Regional Market Area:

$$Z = \begin{cases} AggregateIncome_{t-1} \\ MfgEmploymentShare_{t-1} \\ SvcEmploymentShare_{t-1} \\ RegionalRuralLandProportion_{t-1}. \end{cases}$$

In the urban area equation, $GoodsIndex$ refers to a goods centrality index that captures both the variety and balance of products in each urban area. The goods centrality index is an important source of heterogeneity in the urban system and is how we model the goods hierarchy. $WaterProximity$ measures the distance to the nearest large body of water: either the Great Lakes or an ocean. $Ruggedness$ is defined as either the difference of maximum and minimum elevation within an urban area or as an ordered categorical scale from 1 to 9 such that the lowest value indicates flat plains and the highest value indicates rugged mountains. We explore models that use both measurements of ruggedness. $TemperateClimate$ is measured as either the average yearly temperature in an urban area or as a temperature discomfort index that measures discomfort as a weighted average of the annual temperature deviation from the mildest (warmest) winter and mildest (coolest) summer.[7]

[7] Specifically, the temperature discomfort index comes from Zheng et al. (2009) and is calculated as:

$$TempDiscomfort_{it} = \sqrt{(JanT_{it} - max(JanT_t))^2 + (JulT_{it} - min(JulT_t))^2}.$$

The larger an urban area's deviation from the mildest winter and summer temperatures, the larger the value of the index and the greater the discomfort. These three variables reflect heterogeneity in natural amenities among urban areas.

The variables in the regional market area equation capture regional heterogeneity. Two variables, $MfgEmploymentShare$ and $SvcEmploymentShare$, capture the proportion of the economy in the regional market area that is devoted to manufacturing and services. Economic heterogeneity in the regional market area is measured by $AggregateIncome$, which represents purchasing power and is similar to the income bands of market power in Partridge et al. (2008). Finally, since we create our regional market area variables by aggregating data from urban areas, we include $RegionalRuralLandProportion$ to account for the omitted factors that correspond to the surrounding rural areas. The central place market area equation contains only an intercept and random disturbance.

3 Data

3.1 Unit of Analysis and Control Variables

Defining the Urban Area We define our unit of analysis, the "urban area," in two separate ways that correspond to differences in measurement of the United States census data over recent decades. Starting in the 1950 Census the United States Census Bureau measured urbanity as an "urbanized area" that was constructed from both the area of built environment and legal boundaries (i.e., incorporated places) (US Census Bureau, Geography, 1994b,c). Since the 2000 Census, urbanity was measured differently, primarily to avoid tying urbanity to legal boundaries. Specifically, in the 2000 Census the United States Census Bureau introduced the "urban area" that consisted of two subgroups: urbanized areas and urban clusters. These urbanized areas and urban clusters are constructed from census blocks and tracts based on population size and density to create an urban footprint that spans residential, commercial, and non-residential urban land uses (US Census Bureau, 2011; US Census Bureau, Census History Staff, 2013; US Census Bureau, Geography, 2012), allowing areas to be classified as urban areas based only on population size and density and not on legal municipal status. The "urbanized area" is similar to the earlier 1950 definition in that it requires an area to have at least 50,000 inhabitants, and an "urban cluster" is an urban area that has at least 2,500 inhabitants (and fewer than 50,000).

Our analysis is based on both the urbanized area (which we refer to as metropolitan area, MA) and the urban cluster (which we refer to as urban area, UA). We carefully match geographic areas over time to construct a panel dataset of consistent urban area geographies. The advantage of the metropolitan area is that our final dataset spans three decades, from 1990 to 2010, while the urban area dataset is only available for the decades from 2000 to 2010. The advantage of the

where $JanT$ and $JulT$ is January and July temperature.

urban area dataset is that it incorporates numerous smaller urban areas that are not counted in the metropolitan area dataset; indeed, 88 percent of the urban areas in our urban area dataset are excluded in the metropolitan areas dataset. Figure 1 provides a visual comparison of the geographic areas included in both datasets: it is clear from panel (a) of the figure that the metropolitan area dataset spans all major urban areas in the United States, while the urban area dataset additionally includes many smaller urban areas across the country.

Data Sources Data for the urban area and metropolitan area datasets come from several sources. United States Census Bureau decennial census population and income data and corresponding boundary files come from the National Historic Geographic Information System (NHGIS). However, because urban areas are a sub-county census-defined aggregate geography, there is a paucity of auxiliary data available to support econometric analysis. Hence, we obtain data for additional explanatory variables from various governmental agencies and academic research groups and use ArcGIS to create proximity measures and urban area values for these data. We use small grid cells of land surface forms from the United States Geological Survey (USGS), land elevation from the National Oceanic and Atmospheric Administration (NOAA), and temperature normals from the PRISM Climate Group to create aggregated urban area data. The boundary files for the location of the Atlantic and Pacific Oceans and the Great Lakes come from the Commission for Environmental Cooperation, and establishment and employment data for individual establishments come from the National Establishment Time Series (NETS) database and are aggregated to the urban area level using each establishment's geographic location.

The regional and central place market areas are constructed by assigning each urban area to the closest regional or central place node (i.e., the nearest middle- or high-tier urban area) by network distance (i.e., road distance). The regional market areas are then assigned to the central place node analogously, and hinterlands are created for each urban area using Thiessen polygons which are aggregated to create each market area polygon. The regional and central place market areas are shown in Figure 2, and we note that the regional market area variables are constructed based on the aggregation of all urban area values that belong within that regional area. We specify our neighbor matrix by computing the pairwise distance between any two places and imposing a 400 kilometer inverse distance cutoff. Thus we define an urban area's neighbors as those locations within a 400 kilometer radius.

3.2 The Goods Hierarchy

We use the Shannon-Weiner Index as the base of our measure of the goods hierarchy because this index is designed to reflect both the variety and balance of available goods and services in a particular urban area (Stirling, 1998; Maignan et al., 2003). This index is defined as:

$$SW_i = -\sum_{s=1}^{S} \frac{E_{si}}{E_i} \ln\left(\frac{E_{si}}{E_i}\right) \tag{6}$$

where E_{si} is the number of establishments in sector s of urban area i and E_i is the total number of establishments in urban area i. Here, balance is reflected by the share of each industry in the total population, capturing the extent to which the population of establishments is equally spread across sectors; balance is captured in the ratio E_{si}/E_i. Variety refers to the total number of industries, and is accounted for via the summation. The value of this index is essentially a weighted function of the frequency of an industry, with the weight being an inverse and decreasing (via the logarithm) function of commonality. Intuitively, the Shannon-Wiener index captures the nature of goods within the urban area via a continuum from rare to common (Maignan et al., 2003). At the same time, this index does not account for the availability of rare products at higher levels in the hierarchy, and so we modify the Shannon-Weiner Index to include the range of goods available within the central place market area.

