Geographic variation in US mortality 2004-2008: A spatial Durbin approach

Tse-Chuan Yang[1,2,3], Aggie Noah[2,4], and Carla Shoff[2,3]

Extended Abstract for 2013 Annual Meeting of the Population Association of America
[Please Do Not Cite without Author's Permission]

Short Abstract

Identifying the determinants of mortality in the US counties is not an unexplored area; however, previous studies often ignored the well-documented spatial dependence of mortality and focused on the relationships between mortality and explanatory covariates within a county. We challenge the literature by arguing that the mortality of a certain county should be associated with the features of its neighboring counties, and examine our argument with spatial Durbin modeling. Our theoretical framework is drawn from spillover and relative deprivation perspectives, and substantively, we found that the mortality of a focal county is positively related to health insurance coverage rates and affluence in neighboring counties and negatively associated with neighbors' social capital and income inequality. The former echoes the relative deprivation viewpoint, whereas the latter confirmed the spillover perspective. Methodologically, spatial Durbin modeling outperformed the traditional analytic approaches in ecological mortality research–ordinary least square, spatial error, and spatial lag regression.

[1] Department of Biobehavioral Health, Penn State University
[2] Population Research Institute, Penn State University
[3] Social Science Research Institute, Penn State University
[4] Department of Sociology, Penn State University

Introduction

Mortality is an overall assessment of population health in an area. In the past eight decades, the United States (US) has witnessed an exceptional decrease in mortality, from almost 20 deaths to roughly 8 deaths per 1,000 population.[1] Despite the decrease in overall mortality over the years, disparities in mortality have persisted along various dimensions, such as gender, race/ethnicity, and one understudied dimension, geographic space.[2] For example, Cossman and colleagues reported that the spatial patterns of all-cause mortality in the US counties have persisted in past 35 years and they further suggested that "spatial autocorrelation must be explained by ecological mortality models" so that policymakers could determine "where best to direct limited resources to specific unhealthy regions."[3]

Identifying the determinants of mortality is not an unexplored area. Briggs and Leonard examined the role of ecological structure in predicting variation in mortality, and found that characteristics at the ecological level are strongly associated with mortality.[4] They concluded that ecological socioeconomic characteristics (i.e., socioeconomic disadvantage) were strongly associated with mortality, yet a substantial portion of mortality variation was not adequately explained by socioeconomic characteristics alone.[4] Despite the weak ability to explain mortality variation, socioeconomic conditions at the ecological level (e.g., poverty at tract or county level) have driven the majority of subsequent mortality research following Briggs and Leonard. Studies found that residents in socioeconomically disadvantaged areas have higher rates of mortality compared to their more affluent counterparts.[5-7] Some scholars have also begun to investigate other potential factors, other than socioeconomic characteristics, that may affect mortality[8] by revisiting the earlier finding by Briggs and Leonard. To explain the remaining mortality variation, scholars have started to explore other potential factors, such as social capital,[9-11] income inequality,[12-14] and rurality.[2]

Although the studies discussed above have advanced our understanding of what the determinants of mortality are, they largely overlooked two issues that may either undermine their conclusions or limit the scope of mortality research. <u>First</u>, as noted by Cossman et al.,[3] mortality is an ecological and spatial feature of a population in a specific area, but a large body of mortality research has not yet incorporated a spatial perspective in investigating the relationship between contextual characteristics and mortality rates (*cf.* [2, 15, 16]). Without a spatial perspective, the persistent mortality pattern in the US may not be explained and the previous findings may, thus, have relatively few implications for policymakers. As Voss and his colleagues argued,[17-19] demography is essentially a spatial science, and demographers should be concerned with the potential biases associated with ecological data that have regularly been used since the emergence of population studies. These biases mainly come from the spatial structure of ecological data and may lead to incorrect estimates of the associations between independent and dependent variables, and hence improper conclusions.[2, 18, 20]

