

Dobis, Elizabeth; Delgado, Michael; Florax, Raymond; Mulder, Peter

**Conference Paper**

## Population Growth in American Cities between 1990 and 2010: True Contagion and Urban Hierarchy

55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal

**Provided in Cooperation with:**

European Regional Science Association (ERSA)

*Suggested Citation:* Dobis, Elizabeth; Delgado, Michael; Florax, Raymond; Mulder, Peter (2015) : Population Growth in American Cities between 1990 and 2010: True Contagion and Urban Hierarchy, 55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/124774>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Population Growth in American Cities between 1990 and 2010: True Contagion and Urban Hierarchy

## *Introduction*

Cities and towns are loci of population and production. In 2010, 80.7 percent of the United States population resided in urban areas, and 88.2 percent of those individuals lived in an urban area with at least 50,000 inhabitants, according to the U.S. Census Bureau. These individuals rely on the goods and services available in urban areas, while much of the rural population has to travel to urban areas to gain access to goods and services. The U.S. Bureau of Economic Analysis reported that in 2011, 90.1 percent of gross domestic product (GDP) was produced in metropolitan statistical areas (MSAs), emphasizing that urban areas and their immediate hinterlands are also centers of production.

The high concentration of people and production in urban areas of the United States makes the structure and growth of the urban system a policy focus for national and local governments due to the economic, environmental, and social implications of urbanization. Governments use policy to alter the incentives of firms and individuals, affecting their choices and impacting the structure of the urban system through the number, location, and sizes of cities, as well as urban form and density.

Academic interest in urban population change can be split into two literature branches. The first branch focuses on the growth and structure of individual cities. The second branch focuses on the urban system. Some of the literature in this branch implicitly assumes interconnectedness among cities (e.g., Black and Henderson, 2003; Le Gallo and Chasco, 2008), while other authors explicitly include the spatial proximity between cities (e.g., Partridge et al., 2008; Bosker and Buringh, 2011) or network flows among cities (e.g., Neal, 2011). However, in many studies the difference between contagious and hierarchical interrelations across cities comprised in the urban system are obfuscated. In this paper we clearly distinguish and quantify the effects of both. In other words, we focus on how the structure of the urban system influences population growth. We do this by using central place theory as a theoretical basis for addressing the research question: what natural and man-made locational characteristics influence population growth?

## *Contributions*

We envisage three major contributions to the existing literature. First, we utilize a unique dataset of urban areas with decennial observations from 1990 to 2010. Most of the literature studying urban systems in the United States uses metropolitan statistical areas, which include rural hinterlands that are not part of the urban area, preventing a clear separation between cities and hinterlands and obscuring results. We built a new dataset at the more appropriate and precise geographic level of urban areas, which capture the agglomerated economic activity and built extent of urban locations.

Second, departing from literature that includes either urban hierarchy or continuous urban proximity, our analysis includes both the hierarchical relationship among cities of differing sizes and the continuous nature of proximity to other cities. The novel use of a spatially-lagged hierarchical linear model allows us to include both these critical aspects of the urban system in our analysis. This econometric model captures the influence of the structure of the urban system on population change by allowing city-level explanatory variables to affect population growth differently, given the unique characteristics of each location's regional and central place market areas. Concurrently we account for contiguous effects through the inclusion of neighbors.

Third, we include man-made amenities and characteristics of cities, which have been omitted from previous studies in an effort to avoid endogeneity in the analysis. The inclusion of these

characteristics allows us to include both producer-based and consumer-based characteristics in the analysis, to capture the idea that cities are markets.

### *Model and Data*

The body of empirical literature that explicitly studies dynamics in the urban system utilizes unilevel models. In unilevel models, all variables, whether they capture aspects of cities or market areas, directly and independently influence population change. When a weights matrix is included in the model, the result is a single interconnected system of cities in which all cities influence all other cities, with the strength of that influence dependent on the distance separating each location. In reality, the effect of city-level explanatory variables on population change may vary depending on the unique characteristics of each market area. These unique characteristics arise because, unlike the assumptions of central place theory, the world is not a flat, featureless plain.

To include this possibility in the analysis, we use a multilevel model, also known as a hierarchical level (HLM) model. This will allow the city-level coefficients to vary based on characteristics of the market areas. Due to the nested structure of the model, characteristics of the top level of the hierarchy will influence the lowest level of the model through their effect on intermediate levels.

The number of levels in the model is the same as the number of tiers in the urban hierarchy because the hierarchical classification of cities is directly related to the goods hierarchy. At each level of the goods hierarchy, goods and services that are more specialized and have a higher price elasticity are available, in addition to the goods available at lower tiers. The larger market area required to sustain these more specialized goods subsumes the smaller market areas for the lower tiers, creating a nested structure. Therefore, each higher tier has a larger market area, resulting in an equal number of model levels as city tiers.