We define the rarity of sector s in the central place market area m as $R_s = 1 - (E_{sm}/E_m)$, reflecting the fact that as the number of establishments in the sector decreases relative to the total number of establishments (i.e., as the sector becomes more rare) this index approaches a value of one. This produces a continuous measure of industry rarity at the central place market level that is bounded between 0 and 1; we use this index of rarity to weight the industries included in the Shannon-Wiener Index. The resulting index is our modified goods centrality index:

$$C_i = -\sum_{s=1}^{S} R_s \frac{E_{si}}{E_i} \ln\left(\frac{E_{si}}{E_i}\right) \tag{7}$$

that measures how centralized an urban area is based on the diversity and rarity of the goods and services available to consumers.

We use the NETS database from 1990 and 2000 to determine product diversity in each urban area. The geocoded location and North American Industry Classification System (NAICS) code of each establishment are used to construct the diversity variables for metropolitan and urban areas. We construct the diversity indices for 4-digit (i.e., industry group) classifications in sectors 48 (transportation and warehousing) through 81 (other services). This sectoral restriction aligns with the 12 NAICS service sectors that were used to develop the North American Product Classification System (NAPCS) (US Census Bureau, Business & Industry, 2012). Developers chose these sectors because they produce the majority of products and include the most dynamic industries in Canada, Mexico, and the United States. We focus on these 12 sectors because many of the products in these sectors are produced and purchased in the same urban area. This is less likely to occur with sectors such as utilities (22) or manufacturing (31-33).

3.3 Descriptive Statistics

Tables 2 and 3 contain descriptive statistics for the MA and UA data by metropolitan/urban areas, regional market areas, and central place market areas, and Tables 4 and 5 contain descriptive statistics by tier classification. It is apparent from these tables that there are substantial differences between the MA and UA datasets—the main difference between datasets is the inclusion of numerous smaller urban areas in the UA dataset. In the following paragraphs, we highlight several important insights that come from our descriptive statistical analysis; for additional descriptive statistics that are not reported here, contact the authors.

In 2010, the average population in the metropolitan areas was 574,971 people, while urban areas had an average size of 77,589. This difference comes from the (approximately) 2,800 urban areas that are not included in the MA data. These urban areas are all classified as low-tier areas, bringing the average size of a low-tier urban area in the UA data down to 39,234 inhabitants. The average size of a low-tier metropolitan area in the MA data was 263,939 inhabitants, over 6.7 times the average size in the UA data. Despite the difference in the average size of urban and metropolitan areas in the datasets, both show a similar population growth rate: 11.6 percent in the MA data and 11.8 percent in the UA data. The MA data shows a faster growth rate in average population of 20.2 percent between 1990 and 2000.

More than 90 percent of urban and metropolitan areas are located on either flat plains, smooth plains, or irregular plains, and the average yearly temperature is between 13 and 14 degrees Celsius. The average distance from a higher tier city is also comparable, 166 kilometers for the metropolitan areas and 206 kilometers for urban areas. The similarity in the averages of these geographic variables is unsurprising, as the locations of urban and metropolitan areas are spread throughout the country.

The average centrality indices for the MA and UA data are quite different, with a value of 13.215 for metropolitan areas and 7.456 for urban areas. This variation is particularly noticeable when comparing middle-tier regional nodes to low-tier urban areas. In the MA data, the difference in the centrality index between regional nodes and low-tier metropolitan areas is 1.801, whereas for the UA data it is 7.442. Urban areas with a small centrality index have a much smaller range of services available, and the difference in centrality index for the low-tier urban areas again emphasizes the importance of including small urban areas in city system and urban hierarchy analyses.

The regional market area variables capture economic conditions. The average real aggregate income within regional market areas in the MA data is 132.6 billion dollars, while the same figure in the UA data is 16.3 billion dollars larger. The manufacturing and service employment shares are very similar between the two datasets, with around 12 percent of the population employed in manufacturing and about 60 percent employed in services in 2000. The employment share data from the metropolitan area data indicates that the share in manufacturing has declined and the share in services has increased since 1990, concurrent with a continuing structural transformation in the economy. Regional variation emerges when we analyze the data by market areas. The maps of the regional market areas in Figure 2 show that market areas in the eastern United States tend to be smaller than the market areas in the western United States. These differences indicate that driving distance within a market area is considerably different across the country.

4 Results

4.1 Hierarchical Spatial Lag Regression: Parameter Estimates

We estimate four separate versions of our hierarchical spatial lag regression model, two for each of the MA and UA datasets. All models are structured to include both urban area and regional market area levels, as well as the three-tiered regression error structure. For each dataset, Model 1 measures land ruggedness as the difference in elevation and includes temperature, and Model 2 measures land ruggedness with categorical land type indicators and eschews temperature for the temperature discomfort index.[8]

Tables 6 and 7 report the parameter estimates for the MA and UA estimates. Several facts are readily apparent. First, there is a clear persistence in population levels across time, with the decadal lag in population taking a large positive value (0.97-0.99) and being statistically significant in all four models. This estimate is, of course, as expected because populations do not move much (geographically) over time; it is worth drawing attention to this estimate because we are reminded that the inclusion of the lag of population controls for initial urban area size (in terms of population) and other estimates are to be interpreted relatively. Second, the variance parameters σ_e^2 and ρ_α (the central place market area variance) are statistically significant in all four models, and ρ_μ (the regional market area variance) is insignificant. The significance of the central place market area variance is one indication of support for our hierarchical structure, over a more traditional spatial lag model that does not include the nested error structure. Additionally, we conjecture that the insignificance of the regional market area variance comes from the inclusion of the regional market area variables in the model and, conditional on those variables, there is no longer any statistically detectable variation in population that stems from regional market area heterogeneity. Digging a little deeper, we can use the estimates of σ_e^2, ρ_μ, ρ_α to calculate intraclass correlation coefficients that define the proportion of the error variance that comes from each hierarchical level.[9] We find that variation at the central place market area accounts for about 5 percent (in the MA models) or 13 percent (in the UA models) of the error variance, with the remaining error variance coming from the urban area (first-tier) level.

[8] In addition to these models, we explore traditional (non-hierarchical) spatial lag regressions that correspond to these four models. We find statistical support in favor of the spatial hierarchical regressions; results from these alternative spatial models are available from the authors.

[9] Mathematically, the intraclass correlation coefficient for the central place market area is $\sigma_\alpha^2/(\sigma_e^2+\sigma_\mu^2+\sigma_\alpha^2)$, which is easily adjusted to compute the associated coefficient for the other market areas. See Snijders and Bosker (2012) for more details.

These regressions also reveal some interesting differences in the patterns of population across the MA and UA datasets. It is useful to recall that the MA dataset is the smaller sample of relatively larger urban areas (at least 50,000 residents), while the UA dataset is the larger sample of relatively smaller urban areas in addition to the larger areas in the MA sample. We find four facts to be noteworthy. First, the goods centrality index is insignificant when using the MA dataset (in Table 6), but is significant in both models that use the UA dataset (Table 7). These estimates indicate that goods centrality—which is a key component of the urban hierarchical structure—is important for smaller urban areas but is no longer important for relatively larger urban areas. A large variety of goods and services are generally available in relatively larger urban areas, unlike in many smaller urban areas that appear in the UA dataset, so, for these larger areas, the notion of urban hierarchy is less important.