In addition to the methodological shortcoming, the <u>second</u> issue that has not been addressed in previous work on finding the determinants of mortality is that these studies have limited the context to the "immediate context" without investigating the impacts of the explanatory variables in "adjacent areas." Neglecting adjacent neighbors may undermine the understanding of spatial mortality disparity. This theoretical issue has been emphasized in a review where the author urged health researchers to move beyond typical within-context effects.[21] Most, if not all, previous mortality studies attempted to explain the mortality rate of a given area *only* with the characteristics *within* this area–a micro-demography approach. We argue this approach may not be applicable in ecological mortality research as mortality in the US has been found to be dependent spatially.[2, 22] In this study, we challenge the literature by proposing that the determinants of mortality of a certain area could be explained not only by the features of this area, but also by the characteristics of the surrounding neighbors, which is also known as a spillover effect. To our knowledge, no mortality research has attempted to confirm our argument, though the importance of neighbors has recently drawn researchers' attention.[23, 24]

Our argument is grounded in the following theories. Social processes are spatially embedded, and social relationships are likely to occur across the physical neighborhood boundaries. Thus, "neighbors," broadly defined, play an important role in understanding the social process within an area beyond its own context. There are at least two potential mechanisms in which neighboring locations matter. First, local institutional resources and changes occurring in a place can spill-over to its neighbors. With this perspective, we can hypothesize that a high level of positive social conditions (e.g., social capital) in an area will spillover to its neighboring locations and hence, affect the outcome of interest (i.e., mortality) in neighboring locations. Second, areas may have limited access to various resources and they may compete with one another to secure the limited resources. This can lead to relative deprivation in which the characteristics of neighbors can create dissatisfaction.[25, 26] Based on this perspective, we can speculate that a high level of positive social conditions in an area may create a sense of "relative deprivation" for its neighbors and in turn, influence neighbors' outcome of interest. As discussed previously, whether the social conditions in an area would have a spillover effect (positive effect) or create a relative deprivation effect (negative effect) has yet to be empirically tested.

The goal of this study is to test our argument and address the methodological shortcoming discussed previously with county level mortality data. Specifically, we will investigate if the mortality rate of a county is associated with the features of surrounding counties after accounting for the characteristics of the county. This study will directly respond to the call for actions by Cossman et al. (2007) and advance our understanding of the determinants of mortality by moving beyond the effects of immediate context to those of adjacent places. We will use spatial Durbin modeling[27] to reach our goal.

Method and Data

The spatial Durbin model[27] has been proven to outperform the spatial lag and spatial error model, two widely used spatial analysis methods. Specifically, it has been demonstrated that the spatial Durbin model is "the only means of producing unbiased coefficient estimates," regardless of the true spatial processes underlying the observed data.[28] Moreover, under the spatial Durbin analytic framework, there is no restriction imposed on the magnitude of the spatial effects, and both global and local spillover effects are produced.[29] These advantages have made the spatial Durbin model the state-of-the-art method of spatial econometrics, and should be further promoted in applied research.[28]

Three components comprise a spatial Durbin model: a spatial lagged dependent variable, a set of explanatory variables of a spatial unit, and a set of spatial lagged explanatory variables, which can be expressed as:

$$y = \rho W y + \alpha l_n + X\beta + WX\theta + \varepsilon$$
$$\varepsilon \sim N(0, \sigma^2 I_n),$$

where y denotes an n x 1 vector of the dependent variable (i.e., mortality), W is the spatial weight matrix, Wy represents the spatial lagged dependent variable (endogenous interaction relationships), ρ denotes an 1 x n vector of the effects of Wy, and l_n indicates an n x 1 vector of ones associated with the intercept parameter α. X represents an n x k matrix of k explanatory variables, which are related to the parameters β; WX reflects the spatial lagged explanatory variables (exogenous interaction relationships), and θ denotes an k x 1 vector of the effects of WX. The error tem, ε, follows a normal distribution with a mean 0 and a variance $\sigma^2 I_n$, where I_n is an n x n identify matrix. The formula above clearly demonstrates that the characteristics of a specific place (county in this study) and its neighbors are simultaneously considered in the analysis. This paper will use this approach to explore whether the mortality rate of a county is related to the features of its neighbors and if so, to answer how they are associated.

