



Forthcoming in Michael F. Goodchild and Donald G. Janelle (eds.), *Spatially Integrated Social Science: Examples in Best Practice* (New York: Oxford University Press)

Title: **The Role of Spatial Analysis in Demographic Research**

Author: John R. Weeks
Department of Geography
San Diego State University
San Diego, CA 92182-4493
(john.weeks@sdsu.edu),

August 2001

Acknowledgments: This work was supported by grants from the Andrew Mellon Foundation and the National Science Foundation

The Role of Spatial Analysis in Demographic Research

John R. Weeks
San Diego State University

INTRODUCTION

Demography is an inherently spatial science since it almost always deals with human populations in a defined geographic region, but spatial analysis has thus far played only a small role in the development and testing of demographic theory. There are several reasons for this, including the recency of many of the more useful spatial statistical approaches, and the fact that most people practicing demographic science are not in geography and have not been encouraged to think spatially. Yet, even in geography, few population specialists adopt specifically spatial approaches to their research beyond the measurement of the movement of people from one region to another, or the comparison of demographic trends among different regions.

In the past few decades, demographic research has focused particularly on the analysis of survey data drawn from interviews conducted at the household level and, as a consequence, theory has focused heavily on individual-level influences on demographic behavior. The development of surveys such as the National Survey of Family Growth in the United States, and the U.S.-funded Demographic and Health Surveys in less developed nations represented an important step in demographic research because the previous heavy reliance on aggregated data left gaps in our knowledge about how individuals think and behave. Now, however, the confluence of powerful geographic information system technologies, advances in the design of spatial statistics, and the increasing availability of geo-referenced databases has vastly improved

the ability of demographers to think spatially. As a result, there is a reawakening of interest in models of human behavior that place individuals in the environmental context of space and time.

Demography is not only spatial, but it is also by nature interdisciplinary. The demographic transition, which provides the organizing framework for most demographic research, is really a complex set of transitions, each of which draws upon expertise in differing social science and health-related disciplines. The demographic transition usually begins with the epidemiological transition, which is the shift over time from high death rates with deaths clustered at the younger ages and caused largely by communicable diseases, to low death rates with deaths clustered at the older ages and caused largely by degenerative diseases. This sets in motion a train of other transitions. The fertility transition represents the change from high fertility levels over which people have relatively little direct control to low fertility over which people have considerable control. The migration transition is initially the response to population growth in rural areas which causes people to seek opportunity elsewhere, typically urban places, thus unleashing the urban transition, in which a population moves from being largely rural to being largely urban. The age structure transition is a predictable result of changes in mortality and fertility in which high mortality and high fertility produce a very young age structure that is pyramid-shaped, whereas the declines in both mortality and fertility produce bulges in the young adult ages, leading eventually to a barrel-shaped age structure. The family and household transition represents the change from complex forms of family and household structure when mortality and fertility are both high, to less variability in the middle of the transition, to new forms of complexity when both fertility and mortality are low. Finally, of course, there is the overall transition in population size that occurs when mortality declines sooner than fertility (the

usual pattern in the demographic transition) from which massive changes follow with respect to resource utilization and allocation.

Each of these interrelated aspects of demographic change has a spatiotemporal component which, when understood, adds to our knowledge of how and why these transitions occur. Furthermore, each of these different aspects of the demographic transition draws attention, as appropriate, from sociologists, economists, geographers, regional scientists, public health researchers and practitioners, and a host of other disciplines. In fact, very few people in the field of demography actually have advanced degrees in the named field of demography, and there are very few academic departments of demography in the world. Instead, demographic research is conducted as a sub-discipline of nearly every one of the social and health sciences fields, and researchers in these various disciplines are then frequently associated with academically-based population centers.

The fact that demography is spatial by nature means that much, if not most, of the demographic research that is conducted has a spatial “awareness,” even if relatively little of it engages spatial “analysis” in any formal sense. Spatially aware research understands that demographic behavior will differ by geographic region—that population characteristics and change are different in urban than in rural places; that countries in sub-Saharan Africa with a high proportion of Muslims have lower HIV/AIDS prevalence rates than predominantly non-Muslim nations; that East Asian countries have experienced a different fertility transition than South Asian countries. All migration research—which has historically been the staple of population geographers—has a built-in spatial awareness because the analyses focus on the places from which migrants come and the places to which they go. Migration matrices and multi-regional life tables have been created as tools that increase our quantitative understanding

of these changes involved in migration. But such spatial awareness is not quite the same as spatial analysis because it is not typically associated with underlying theories and hypotheses about spatial patterns that are designed to be tested for their specific spatial content.