Second, we do not find broad statistical support for the natural amenities (i.e., distance to large water bodies, land ruggedness, and temperature) as being important drivers of population in an urban area. It is interesting that we see more support for the natural amenities in the UA models (distance to water bodies and elevation difference in Model 1 and smooth plains and temperature discomfort in Model 2), which suggests that these amenities may be more important for populations locating in smaller urban areas. These estimates are consistent with historical studies of urban development (Dobis et al., 2017), that show that the natural amenities were historically important for the development of the urban system but over time became less important given structural technological change. So, in larger urban areas in which there is a greater variety of economic and technological amenities, the natural amenities have become less important for populations, whereas these factors may remain important in areas that lack richness in other non-natural amenities.

Third, we see that regional aggregate income and the share of manufacturing have positive and negative effects on population in the MA dataset, which is somewhat counterintuitive; yet, in the UA models, regional aggregate income becomes insignificant and manufacturing becomes negative and significant (in Model 2 of Table 7). This latter estimate is consistent with recent structural changes in the United States economy: as the manufacturing sector has continued to decline, populations have gradually relocated away from areas that were traditionally known for manufacturing.

Fourth, the spatial lag parameters (λ_{2000} and λ_{2010} in the MA models and λ in the UA models) are significantly negative in the MA models, which indicates spatial competition (or "negative spatial contagion") among urban areas, but significantly positive in the UA dataset, which indicates spatial complementarity ("positive spatial contagion"). These estimates also come from the fact that the urban areas in the MA sample are relatively larger and, hence, more likely to compete for populations (or resources more generally), whereas smaller urban areas are more likely to depend on each other for a sharing of resources.

4.2 Spatial Contagion and Urban Hierarchy

The most interesting insights we can draw from our models correspond to the effects of spatial contagion and urban hierarchy, which come from the marginal effects implied by the estimated regression parameters. It is well known that spatial lag models generate estimates of direct, indirect, and total marginal effects. We compute these effects for each of our models and report the average estimates in Tables 8 and 9. Spatial contagion manifests in significant indirect effects, and urban hierarchy manifests in statistical significance of the centrality index and the regional market area variables (in addition to the structured error variances, discussed earlier).

There are several features of our estimates that are noteworthy and correspond to the ideas of spatial contagion and urban hierarchy. The first is that the average indirect effect of the initial population level is negative and statistically significant in both versions of the model using the MA dataset (Table 8), but the same effect is relatively quite large, positive, and significant for both models when using the UA dataset (Table 9). These estimates indicate that a higher initial population in a relatively large metropolitan area corresponds to a smaller subsequent population in the proximate metropolitan areas; in other words, larger urban areas compete for populations. Yet, the same is not true of smaller urban areas. The corresponding positive and significant estimate indicates that smaller urban areas benefit from the influx of populations to proximate areas. These effects, differential by urban area size, are intuitive: smaller urban areas mutually benefit as their neighbors grow because of the agglomerative effects of that proximate growth. On the other hand, relatively larger urban areas have already grown to the extent that there is agglomeration within the urban area, so the area must compete with neighboring areas for further population resources.

The second noteworthy finding that corresponds to both spatial contagion and urban hierarchy is that the goods centrality index is not significant in any way (directly or indirectly) in either of the MA models but is positive and significant in both ways (directly and indirectly) in both of the UA models. These results are consistent with the principle of our first argument: that the relatively larger urban areas are more independent from others in the hierarchy, afforded by their size, so being central is not important. Yet, smaller urban areas depend crucially on their place in the urban system, such that being more central (or having a directly proximate neighbor that is more central) increases population levels within. Further, the significance of the indirect effects of the centrality index in the UA specifications is evidence of interaction between the forces of spatial contagion and the structure of the urban hierarchy.

Finally, our analysis indicates a significant difference in terms of both spatial contagion and urban hierarchy across large and small urban areas. In recognizing this fact, it is clear that focusing entirely on large urban areas leads to a different picture of the urban system compared to an analysis focusing on both large and small (or just small) urban areas.

4.3 Trends in the Urban System

Looking at the coefficient estimates and the direct and indirect spatial effects provides a lens for understanding the interaction between spatial contagion and urban hierarchical structure. To complete our discussion, we describe four trends in the urban system that are visible from our separate focus on spatial contagion and urban hierarchy.

1. Though the spatial distribution of the urban population changes gradually over time, our analysis of spatial contagion reveals competition for populations among proximate (large) metropolitan areas and complementarity among (small) urban areas.

The strong direct effect of the lagged population level reflects gradual spatial change in the population over time. Our lens of spatial contagion clearly reveals competition for population among metropolitan areas, which reflects agglomeration benefits within a large urban area. Yet, small urban areas mutually benefit as proximate populations increase.

2. The structure of the urban hierarchy significantly affects the level of the population in (small) urban areas but does not significantly affect the population level in (large) metropolitan areas. Furthermore, spatial contagion ensures that the significance of the urban hierarchy in (small) urban areas has both direct and indirect effects on proximate areas.

It is useful to recall that the centrality index reflects the urban hierarchy in the model such that a larger value of the centrality index indicates a wider variety of goods and services as is characteristic of higher-tier locations. A priori, we expect this variable to positively affect the population of a settlement, though the centrality index is not significant for metropolitan areas. This suggests that there is no direct relationship between the size of a metropolitan area and the balance and variety of products available. Our descriptive statistics for the Seattle central place market region provide support for this suggestion: the average size of a metropolitan area in the Seattle market area was the fourth largest in the country in 2000 (472,162 inhabitants), while the average centrality index in the market area was the largest in the country (14.268). From the regression results and the Seattle example, it is clear that the goods centrality index does not correlate well with population levels in large metropolitan areas. At the same time, the centrality index has the expected positive effect on the population size of urban areas, showing that a one unit increase in the centrality index leads to nearly a 0.1 percent increase in the population within the urban area and a 0.002 percent increase in the population in proximate urban areas.

3. Natural amenities have only a limited effect on the level of the population in (small) urban and (large) metropolitan areas.

The relationship between the natural amenities and the population level of an urban/metropolitan area is weak, largely because the foundations of this relationship is historical and our regressions control for initial population levels. Historical studies of the United States urban system (e.g., Bosker and Buringh, 2017; Dobis et al., 2017) indicate the importance of natural amenities prior to the modern economic/technological era. Therefore, in our contemporary analysis, the natural amenities play a limited role.

It is worth pointing out that temperature discomfort is a natural amenity used by other authors to capture a preference for a temperate climate (e.g., Glaeser et al., 2001; González-Val, 2015), and counterintuitively, we find that urban areas are positively affected by more temperature discomfort. Yet, these effects are very small in magnitude.

4. The proportion of the population employed in the manufacturing sector is inversely associated with the level of the population in (small) urban areas but has a positive direct effect and negative indirect effect in (large) metropolitan areas.