Drawing from the Compressed Mortality Files maintained by the National Center for Health Statistics,[30] we calculated the county-level five-year (2004-2008) age-sex adjusted mortality rate as our dependent variable. We identified six groups of independent variables. *Rural/urban residence* is measured with the rural-urban continuum codes developed by the Economic Research Service (ERS) of the US Department of Agriculture.[31] The coding scheme is from 1 to 9 where 1 indicates the most urbanized county and 9 is the most rural county. The percentages of non-Hispanic Black, Hispanic, and non-Hispanic other races were included in the group of *racial compositions.* The percentage of non-Hispanic white was not included to avoid multicollinearity. We also considered the *health care infrastructure* of a county and measured this concept with the percent of population with health insurance, the total number of medical doctors per 1,000 population, and the total number of hospital beds per 1,000 population. Following Sampson et al. (1997), we created two variables to describe the *socioeconomic conditions* of a county: social affluence and aggregate disadvantage.[32] As for the concept of social capital, four indicators were used: social capital index developed by Rupasingha and colleagues (2006), violent and property crime rates, and residential stability.[33] These variables have been found to be related to county-level mortality and to account for the geographic mortality disparity.[2] The last independent variable is *income inequality*, which was measured with the Gini coefficient to understand the income distribution in a county.

Preliminary Findings

While the spatial Durbin model is the major model used in this study, we estimated three additional regression models (OLS, spatial lag and spatial error model) to demonstrate that spatial Durbin model empirically outperforms the conventional spatial regression methods. The spatially lagged independent variables (features of neighboring counties) will be created based on the first-order Queen adjacency matrix where two counties are defined as neighbors if they share a boundary or a vertex geographically. As the spatial weights are assigned equally to each neighbor, the spatial lag independent variables could be understood as the average value of the independent variables among neighbors. We summarized the four regression models in Table 1. Several findings are notable: <u>First and foremost</u>, spatial Durbin model has the lowest Akaike Information Criterion (AIC) value, indicating that it fits our data best. The likelihood ratio tests further confirmed this conclusion. That said, taking the features of neighboring counties into account provides a statistically meaningful improvement and explains the spatial mortality variation better than other conventional methods. <u>Second</u>, we found strong evidence to support our argument that the features of surrounding counties are important. Percent of population with health insurance and affluence followed our relative deprivation argument. Specifically, within a particular county, percent of population with health insurance was negatively related to mortality; however, the percent of population with health insurance in neighboring counties was positively associated with the mortality in this particular county. That means, in contrast to neighbors, if a county was featured by a lower health insurance coverage rate, the mortality rate of this county is likely to increase; and this relationship holds, even after controlling for other county level features. In the same vein, one unit increase in the affluence score was associated with a 0.5 death per 1,000 population decrease within a county, but the same increase in the affluence score of neighboring counties would increase mortality rate by 0.2 death per 1,000 population. Third, the Gini coefficient seemed to have a "spillover effect," which is the most profound impact from neighboring counties on mortality. If the inequality in neighboring counties increased by 0.1 unit, the mortality rate of a specific county would decrease by 0.3 death per 1,000 population. It should be noted that inequality was positively associated with mortality within a county as the literature suggested.

Fourth, comparing the spatial Durbin model with the other three regression models, we found that the positive association between percent of Black population and mortality disappeared after accounting for the features of neighboring counties. This finding implied that the positive association found in other regression methods may be due to the spatial clustering of various social dimensions. Fifth, we did not find the spillover effects for the total number of medical doctors and hospital beds per 1,000 population; and this may be explained by the fact that patients prefer local providers to avoid an additional burden and inconvenience coming along with travel. Finally, high social capital index was irrelevant to mortality within a county. However, the mortality of a county would decrease by roughly 0.2 death per 1,000 population with an increase of 1 unit in neighboring social capital index score, another piece of evidence to bolster our hypothesis that spillover effect from neighbors really counts.