In this chapter I first offer a general framework for the application of spatial analysis to demographic research as a way of integrating and better understanding the different transitional components of the overall demographic transition. Then I discuss the kinds of data that are required for spatial demographic analysis, allowing researchers to use the concepts and tools of spatial analysis to test theories growing out of the general framework that I have laid out. Finally, I summarize some of the work that I and my colleagues have been doing in Egypt, searching for an improved understanding of the Arab fertility transition through the testing of explicitly spatial hypotheses about the timing and tempo of fertility change.

A GENERAL FRAMEWORK FOR SPATIAL ANALYSIS IN DEMOGRAPHY

Spatial analysis can be defined as a quantitative data analysis in which the focus is on the role of space and which relies on explicitly spatial variables in the explanation or prediction of the phenomenon under investigation (Cressie 1993; Chou 1997). Spatial analysis in the social sciences tests theories that where you are makes a difference in social attitudes and behavior, and that observed differences in the social world are not distributed in a spatially random pattern. Cressie (1993) argues that the classical, nonspatial data analysis should actually be seen as a special case of spatial data analysis. Viewed in this way, the underlying logic is that each random variable (z) is associated with locational attributes (x and y). In spatial data analysis, the researcher uses geostatistics to glean information from the x and y coordinates, whereas in classical statistical analysis the researcher ignores those coordinates (often not even realizing that

they might exist). More to the point, in classical statistical analysis, the locational attributes are considered to be a nuisance, rather than representing useful information. Spatial autocorrelation follows Tobler's "First Law of Geography": Everything is related to everything else, but near places are more related than far places (Tobler 1970). In classical statistical analysis, this is something to be gotten rid of, or controlled for, whereas in spatial data analysis it becomes an object of investigation. If spatial autocorrelation exists, then there may be spatial dependence; and thus something of interest spatially that is occurring.

The comments about spatial autocorrelation also apply to temporal autocorrelation (things that are close to one another temporally are more likely to be similar than things that are more temporally distant). Econometricians have developed autoregressive models to account for the temporal autocorrelation that is typically found in time-series data that comprise the backbone of much of economic analysis. Time is a disturbance to be controlled, not an effect to be studied.

In order to think spatially with regard to demographic research, it is useful to keep in mind the suggestion of Star and Estes (1990) that spatial analysis can be divided into two "families": (1) analysis that is concerned with *local or neighborhood characteristics*; and (2) analysis that is concerned with *connections* between locations. This distinction provides a useful way of organizing our thinking into a general framework, as I show in Figure 1. In demographic research we can think of the neighborhood characteristics as representing aspects of the context in which demographic decisions are made and demographic behavior is manifested. Spatial analysis then looks for place-specific factors that influence the behavior of otherwise similar people. The connections relate to the kinds of networking and interaction that lead both to *diffusion* (the spread of ideas) and *dispersal* (the geographic redistribution of people). Spatial

analysis then searches for the timing and direction of the connections, and seeks to understand their causes and consequences. Let me discuss these general concepts in more detail.

FIGURE 1 ABOUT HERE

Spatial Analysis Based on Environmental Context

One of the theoretically more robust ways in which spatial analysis is beginning to enter demography theory is in an updated version of human ecology that is often referred to in the literature as “environmental context.” From a human ecological perspective, this means that population size and characteristics interact with social organization, and with the environment and technology, to produce the behavior that constitutes human society. In turn, human behavior influences population, organization, the environment, and technology and for this reason the concept is that of a system, a human ecosystem (Namboodiri 1988; Micklin and Sly 1998).

Social scientists tend to focus on the population and social organizational parts of this system, and spend less time thinking about the environment in which these parts are embedded. In particular, sociologists and demographers tend to be vague, if not dismissive, of the built environment--of the buildings, parks, roads, bridges, and the associated infrastructure that humans create out of the natural environment and which become the places where everyday life takes place. Micklin and Sly (1998) put the built environment under technology, representing one set of "tools" available to human society. Yet, the built environment is more than that--it is the actual environment in which a large fraction of humans spend their entire lives. The natural environment is so transformed by urbanization that the majority of urban residents spend little time touching soil and interacting with flora and fauna. Even more importantly, the built environment is not just a product of human activity; it is also a very important element of what

Namboodiri (1988) has called the goal of human ecology, which is "to identify the linkage between the dynamics of human interdependence and the pursuit of the art of living" (p. 622).