Our finding that the share of employment in manufacturing in the regional market area is negatively correlated with population levels in (small) urban areas is consistent with evidence of population decline in areas affected by recent structural economic transformations away from manufacturing. Our lens of spatial contagion affords us unique insight, revealing a positive direct effect, but negative indirect effect, of manufacturing for (large) metropolitan areas. Our results indicate that a one percent increase in the share of employment in manufacturing leads to a 1.4 percent increase in the level of population within the metropolitan area, and this same increase leads to a 0.1 percent decrease in the level of the population in proximate areas. We believe this trend reflects the benefits of agglomeration within the metropolitan area.

5 Conclusion

In this paper we disentangle the unique effects of contagion and hierarchy on the population level of cities in the urban system using a spatial lag hierarchical linear model. We develop this model empirically using a unique dataset of urban areas in the United States that includes small urban locations that are often excluded from analyses of the urban system. Our results clearly indicate heterogeneity in terms of the effects of both spatial contagion and urban hierarchical structure on population levels. We draw numerous conclusions describing the urban system in the United States. A concise summary of our findings is that (large) metropolitan areas are characterized by urban agglomeration and spatial competition for population, while (small) urban areas are characterized by both spatial complementarity (as a manifestation of spatial contagion) and position in the urban hierarchy.

References

Abdel-Rahman, H. M. (1988). Product differentiation, monopolistic competition and city size. *Regional Science and Urban Economics 18*(1), 69–86.

Abdel-Rahman, H. M. (1990). Agglomeration economies, types, and sizes of cities. *Journal of Urban Economics 27*(1), 25–45.

Abdel-Rahman, H. M. and M. Fujita (1993). Specialization and diversification in a system of cities. *Journal of Urban Economics 33*(2), 189–222.

Alonso, W. (1964). *Location and Land Use: Toward a General Theory of Land Rent.* Harvard University Press.

Anas, A. and K. Xiong (2003). Intercity trade and the industrial diversification of cities. *Journal of Urban Economics 54*(2), 258–276.

Antweiler, W. (2001). Nested random effects estimation in unbalanced panel data. *Journal of Econometrics 101*(2), 295–313.

Baltagi, B. H., G. Bresson, and J.-M. Etienne (2015). Hedonic housing prices in Paris: An unbalanced spatial lag pseudo-panel model with nested random effects. *Journal of Applied Econometrics 30*(3), 509–528.

Baltagi, B. H., B. Fingleton, and A. Pirotte (2014). Spatial lag models with nested random effects: An instrumental variable procedure with an application to English house prices. *Journal of Urban Economics 80*, 76–86.

Black, D. and V. Henderson (2003). Urban evolution in the USA. *Journal of Economic Geography 3*(4), 343–372.

Borchert, J. R. (1967). American metropolitan evolution. *Geographical Review 57*(3), 301–332.

Bosker, M. and E. Buringh (2017, March). City seeds: Geography and the origins of the European city system. *Journal of Urban Economics 98*, 139–157.

Burger, M. J. and E. J. Meijers (2016). Agglomerations and the rise of urban network externalities. *Papers in Regional Science 95*(1), 5–15.

Camagni, R., R. Capello, and A. Caragliu (2015). The rise of second-rank cities: What role for agglomeration economies? *European Planning Studies 23*(6), 1069–1089.

Capello, R. (2000). The city network paradigm: Measuring urban network externalities. *Urban Studies 37*(11), 1925–1945.

Castells-Quintana, D. (2017). Malthus living in a slum: Urban concentration, infrastructure and economic growth. *Journal of Urban Economics 98*, 158–173.

Christaller, W. (1933). *Die Zentralen Orte in Süddeutschland* (1966 Translation ed.). Englewood Cliffs: Prentice Hall.

Coşar, A. K. and P. D. Fajgelbaum (2016). Internal geography, international trade, and regional specialization. *American Economic Journal: Microeconomics 8*(1), 24–56.

Cress, J., R. Sayre, P. Comer, and H. Warner (2009). Terrestrial ecosystems—land surface forms of the conterminous United States [Dataset].

de Vries, J. (1990). Problems in the measurement, description, and analysis of historical urbanization. In A. van der Woude, J. de Vries, and A. Hayami (Eds.), *Urbanization in History: A Process of Dynamic Interactions*, pp. 371. Oxford: Clarendon Press.

Desmet, K. and J. Rappaport (2017, March). The settlement of the United States, 1800–2000: The long transition towards Gibrat's Law. *Journal of Urban Economics 98*, 50–68.

Desmet, K. and E. Rossi-Hansberg (2014). Spatial development. *The American Economic Review 104*(4), 1211–1243.

Dobis, E. A., M. S. Delgado, R. J. Florax, and P. Mulder (2017). Urban settlement in the United States: Birth, location, and growth of incorporated places. Working paper.

Dobkins, L. and Y. M. Ioannides (2001). Spatial interactions among US cities: 1900–1990. *Regional Science and Urban Economics 31*(6), 701–731.

Duranton, G. and D. Puga (2014). The growth of cities. In *Handbook of Economic Growth*, Volume 2, pp. 781–853. Elsevier.

Duranton, G. and M. A. Turner (2012). Urban growth and transportation. *Review of Economic Studies 79*(4), 1407–1440.

Frick, S. A. and A. Rodríguez-Pose (2017). Big or small cities? on city size and economic growth. *Growth and Change*.

Fujita, M. (1988). A monopolistic competition model of spatial sgglomeration: Differentiated product approach. *Regional Science and Urban Economics 18*(1), 87–124.

Fujita, M. (2012). Thünen and the new economic geography. *Regional Science and Urban Economics 42*(6), 907–912.

Fujita, M., P. Krugman, and T. Mori (1999). On the evolution of hierarchical urban systems. *European Economic Review 43*(2), 209–251.

Fujita, M., P. R. Krugman, and A. J. Venables (1999). *The Spatial Economy: Cities, Regions and International Trade*, Volume 213. Wiley Online Library.

Glaeser, E. L. and J. E. Kohlhase (2004). Cities, regions and the decline of transport costs. *Papers in Regional Science 83*(1), 197–228.

Glaeser, E. L., J. Kolko, and A. Saiz (2001). Consumer city. *Journal of Economic Geography 1*(1), 27–50.

GLOBE Task Team (1999). The Global Land One-kilometer Base Elevation (GLOBE) Digital Elevation Model, version 1.0 [Database].

González-Val, R. (2010). The evolution of U.S. city size distribution from a long-term perspective (1900–2000). *Journal of Regional Science 50*(5), 952–972.

González-Val, R. (2015). Cross-sectional growth in US cities from 1990 to 2000. *Journal of Geographical Systems 17*(1), 83–106.

González-Val, R. and L. Lanaspa (2016). Patterns in US urban growth, 1790–2000. *Regional Studies 50*(2), 289–309.

Helsley, R. W. and W. C. Strange (2014). Coagglomeration, clusters, and the scale and composition of cities. *Journal of Political Economy 122*(5), 1064–1093.

Henderson, J. V. (1974). The sizes and types of cities. *The American Economic Review*, 640–656.

Horizon Systems Corporation (2012). National Hydrology Dataset (NHD) Plus: Version 2.

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*(1), 99–121.