Table 1. Preliminary results of spatial Durbin model

	OLS Model	Spatial Lag Model†	Spatial Error Model‡	Spatial Durbin Model	
	Estimate	**Estimate**	**Estimate**	**Estimate**	**Lag Estimate**
Intercept	7.969***	4.825***	8.703***	5.625***	
Rural/Urban Residence					
Rurality	-0.055***	-0.040***	-0.030**	-0.007	-0.027
Racial/Ethnic Composition					
Non-Hispanic Black	1.393***	0.849***	1.342***	0.421	-0.045
Hispanic	-3.019***	-1.449***	-1.769***	-0.904**	-1.516***
Other Races	0.679*	1.498***	1.922***	2.429***	-3.407***
Health Care Infrastructure					
Health Insurance	-0.001	-0.014***	-0.030***	-0.041***	0.057***
MD per 1,000	-0.035*	-0.032*	-0.032*	-0.043**	-0.019
Hospital Beds per 1,000	0.021***	0.017***	0.014***	0.015***	0.010
Socioeconomic Status					
Affluence	-0.658***	-0.495***	-0.591***	-0.529***	0.208***
Disadvantage	0.390***	0.294***	0.333***	0.316***	0.006
Social Capital					
Social Capital Index	-0.189***	-0.099***	-0.078***	-0.027	-0.169***
Violent Crime Rate	0.039*	0.017	0.025	0.016	0.026
Property Crime Rate	0.003	0.003	0.003	-0.001	-0.012
Stability	-0.040	-0.028	-0.015	0.003	-0.027
Income Inequality					
Gini Coefficient	2.840***	1.496**	1.782**	2.071***	-3.409***
Rho (Spatial Lag)		0.428***		0.439***	
Lambda (Spatial Error)			0.586***		
AIC	8752.7	8243.6	8229.4	8066.4	

*$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$

†Likelihood Ratio (LR) tests indicated that Spatial Durbin fits data better then Spatial Lag Model (LR=205.11, DF=14, p≤0.001)

‡ Likelihood Ratio (LR) tests indicated that Spatial Durbin fits data better then Spatial Error Model (LR=190.93, DF=14, p≤0.001)

References

1. Hoyert DL. *75 Years of Mortality in the United States, 1935–2010.* Hyattsville, MD: National Center for Health Statistics;2012.
2. Yang TC, Jensen L, Haran M. Social Capital and Human Mortality: Explaining the Rural Paradox with County-Level Mortality Data. *Rural Sociology.* 2011;76(3):347-374.
3. Cossman JS, Cossman RE, James WL, Campbell CR, Blanchard TC, Cosby AG. Persistent clusters of mortality in the United States. *American Journal of Public Health.* 2007;97(12):2148-2150.
4. Briggs R, Leonard WA. Mortality and ecological structure: a canonical approach. *Social Science & Medicine (1967).* 1977;11(14):757-762.
5. Kaplan GA, Pamuk ER, Lynch JW, Cohen RD, Balfour JL. Inequality in income and mortality in the United States: analysis of mortality and potential pathways. *BMJ: British Medical Journal.* 1996;312(7037):999-1003.
6. Ezzati M, Friedman AB, Kulkarni SC, Murray CJL. The reversal of fortunes: trends in county mortality and cross-county mortality disparities in the United States. *PLoS medicine.* 2008;5(4):e66.
7. Mansfield CJ, Wilson JL, Kobrinski EJ, Mitchell J. Premature mortality in the United States: the roles of geographic area, socioeconomic status, household type, and availability of medical care. *American Journal of Public Health.* 1999;89(6):893-898.
8. Kawachi I, Berkman LF. *Neighborhoods and Health*. New York: Oxford University Press; 2003.
9. Kawachi I, Kennedy BP. The relationship of income inequality to mortality: Does the choice of indicator matter? *Social Science & Medicine.* 1997;45(7):1121-1127.
10. Kennedy BP, Kawachi I, Glass R, Prothrow-Stith D. Income Distribution, Socioeconomic Status and Self-rated Health in the United States: Multi-level Analysis. *British Medical Journal.* 1998;317(7163):917-921.
11. Lochner KA, Kawachi I, Brennan RT, Buka SL. Social capital and neighborhood mortality rates in Chicago. *Social Science & Medicine.* 2003;56(8):1797-1805.
12. Kawachi I, Kennedy BP. Income inequality and health: Pathways and mechanisms. *Health Services Research.* 1999;34(1 Pt 2):215-227.
13. Marmot M. *The Status Syndrome: How Social Standing Affects Our Health and Longevity*. New York: Henry Holt and Company; 2004.
14. Yang TC, Yi-Ju Chen V, Shoff C, Matthews SA. Using quantile regression to examine the effects of inequality across the mortality distribution in the US counties. *Social Science & Medicine.* 2012;74(12):1900-1910.
15. Sparks JP, Sparks CS, Campbell JJA. An application of Bayesian spatial statistical methods to the study of racial and poverty segregation and infant mortality rates in the US. *GeoJournal.* 2012:1-17.
16. Sparks PJ, Sparks CS. An application of spatially autoregressive models to the study of US county mortality rates. *Population, Space and Place.* 2010;16(6):465-481.
17. Voss PR, Curtis White KJ, Hammer RB. Explorations in Spatial Demography. In: Kandel WA, Brown DL, eds. *Population Change and Rural Society*. Dordrecht, Netherlands: Springer; 2006:407-429.
18. Voss PR, Long DD, Hammer RB. County Child Poverty Rates in the US: A Spatial Regression Approach. *Population Research Policy Review.* 2006;25:369-391.
19. Voss PR. Demography as a spatial social science. *Population Research and Policy Review.* 2007;26(5):457-476.