There can be little doubt, of course, that national and regional events affect things like fertility levels regardless of where a person lives (see, for example, Fargues 1997), and that events outside an area can be instrumental in producing change at the local level (Courbage 1994). But ultimately it is at the local level that the actual decisions are made that lead to the specific behavior that actually determines what the regional and national fertility levels are going to be. Duncan made the classic statement of this in 1959:

A concrete human population exists not in limbo but in an environment. Moreover, to continue to exist, it must cope with the problems posed by an environment which is indifferent to its survival but offering (in varying degree) resources potentially useful for the maintenance of life. By mere occupancy of an environment, as well as by the exploitation of its resources, a human population modifies its environment to a greater or lesser degree, introducing environmental changes additional to those produced by other organisms, geological processes, and the like. Thus, in the language of bioecology, one may say that not only does the environment "act" upon the population but also the human population "reacts" upon its environment...The "adjustment of a population its environment, therefore, is not a state of being or static equilibrium but a continuing, dynamic process (Duncan 1959:681-682).

When Duncan uses the word "environment" he is referring to the natural environment, in the way that human ecologists have tended to do. But, a substitution of "built environment" for

"environment" keeps the meaning while applying it specifically to human life as organized in cities, towns and villages throughout the world. And when we use the term local context or local environment we mean the complex of social activities that are taking place within a given built environment, situated within a specific natural environment.

Local context has emerged as an important way of conceptualizing inequalities in the social world (Tickamyer 2000), and this approach is exemplified in this volume by the chapters by Logan, by Sampson and Morenoff, and by Messner and Anselin. With respect to demographic research, this approach offers a way of exploring differences and inequalities in fertility, mortality, age structure, and family and household structure. Note that I have put a two-headed arrow in Figure 1 to indicate the reciprocal nature of these relationships. If you want to understand spatial inequality in fertility, for example, you must realize that part of the context of that fertility behavior will be in terms of local health and mortality levels, the local age structure (influencing cohort size distortions), and the prevailing family and household structures, all of which will be related to local issues of gender equality and empowerment.

Relatively little attention has been paid thus far to analyses of fertility at the local spatial scale. There is a vast literature on fertility differentials, to be sure, but attention is paid largely to characteristics of individuals without regard to where they live. At first glance, it might seem that this is simply evidence that population geographers believe that where you live is not related to fertility. There is the nearly universal finding that fertility differs by social class (defined sometimes by income differences, by occupational distributions, by education, and sometimes by racial/ethnic differences). Since there is a tendency for there to be a geographic sorting process by social class, the spatial dimension of fertility is implicitly incorporated into that model, but little attention is otherwise given to the demographic and social variability in fertility across

space. That is to say, little attention is paid to the ecology of fertility, even among human ecologists. Rather, the emphasis is on examining fertility levels at the individual level, using data from surveys that by and large do not permit a neighborhood analysis. These studies of necessity focus attention on national comparisons or at regional differences within a country.

Scale becomes an important issue in such research, because as fertility declines over time, regional differences may disappear even though local variations remain. Indeed, even under conditions of relatively high fertility, this confounding may occur. Wilson and Wood (1991) show that in Victorian England and Wales “demographic conditions were local, rather than regional. These local patterns tend to be masked when counties, or combinations of counties, are made the framework for analysis” (p. 414). Weeks et al (2001) found the same issues of scale in the spatial patterning of fertility in Egypt, based on an analysis of Demographic and Health Survey data. Using data at the governorate (state) level, only the most general pattern of a north-south fertility differential was observable. However, within Cairo, where the sample size was sufficient for a within-state analysis, clear spatial patterns did in fact emerge.