Rather than arbitrarily determining the tiers based on population size—a frequent practice in the literature—we use a modified version of the nodal functional hierarchy used by Overman and Ioannides (2001) and Dobkins and Ioannides (2001) for our hierarchy. This hierarchy is exogenous to the model and fixed throughout the analysis.

Our dataset of urban areas from 1990 to 2010 was constructed using geographic information systems and statistical programs to compile data available from the U.S. Census Bureau, U.S. Geological Survey, PRISM Climate Group, National Establishment Time Series (NETS) Database, and boundary files of physical and made-made features of the United States. The geography’s lower threshold of 2,500 inhabitants ensures that cities as well as all but the smallest rural communities are included in the analysis.

Our dependent variable is the change in the population of an urban area, calculated as the difference of natural logarithms. Our model represents changes in the population of urban areas via hierarchy, heterogeneity, and proximity. Hierarchy is incorporated through our econometric setup, while our explanatory variables address heterogeneity and proximity. This results in the following base model:

Urban Area Level:

$$y_{ijk} = \sum_{p=0}^P \pi_{pjk} A_{pijk} + e_{ijk} \tag{1}$$

Regional Market Area Level:

$$\pi_{pjk} = \sum_{q=0}^{Q_p} \beta_{pqk} X_{qjk} + r_{pjk}, \quad \forall p = 0, \dots, P \quad (2)$$

Central Place Market Area Level:

$$\beta_{pqk} = \sum_{s=0}^{S_{pq}} \gamma_{pqs} Z_{sk} + u_{pqk}, \quad \forall p = 0, \dots, P \text{ and } q = 0, \dots, Q_p \quad (3)$$

where  $i$  is an urban area,  $j$  is a region, and  $k$  is a central place.  $Y$  is the dependent variable, population change in city  $i$ , and the independent variables,  $A$ ,  $X$ , and  $Z$  contain variables that address heterogeneity and proximity. The errors are  $e$ ,  $r$ , and  $u$  for each level, respectively.

The urban area-level equation includes time-lagged population and goods centrality index variables. The goods centrality index includes market heterogeneity through the variety and balance of products in each urban area. The heterogeneity of natural amenities among urban areas is captured by variables addressing the temperature, ruggedness, and proximity to oceans, and the road distance from the centroid of each urban area to the nearest regional node is included as a proximity variable.

The regional market area equation emphasizes economic heterogeneity related to purchasing power and industry by including time-lagged variables for aggregate income and the regional employment share devoted to agriculture, manufacturing, and services. Proximity is once again included by the travel distance from the center of the regional node to the central place node. The central place level equation is sparse, including only a time-lagged heterogeneity variable measuring the variety of goods available in the central place region, to indicate which of these principal urban areas may have more types of goods than the others.

We estimate cross-sectional and panel models. In addition to the base models, we include proximity to other settlements in the model through a weights matrix containing inverse distances in the urban area level of the model. Following Baltagi et al. (2014), we write the spatial lag panel data model as:

$$y_{ijkt} = \rho \sum_{g=1}^N \sum_{h=1}^{M_g} \sum_{l=1}^{O_{gh}} w_{ijk,ghl} y_{ghlt} + A_{ijkt} \pi_{jk} + \varepsilon_{ijkt} \quad (4)$$

where  $i = 1, \dots, N$ ,  $j = 1, \dots, M$ ,  $k = 1, \dots, O$ , and similarly for  $g$ ,  $h$ , and  $l$ , as well as  $t = 1, \dots, T$ . The dependent variable  $y_{ijkt}$  is the population change of urban area  $i$ . The interaction with regional market areas  $j$  and central place market areas  $k$ , at time  $t$ , is modeled through equations similar to equations (2) and (3) above. The symbols  $A_{ijkt}$ ,  $X_{jk}$ , and  $Z_k$  represent the explanatory variables at the urban area, regional market area, and central place market area levels, respectively, and  $\pi_{jk}$ ,  $\beta_k$ , and  $\gamma$  are corresponding vectors of parameters to be estimated. The spatial lag is included with the scalar  $\rho$  and the weights matrix is created from the individual  $w_{ijk,ghl}$  elements. Finally, the error component structure of the panel model,  $\varepsilon_{ijkt} = u_k + r_{jk} + e_{ijk} + v_{ijkt}$ , introduces nested random effects to the model.

## References

- 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.
- Bosker, M. and E. Buringh (2011). City seeds: Geography and the origins of European cities. Technical report, CEPR Discussion Paper 8066.
- Dobkins, L. and Y. M. Ioannides (2001). Spatial interactions among US cities: 1900–1990. *Regional Science and Urban Economics* 31(6), 701–731.
- Le Gallo, J. and C. Chasco (2008). Spatial analysis of urban growth in Spain, 1900–2001. *Empirical Economics* 34(1), 59–80.
- 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.
- 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.