Kim, S. (2000). Urban development in the United States, 1690–1990. *Southern Economic Journal 66*(4), 855–880.

Knox, P. L. and L. M. McCarthy (2005). *Urbanization: An Introduction to Urban Geography.* Pearson Prentice Hall.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy 99*(3), 483–499.

Le Gallo, J. and C. Chasco (2008). Spatial analysis of urban growth in Spain, 1900–2001. *Empirical Economics 34*(1), 59–80.

Lorenzen, M. and K. V. Andersen (2009). Centrality and creativity: Does Richard Florida's creative class offer new insights into urban hierarchy? *Economic Geography 85*(4), 363–390.

Lösch, A. (1943). *Die räumliche Ordnung der Wirtshcaft* (2nd Revised Edition ed.), Volume 1967. New York: John Wiley & Sons, Inc.

Maignan, C., G. Ottaviano, D. Pinelli, and F. Rullani (2003). Bio-ecological diversity vs. socio-economic diversity: A comparison of existing measures. Working Paper 13, Fondazione Eni Enrico Mattei (FEEM).

Maliszewski, P. J. and B. Ó hUallacháin (2012). Hierarchy and concentration in the American urban system of technological advance. *Papers in Regional Science 91*(4), 743–758.

McCann, P. and Z. J. Acs (2011). Globalization: Countries, cities and multinationals. *Regional Studies 45*(1), 17–32.

Meijers, E. J., M. J. Burger, and M. M. Hoogerbrugge (2016). Borrowing size in networks of cities: City size, network connectivity and metropolitan functions in Europe. *Papers in Regional Science 95*(1), 181–198.

Minnesota Population Center (2011). National Historical Geographic Information System: Version 2.0 [Database].

Neal, Z. P. (2011). From central places to network bases: A transition in the US urban hierarchy, 1900–2000. *City & Community 10*(1), 49–75.

Nzaku, K. and J. O. Bukenya (2005). Examining the relationship between quality of life amenities and economic development in the southeast USA. *Review of Urban & Regional Development Studies 17*(2), 89–103.

Olfert, M. R., M. Jelinski, D. Zikos, and J. Campbell (2012). Human capital drift up the urban hierarchy: Veterinarians in western Canada. *Annals of Regional Science 49*(2), 551–570.

Overman, H. G. and Y. M. Ioannides (2001). Cross-sectional evolution of the U.S. city size distribution. *Journal of Urban Economics 49*(3), 543–566.

Partridge, M. D., D. S. Rickman, K. Ali, and M. R. Olfert (2008). Lost in space: Population growth in the American hinterlands and small cities. *Journal of Economic Geography 8*(6), 727–757.

Partridge, M. D., D. S. Rickman, K. Ali, and M. R. Olfert (2009a). Agglomeration spillovers and wage and housing cost gradients across the urban hierarchy. *Journal of International Economics 78*(1), 126–140.

Partridge, M. D., D. S. Rickman, K. Ali, and M. R. Olfert (2009b). Do New Economic Geography agglomeration shadows underlie current population dynamics across the urban hierarchy? *Papers in Regional Science 88*(2), 445–466.

Polèse, M. and R. Shearmur (2004). Culture, language, and the location of high-order service functions: The case of Montreal and Toronto. *Economic Geography 80*(4), 329–350.

PRISM Climate Group, Oregon State University (2014). Temperature 30-Year Normals [Database].

Redding, S. J. and E. A. Rossi-Hansberg (2017). Quantitative spatial economics. *Annual Review of Economics 9*(1).

Rossi-Hansberg, E. (2005). A spatial theory of trade. *American Economic Review 95*(5), 1464–1491.

Snijders, T. A. and R. J. Bosker (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling* (2nd ed.). Sage Publications.

Stirling, A. (1998). On the economics and analysis of diversity. Electronic Working Papers Series 28, Science Policy Research Unit (SPRU).

Tabuchi, T. and J. F. Thisse (2011). A new economic geography model of central places. *Journal of Urban Economics 69*(2), 240–252.

US Census Bureau (2011). Urban area criteria for the 2010 census. *Federal Register Notice 76*(164), 53030–53043.

US Census Bureau (2014). Places 1960–2010. Obtained from Todd Gardner, Survey Statistician, Center for Economic Studies, US Census Bureau.

US Census Bureau, Business & Industry (2012). North American Product Classification System, overview and progress report.

US Census Bureau, Census History Staff (2013). Urban and rural areas - history. A description of how the Census Bureau defines the geographic terms urban and rural.

US Census Bureau, Geography (1994a). *Geographic Areas Reference Manual*, Chapter 9. Places.

US Census Bureau, Geography (1994b). *Geographic Areas Reference Manual*, Chapter 12. The Urban and Rural Classifications.

US Census Bureau, Geography (1994c). *Geographic Areas Reference Manual*, Chapter 13. Metropolitan Areas.

US Census Bureau, Geography (2012). 2010 urban area FAQs.

Van Meeteren, M., Z. Neal, and B. Derudder (2016). Disentangling agglomeration and network externalities: A conceptual typology. *Papers in Regional Science 95*(1), 61–80.

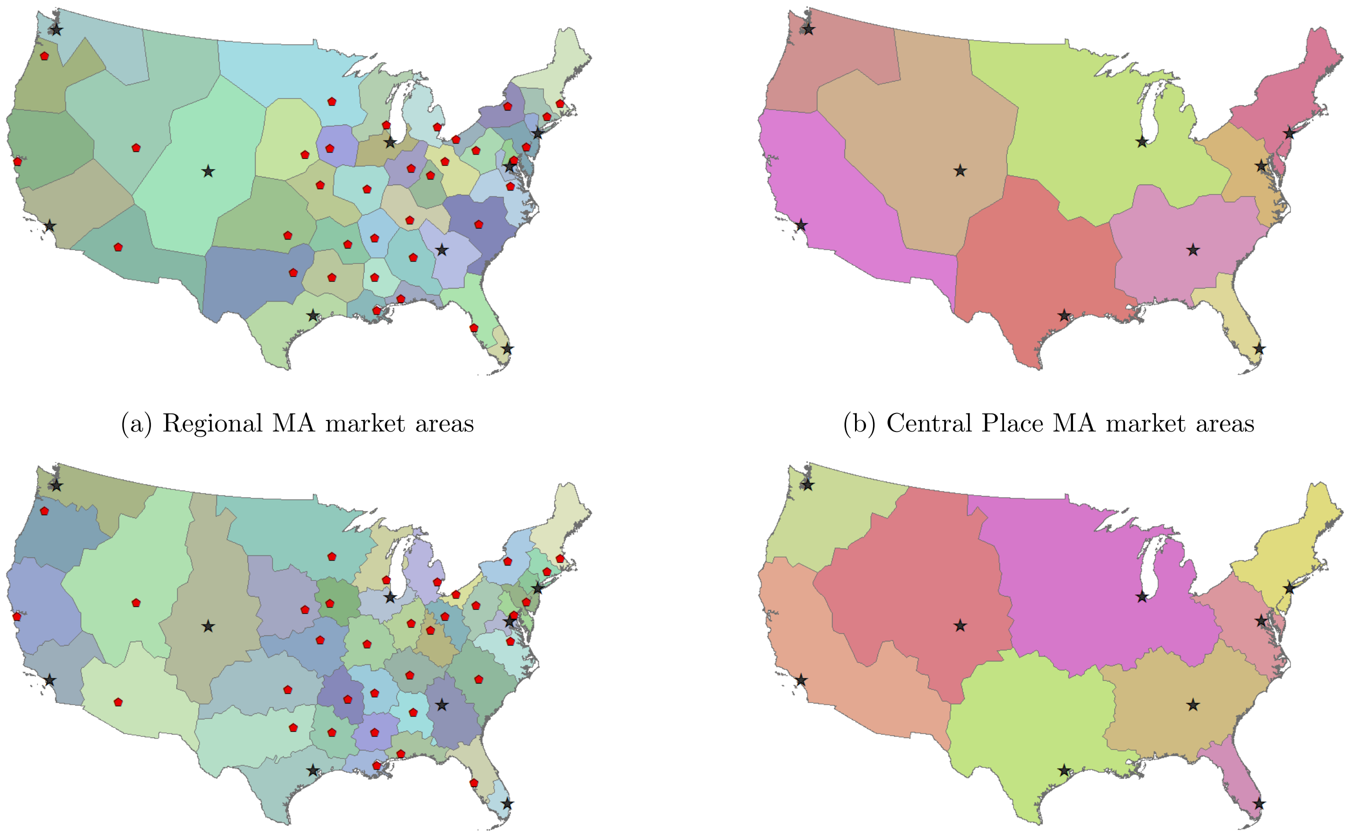
Walls & Associates (2012). National Establishment Time-Series (NETS) Database.