20. Haining R. *Spatial data analysis: theory and practice*. Cambridge: Cambridge University Press; 2003.

21. Dietz RD. The estimation of neighborhood effects in the social sciences: An interdisciplinary approach. *Social Science Research.* 2002;31(4):539-575.

22. James WL, Cossman RE, Cossman JS, Campbell C, Blanchard T. A brief visual primer for the mapping of mortality trend data. *International Journal of Health Geographics.* 2004;3(7):1-17.

23. Auchincloss AH, Roux AVD, Brown DG, O'Meara ES, Raghunathan TE. Association of insulin resistance with distance to wealthy areas: the multi-ethnic study of atherosclerosis. *American Journal of Epidemiology.* 2007;165(4):389-397.

24. Takagi D, Ikeda K, Kawachi I. Neighborhood social capital and crime victimization: Comparison of spatial regression analysis and hierarchical regression analysis. *Social Science & Medicine.* 2012;Forthcoming.

25. Ginther D, Haveman R, Wolfe B. Neighborhood attributes as determinants of children's outcomes: how robust are the relationships? *Journal of Human Resources.* 2000;35(4):603-642.

26. Firebaugh G, Schroeder MB. Does Your Neighbor’s Income Affect Your Happiness? *American Journal of Sociology.* 2009;115(3):805-831.

27. Anselin L. *Spatial Econometrics: Methods and Models*. Cordrecht: Kluwer Academic Publishers; 1988.

28. Elhorst JP. Applied spatial econometrics: raising the bar. *Spatial Economic Analysis.* 2010;5(1):9-28.

29. LeSage JP, Pace RK. *Introduction to spatial econometrics*: Chapman & Hall/CRC; 2009.

30. NCHS. Compressed Mortality File, 2004-2008 (machine readable data file and documentation, CD-ROM series 20, No.2M)Hyattsville, Maryland: National Center for Health Statistics; 2011.

31. Rural-Urban Continuum Codes. United States Department of Agriculture; 2003. http://www.ers.usda.gov/briefing/rurality/ruralurbcon/.

32. Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science.* 1997;277(5328):918-924.

33. Rupasingha A, Goetz SJ, Freshwater D. The production of social capital in US counties. *Journal of Socio-Economics.* 2006;35(1):83-101.