Future Directions in Spatial Demography

KATHERINE J. CURTIS
Department of Sociology
University of Wisconsin, Madison
Email: kcurtis@ssc.wisc.edu

My primary concerns within each of the areas to be elaborated in the workshop on Future Directions in Spatial Demography are reflected in two specific issues: advances in spatially-explicit theory; and advances in space-time analysis.

With regard to advances in spatially-explicit theory, I suggest that spatial heterogeneity is too often overlooked in demographic research aiming to address spatial effects. In my research and teaching, I build from the position that spatial effects are either a heterogeneity process or a dependence process (e.g., Voss, Curtis, and Hammer 2006). In my read of the literature, spatial dependence is given more attention relative to spatial heterogeneity. However, demographers and other social scientists do not always theoretically justify or fully explain the theoretical implications of using a spatial dependence model. There are exceptions, of course, as some theories of fertility, mortality, and migration invoke a dependence argument (i.e., the diffusion of fertility practices, the spread of infection, and the development of migration networks). Yet there is too often a mismatch between the theoretical model asserted and the statistical model used to test it, where theory suggests heterogeneity while the analytical approach suggests dependence.

There are two potential reasons for the mismatch, based on my interactions with graduate students as well as early- and mid-career scholars and researchers in my seminar and workshops. First, spatial heterogeneity is relatively more conceptually complicated. I do not claim that spatial dependence is simple to statistically model or substantively explain. However, in my experience, many if not most students and workshop participants come primed to analyze their data using a spatial lag regression model, assuming some “neighborhood effect” is driving the outcome. Even those who come with little interest in the spatial effect, per se, but with an eye toward space as nuisance approach the “problem” as one of spatial dependence, often thinking the spatial lag regression model is most appropriate because the problem is derived from dependence in the data structure. This approach can result in a statistically poor-fitting model. Of equal if not greater importance to population and health research is this: immediately resorting to a spatial dependence frame excludes the potential for a deeper substantive knowledge of the spatial process at play and, correspondingly, meaningful theoretical development.

Spatial heterogeneity can be conceptually complicated because of its multiple meanings. Spatial heterogeneity can refer to spatial variation in the distribution of an attribute. It also can refer to spatial variation in relationships between attributes. I assert that there is much promise in exploring spatial effects through both forms of spatial heterogeneity. In the first form, one is positioned to “explain” the spatial effect through measurable attributes. In the second form,
one is positioned to reconcile inconsistent findings and identify contexts in which processes differ (e.g., Curtis, Voss, and Long 2011).

Accomplishing either task is not necessarily simple or straightforward, yet it is essential to develop theory. In the first form, one must think carefully about the potential source of heterogeneity; it is measureable or immeasurable? Is it a product of a demographic process or is it an artifact of the data (measurement error)? In the second form, one must think carefully about the interplay among multiple attributes that give rise to different contexts and, ultimately, to different relationships between key demographic factors. In both forms, the researcher is challenged to think about the qualities of spatial units in terms of social, economic, environmental, political or geophysical attributes that are relevant to the demographic process or event of interest. That is, what precisely is “space” and why is it relevant? An important consequence of fully exploring spatial heterogeneity is more fully developed theory.

The second source of mismatch between conceptual and statistical models is that there are many ways to explore spatial heterogeneity and not all of them are considered “spatial” methods. Consequently, researchers exploring spatial dimensions seem to limit themselves to “spatial” modeling strategies despite the conceptual model. At the same time, researchers uninterested in spatial dimensions do not speak in terms of space, leaving the conceptual model underdeveloped. A simple illustration is the student who finds evidence of spatial autocorrelation in his data through a Moran’s I or similar statistic. He analyzes his data with an ordinary least squares regression and finds no evidence of spatial autocorrelation in the residuals. He is disappointed because there is no longer a spatial effect. In fact, this outcome would be disappointing if the theoretical model claimed spatial dependence. Yet this outcome could be consistent with a theoretical model asserting spatial heterogeneity; the independent variables have identified the spatial patterning in the dependent variable. One does not equate OLS with spatial data analysis. However, an OLS can be used to test a spatially-informed conceptual model, one in which the spatial effect is a form of spatial heterogeneity.

The mismatch between conceptual and statistical models is rooted in training. I assert we, spatial demographers, should prioritize conceptual development in our training by integrating demographic theory and spatial analytical techniques. We have made significant advances in providing training in spatial analytical methods through courses and workshops. I suggest we integrate spatially explicit or implicit theories into already existing offerings and/or develop companion courses and workshops that address spatial theory. It is within this domain that we can make significant progress in addressing a variety of issues including, for example, measuring place and the interrelationships among places, assessing the validity of measures of neighborhood and context, and determining the most effective way to visualize and analyze data (including tools not yet available).

With regard to advances in space-time analysis, I suggest that we still lack the necessary tools to address the most basic demographic questions. Much of demographic research concerns trends—changes in trends, divergent trends, and factors underlying trends. We have the ability to describe trends among spatial units. However, we do not have available accessible,
statistically rigorous tools to draw statistical inference or to simultaneously analyze temporal and spatial processes. The former is necessary to assess whether changes or differences are meaningful or negligible. The latter is necessary to correctly identify what is driving the trend. There have been significant advances in spatial regression in recent decades, in types and accessibility, yet space-time regression has not matured at the same rate or to the same degree. Moreover, statisticians developing the analytical techniques are disconnected from the demographic user community wanting and needing to apply the techniques.

This disconnect has stalled scientific progress. Researchers are unable to answer questions relevant to the discipline and to social policy because the analytical tools are not accessible to them. In my own experience, I began exploring spatial data analysis techniques in 2003. At that time, I took advantage of the descriptive tools available, engaged the literature on advances in space-time regression, and was informed that such advances would soon be implemented into existing software programs tailored to spatial analysis. In 2008, five years later, I published the work I began in 2003 without the intended space-time regression approach because the tools were still unavailable. Also in 2008, I sought a statistics collaborator to develop techniques that would enable me to answer questions about change in space over time. Now, in 2011, the collaboration is bearing fruit; we have one manuscript under review developing the approach (e.g., Reyes, Zhu, and Curtis 2011, developed in R).3

Such synergistic collaborations are important, appear to be necessary, should yield publically available data analysis tools, and should be sponsored by funding agencies or collectives already supported by these agencies. I argue priority should be given to space-time techniques that enable statistical inference because space-time questions are foundational to population and health research. At present, we are unable to address the most basic questions at demography’s core because tools that are within reach are not available to all members of the demographic community.

Citations