Zheng, S., Y. Fu, and H. Liu (2009). Demand for urban quality of living in China: Evolution in compensating land-rent and wage-rate differentials. *Journal of Real Estate Finance and Economics 38*(3), 194–213.

(a) MA Dataset

(b) UA Dataset

Figure 1: Metropolitan and Urban Areas in the United States for the MA and UA datasets.



(a) Regional MA market areas

(b) Central Place MA market areas

(c) Regional UA market areas

(d) Central Place UA market areas

Figure 2: MA and UA market areas for middle- and high-tier nodes. In the regional node maps, panels (a) and (c), the regional nodes (e.g., Indianapolis, San Francisco) are represented by red pentagons and the central place nodes (e.g., New York City, Denver) are represented by black stars.

6 Tables

Table 1: Overman and Ioannides (2001) and Dobkins and Ioannides (2001) Tier Classification

National Nodal (Tier 1)	Regional Nodal (Tier 2)	Subregional Nodal (Tier 3)
Atlanta	Baltimore	Birmingham
Chicago	Boston	Charlotte
Denver	Cincinnati	Des Moines
Houston	Cleveland	Detroit
Los Angeles	Columbus	Hartford
Miami	Dallas	Jackson, MS
New York	Indianapolis	Little Rock
San Francisco	Kansas City	Memphis
Seattle	Minneapolis	Milwaukee
Washington, DC	New Orleans	Mobile
	Philadelphia	Nashville
	Phoenix	Oklahoma City
	Portland	Omaha
	St. Louis	Pittsburgh
		Richmond
		Salt Lake City
		Shreveport
		Syracuse
		Tampa

Table 2: Descriptive Statistics for the Metropolitan Area Data, 1990–2010

Variable	Mean	Std. Dev.	Min.	Max.
	Metropolitan Areas			
Total Pop 1990	428, 582.800	1, 235, 902.000	50,066	16,044,012
Total Pop 2000	515, 174.700	1, 397, 374.000	50,902	17,832,182
Total Pop 2010	574, 970.800	1, 480, 237.000	44,022	18,388,132
ln(Pop 1990)	12.008	1.112	10.821	16.591
ln(Pop 2000)	12.225	1.115	10.838	16.697
ln(Pop 2010)	12.347	1.128	10.692	16.727
Avg Temp (° C, 1990)	13.867	4.583	4.408	24.344
Avg Temp (° C, 2000)	13.852	4.583	4.385	24.307
Temp Discomfort (1990)	19.817	4.883	7.179	33.549
Temp Discomfort (2000)	19.672	4.914	7.020	33.456
Dist to GL/Ocean (km, 1990)	232.318	277.499	0.000	1,232.600
Dist to GL/Ocean (km, 2000)	232.576	277.507	0.000	1,237.063
Ruggedness (category, 1990)	1.886	1.247	1.000	8.000
Ruggedness (category, 2000)	1.913	1.243	1.000	8.000
Elev Diff (m, 1990)	170.698	195.273	2.000	1,388.000
Elev Diff (m, 2000)	176.185	186.649	4.000	1,340.000
Dist to Middle/High Tier (km)	165.546	142.441	0.000	914.290
Centrality Index (1990)	11.980	1.726	6.592	15.256
Centrality Index (2000)	13.215	1.365	8.038	15.676
	Regional Market Areas			
Real Agg Income 1989 (B)	100.963	110.559	7.692	527.125
Real Agg Income 1999 (B)	132.556	131.223	9.381	617.600
Dist to High Tier (km)	354.959	251.030	0.000	780.711
Mfg Employment Share 1990	0.151	0.044	0.034	0.243
Svc Employment Share 1990	0.555	0.028	0.492	0.619
Mfg Employment Share 2000	0.114	0.038	0.034	0.208
Svc Employment Share 2000	0.608	0.027	0.563	0.692
Urban Area 1990 (km^2)	3, 693.497	2, 255.076	703.259	9,116.381
Rural Area 1990 (km^2)	177, 189.600	186, 161.700	9,317.770	889,139.500
Urban Area 2000 (km^2)	4, 297.963	2, 927.512	587.548	11,149.070
Rural Area 2000 (km^2)	176, 585.100	186, 339.500	7,921.183	888,812.900
Total Area (km^2)	180, 883.100	186, 300.400	17,482.970	892,361.600
Prop of Rural Area 1990	0.955	0.073	0.533	0.998
Prop of Rural Area 2000	0.946	0.087	0.453	0.997
	Central Place Market Areas			
Urban Area 1990 (km^2)	17, 646.710	9, 695.224	4,841.911	33,930.010
Rural Area 1990 (km^2)	846, 572.500	590, 115.400	141,397.100	1,694,827.000
Urban Area 2000 (km^2)	20, 534.710	11, 438.140	5,627.390	37,078.080
Rural Area 2000 (km^2)	843, 684.500	591, 059.900	138,990.800	1,691,679.000
Total Area (km^2)	864, 219.200	592, 324.400	152,591.300	1,728,757.000
Prop of Rural Area 1990	0.967	0.027	0.927	0.997
Prop of Rural Area 2000	0.960	0.034	0.905	0.997

The MA data includes Census-defined urban areas with a population greater than 50,000 residents in 1990. There are 367 metropolitan areas. Each of the 43 regional market areas contain all metropolitan areas closest to the regional node by network distance, and the 9 central place market areas contain all metropolitan areas closest to central place nodes by network distance.

Table 3: Descriptive Statistics for the Urban Area Data, 2000–2010

Variable	Mean	Std. Dev.	Min.	Max.
		Urban Areas		
Total Pop 2000	69,358.930	493,711.400	2,501.000	17,832,182
Total Pop 2010	77,589.230	525,599.100	2,503.000	18,388,132
ln(Pop 2000)	9.313	1.366	7.824	16.697
ln(Pop 2010)	9.396	1.398	7.825	16.727
Avg Temp (° C)	13.004	4.457	1.589	25.442
Temp Discomfort	23.454	4.408	8.834	36.558
Dist to GL/Ocean (km)	308.871	281.172	0.000	1,275.840
Ruggedness (category)	2.085	1.411	1.000	8.000
Elev Diff (m)	86.206	113.442	0.000	1,340.000
Dist to Middle/High Tier (km)	205.930	146.456	0.000	1,133.459
Centrality Index (2000)	7.456	2.998	0.000	15.684
		Regional Market Areas		
Real Agg Income 1999 (B)	148.870	132.853	16.640	624.207
Dist to High Tier	354.961	251.219	0.000	780.711
Mfg Employment Share 2000	0.121	0.036	0.039	0.208
Svc Employment Share 2000	0.597	0.024	0.557	0.658
Urban Area 2000 (km^2)	5,545.067	3,223.160	1,341.056	13,986.150
Rural Area 2000 (km^2)	175,338.000	187,313.400	12,641.060	870,714.500
Total Area (km^2)	180,883.100	187,224.100	22,690.070	873,979.800
Prop of Rural Area 2000	0.936	0.077	0.557	0.996
		Central Place Market Areas		
Urban Area 2000 (km^2)	26,493.100	14,505.270	8,136.151	50,419.700
Rural Area 2000 (km^2)	837,726.100	592,376.300	114,723.700	1,746,183.000
Total Area (km^2)	864,219.200	595,934.900	130,256.900	1,796,603.000
Prop of Rural Area 2000	0.950	0.042	0.881	0.995

The UA data includes Census-defined urban areas that exist in both 2000 and 2010. There are 3,174 urban areas. Each of the 43 regional market areas contain all urban areas closest to the regional node by network distance, and the 9 central place market areas contain all urban areas closest to central place nodes by network distance.

Table 4: Descriptive Statistics for Metropolitan Areas by Tier, 1990–2010

Variable	Mean	Std. Dev.	Min.	Max.
High Tier				
Total Pop 1990	5,658,645	5,007,423	1,517,977	16,044,012
Total Pop 2000	6,642,846	5,315,725	2,010,212	17,832,182
Total Pop 2010	7,236,398	5,266,982	2,374,203	18,388,132
Avg Temp (° C)	15.019	5.176	9.889	24.307
Temp Discomfort	17.362	4.981	9.332	24.977
Dist to GL/Ocean (km)	186.109	394.066	0.276	1,193.023
Ruggedness (category)	1.889	1.054	1.000	3.000
Elev Diff (m)	419.111	435.460	8.000	1,340.000
Centrality Index (2000)	15.317	0.282	14.942	15.611
Middle Tier				
Total Pop 1990	1,396,655	1,137,382	256,489	4,671,827
Total Pop 2000	1,626,009	1,301,799	275,213	5,190,255
Total Pop 2010	1,775,598	1,393,913	298,317	5,441,567
Avg Temp (° C)	13.853	4.071	7.586	22.603
Temp Discomfort	20.418	4.186	10.374	29.477
Dist to GL/Ocean (km)	254.496	253.067	2.069	928.005
Ruggedness (category)	2.000	1.101	1.000	6.000
Elev Diff (m)	209.941	162.260	10.000	718.000
Dist to High Tier (km)	448.918	191.686	67.809	780.710
Centrality Index (2000)	14.792	0.459	13.784	15.676
Low Tier				
Total Pop 1990	181,715.500	235,888.100	50,066	2,348,417
Total Pop 2000	228,392.500	284,561.400	50,902	2,681,237
Total Pop 2010	263,939.400	334,629.200	44,022	2,956,746
Avg Temp (° C)	13.819	4.627	4.385	23.624
Temp Discomfort	19.658	4.975	7.020	33.456
Dist to GL/Ocean (km)	231.566	277.059	0.000	1,237.063
Ruggedness (category)	1.904	1.264	1.000	8.000
Elev Diff (m)	165.895	173.835	4.000	1,149.000
Dist to Middle/High Tier (km)	187.517	137.324	34.692	914.290
Centrality Index (2000)	12.991	1.286	8.038	15.525

Table 5: Descriptive Statistics for Urban Areas by Tier, 2000–2010

Variable	Mean	Std. Dev.	Min.	Max.
High Tier				
Total Pop 2000	6,549,324	5,231,462	2,010,212	17,832,182
Total Pop 2010	7,125,998	5,172,898	2,374,203	18,388,132
Avg Temp (° C)	15.029	5.176	9.889	24.307
Temp Discomfort	20.061	4.799	12.471	27.353
Dist to GL/Ocean (km)	186.136	394.058	0.276	1,193.020
Ruggedness (category)	1.889	1.054	1.000	3.000
Elev Diff (m)	420.889	434.964	8.000	1,340.000
Centrality Index (2000)	15.318	0.277	14.949	15.618
Middle Tier				
Total Pop 2000	1,597,923	1,265,224	275,213	5,190,255
Total Pop 2010	1,743,944	1,358,097	298,317	5,441,567
Avg Temp (° C)	13.848	4.070	7.586	22.602
Temp Discomfort	23.088	4.030	12.236	31.713
Dist to GL/Ocean (km)	254.465	253.362	0.000	928.005
Ruggedness (category)	2.000	1.101	1.000	6.000
Elev Diff (m)	206.882	157.723	10.000	718.000
Dist to High Tier (km)	448.920	192.000	67.809	780.710
Centrality Index (2000)	14.795	0.458	13.785	15.686
Low Tier				
Total Pop 2000	34,133.490	113,630.600	2,501	2,681,237
Total Pop 2010	39,233.530	133,138.900	2,503	2,956,746
Avg Temp (° C)	12.989	4.459	1.589	25.442
Temp Discomfort	23.468	4.409	8.834	36.558
Dist to GL/Ocean (km)	309.815	281.058	0.000	1,275.840
Ruggedness (category)	2.087	1.415	1.000	8.000
Elev Diff (m)	83.934	108.695	0.000	1,231.000
Dist to Middle/High Tier (km)	208.758	145.442	23.808	1,133.459
Centrality Index (2000)	7.353	2.887	0.000	15.575

Table 6: Metropolitan Area Spatial Hierarchical Regression Results

Dependent variable:	*Model 1*		*Model 2*	
$\ln(Pop_t)$	Coefficient	Std. Error	Coefficient	Std. Error
Constant	1.479**	(0.599)	1.081*	(0.588)
Urban Area Level:				
$\ln(Pop_{t-1})$	0.971***	(0.011)	0.977***	(0.010)
Centrality Index	0.006	(0.008)	0.007	(0.008)
Dist to GL/Ocean (100km)	−0.010**	(0.004)	−0.005	(0.004)
Elev Diff (100m)	0.012***	(0.004)		
Land Surface Forms				
Flat Plains			base	base
Smooth Plains			−0.008	(0.017)
Irregular Plains			0.007	(0.017)
Hills			0.026	(0.052)
Foothills			0.126**	(0.054)
Low Mountains			0.149	(0.098)
Avg Temp (° C)	−0.003	(0.003)		
Temp Discomfort			−0.001	(0.002)
Time Period	−0.756	(0.573)	−0.873	(0.575)
Regional Market Area Level:				
Real Agg Income ($100B)	−0.026**	(0.010)	−0.022**	(0.010)
Mfg Emp Share	1.186***	(0.415)	1.409***	(0.418)
Svc Emp Share	0.495	(0.508)	0.557	(0.516)
Rural Land Proportion	−0.149	(0.200)	−0.104	(0.203)
λ_{2000}	−0.136***	(0.037)	−0.122***	(0.037)
λ_{2010}	−0.063*	(0.034)	−0.039	(0.033)
σ_e^2	0.022***	(0.001)	0.022***	(0.001)
$\rho_\mu = \sigma_\mu^2/\sigma_e^2$	0.00001	(0.046)	0.00001	(0.034)
$\rho_\alpha = \sigma_\alpha^2/\sigma_e^2$	0.165***	(0.044)	0.148***	(0.043)
Observations	734		734	

Statistical significance of the estimates at the 10, 5, and 1 percent levels is denoted as *, **, and ***.

Table 7: Urban Area Spatial Hierarchical Regression Results

Dependent variable:	*Model 1*		*Model 2*	
ln(Pop 2010)	Coefficient	Std. Error	Coefficient	Std. Error
Constant	-1.679^{***}	(0.448)	-1.602^{***}	(0.430)
Urban Area Level:				
ln(Pop 2000)	0.993^{***}	(0.007)	0.990^{***}	(0.007)
Centrality Index	0.009^{***}	(0.003)	0.008^{***}	(0.003)
Dist to GL/Ocean (100km)	0.007^{***}	(0.003)	0.004	(0.003)
Elev Diff (100m)	-0.013^{***}	(0.004)		
Land Surface Forms				
Flat Plains			base	base
Smooth Plains			0.016^{*}	(0.009)
Irregular Plains			−0.0004	(0.009)
Escarpments			−0.067	(0.067)
Hills			0.010	(0.025)
Foothills			−0.013	(0.024)
Low Mountains			−0.028	(0.031)
Avg Temp (° C)	0.0004	(0.002)		
Temp Discomfort			0.005^{***}	(0.002)
Regional Market Area Level:				
Real Agg Income ($100B)	−0.005	(0.009)	−0.002	(0.008)
Mfg Emp Share	−0.328	(0.367)	-0.970^{***}	(0.345)
Svc Emp Share	−0.069	(0.520)	−0.678	(0.505)
Rural Land Proportion	−0.028	(0.131)	0.062	(0.116)
λ	0.198^{***}	(0.020)	0.217^{***}	(0.020)
σ_e^2	0.035^{***}	(0.001)	0.035^{***}	(0.001)
$\rho_\mu = \sigma_\mu^2/\sigma_e^2$	0.00001	(0.008)	0.00001	(0.010)
$\rho_\alpha = \sigma_\alpha^2/\sigma_e^2$	0.054^{***}	(0.019)	0.052^{***}	(0.017)
Observations	3,174		3,174	

Statistical significance of the estimates at the 10, 5, and 1 percent levels is denoted as *, **, and ***.

Table 8: Metropolitan Area Spatial Hierarchical Marginal Effects

Dependent variable:	*Model 1*			*Model 2*		
$\ln(Pop_t)$	Direct Effect	Indirect Effect	Total Effect	Direct Effect	Indirect Effect	Total Effect
Urban Area Level:						
$\ln(Pop_{t-1})$	0.971***	−0.088***	0.884***	0.977***	−0.072***	0.905***
Centrality Index	0.006	−0.001	0.005	0.007	−0.001	0.007
Dist to GL/Ocean	−0.010**	0.001*	−0.009**	−0.005	0.0004	−0.005
Elev Diff	0.012***	−0.001**	0.011***			
Land Surface Forms						
Flat Plains				base	base	base
Smooth Plains				−0.008	0.001	−0.007
Irregular Plains				0.007	−0.001	0.007
Hills				0.026	−0.002	0.024
Foothills				0.126**	−0.009**	0.117**
Low Mountains				0.149	−0.011	0.138
Avg Temp	−0.003	0.0003	−0.003			
Temp Discomfort				−0.001	0.0001	−0.001
Time Period	−0.757	0.068	−0.689	−0.873	0.064	−0.809
Regional Market Area Level:						
Real Agg Income	−0.026**	0.002**	−0.024**	−0.022**	0.002*	−0.021**
Mfg Emp Share	1.187***	−0.107**	1.080***	1.410***	−0.104**	1.306***
Svc Emp Share	0.495	−0.045	0.451	0.557	−0.041	0.516
Rural Land Proportion	−0.149	0.013	−0.136	−0.104	0.008	−0.097

Statistical significance of the estimates at the 10, 5, and 1 percent levels is denoted as *, **, and ***.

Table 9: Urban Area Spatial Hierarchical Marginal Effects

Dependent variable:	*Model 1*			*Model 2*		
ln(Pop2010)	Direct Effect	Indirect Effect	Total Effect	Direct Effect	Indirect Effect	Total Effect
Urban Area Level:						
ln(Pop 2000)	0.932***	0.245***	1.238***	0.991***	0.274***	1.265***
Centrality Index	0.009***	0.002***	0.011***	0.008***	0.002***	0.011***
Dist to GL/Ocean	0.007***	0.002***	0.009***	0.004	0.001	0.005
Elev Diff	−0.013***	−0.003***	−0.016***			
Land Surface Forms						
Flat Plains				base	base	base
Smooth Plains				0.016*	0.005*	0.021*
Irregular Plains				−0.0004	−0.0001	−0.0005
Escarpments				−0.067	−0.018	−0.085
Hills				0.010	0.003	0.013
Foothills				−0.013	−0.004	−0.017
Low Mountains				−0.028	−0.008	−0.036
Avg Temp	0.0004	0.0001	−0.001			
Temp Discomfort				0.005***	0.001**	0.007***
Regional Market Area Level:						
Real Agg Income	−0.005	−0.001	−0.006	−0.002	−0.001	−0.003
Mfg Emp Share	−0.328	−0.081	−0.409	−0.970***	−0.269***	−1.239***
Svc Emp Share	−0.069	−0.017	−0.086	−0.678	−0.188	−0.866
Rural Land Proportion	−0.028	−0.007	−0.035	0.062	0.017	0.079

Statistical significance of the estimates at the 10, 5, and 1 percent levels is denoted as *, **, and ***.