More attention has been paid to spatial differences in various causes of disease and death. Communicable diseases by their very nature are susceptible to local differences in infrastructure and population structure. Langford et al (1999) have examined some general issues related to spatial patterns of mortality in Scotland, whereas Wallace and Fullilove (1991) examined the local spatial context of AIDS deaths in the Bronx in the 1980s. Reid (1997) used data from the 1911 census of England and Wales to show that infant and child mortality at the beginning of the twentieth century was influenced by a combination of the parents’ characteristics (especially the father’s occupation) and where the child lived. LeClere, Rogers and Peters (1998) show that neighborhood context is related to death from heart disease among women in the United States;

Robert (1998) has shown that the socioeconomic status of a person's community affects adult health over and above an individual's own sociodemographic characteristics; and Peak and Weeks (2001) have demonstrated that living in a Mexican ethnic enclave improves reproductive outcomes of Mexican-origin women in California. Similar types of studies are discussed by Gatrell and Rigby in this volume.

Although the local environmental context may be the focus of a particular spatial analysis, there may well be overlap with the diffusion/dispersal aspects of analysis. To be sure, one of the reasons why change may occur at the local level is because of the influence of networks and connections between different places. For this reason, in Figure 1 I have shown a link between these types of analysis.

Spatial Analysis Based on Networks and Connections

Demographic transition theory rests as least partially on the concept of the diffusion of innovations—including ways of preventing death, preventing births, and organizing migration flows and chains. These are processes that spread over space and across time, and are naturally prone to spatial analysis. Early references to this type of spatially demographic way of viewing the world include a study of fertility change in Nigeria by van de Walle (1965), the study of fertility differences in Spain by Leasure (1962) that actually ignited the Princeton European Fertility Project, and the later results from that project that emphasized the process of diffusion in the European fertility transition (Coale and Watkins 1986; Watkins 1991). However, all of these studies would fall into the category of “spatial awareness” rather than spatial analysis. Casterline and his colleagues used pooled time series data to demonstrate the diffusion of the fertility transition in Taiwan (Montgomery and Casterline 1993) and in Costa Rica (Rosero-Bixby and Casterline 1994). More recent studies have benefited from the availability of GIS, including

especially the work of Bocquet-Appel and Jakobi (1998) who used kriging and other geostatistical techniques to show that Peebles, in lowland Scotland, quite possibly served in the mid-nineteenth century as the “detonator to what appears to be a ‘Big Bang’ in the introduction of contraception in Great Britain” (p. 199).

A potentially important scalar subset of diffusion is the idea of networks. Diffusion is usually measured, as in the Bocquet-Appel and Jakobi study, at the aggregate level. If area A and B are not alike on characteristic z at time 1, but are more alike at time 2, even after controlling for endogenous sociocultural changes, then we infer that diffusion has occurred from area A to area B (Tolnay 1995). Some sources of change do occur at the regional level, especially those affecting mortality (such as cleaning up the water supply, or controlling mosquito populations), but some occur at the individual level, especially innovative behavior like fertility limitation, or migration decisions. The influences here may be linked to networks, which have an important spatial component, even though most theorizing about human networks has been done without taking location into account (see 1992; Burt 1999), and most analyses in demographic research can be best described as spatially aware, rather than spatially analytic (see, for example, Kohler, Behrman, and Watkins 2001). A major exception to this statement is the study by Entwisle *et al* (1997) which used spatial network analysis to study the accessibility of family planning services in a rural population in Thailand.

DATA REQUIREMENTS FOR SPATIAL ANALYSIS IN DEMOGRAPHIC RESEARCH

Demographic research that employs spatial analysis obviously requires data that are georeferenced. If data are not assigned to a location, then spatial analysis is not possible.

Virtually all demographic have some georeferencing associated with them: it is largely a question of scale. Data are almost always recorded at least at the country level (although some United Nations publications aggregate data only to the regional level), and so this provides a way of assessing spatial patterns among and between different regions of the world. Indeed, spatial clustering is a way of *defining* regions, if by that we follow the Oxford English Dictionary definition of “an area of the world made up of neighbouring countries which are considered socially, economically, or politically interdependent” (Brown 1993).

There is a limit to the usefulness of data at the national level for spatial analysis, especially given the relatively small number of countries in the world. Demographic theories are likely to be more readily tested using geographic units at the subnational level, especially if such data are available for more than one country, so that regional patterns can still be discerned, but at a finer scale. Such data work their way down the level equivalent to the census tract in the United States, or in some instances, a sub-unit of the tract, such a block or block group. Most data that are georeferenced to these types of administrative units are also aggregated statistically, creating classical problems of ecological inference: Do the results summarized for many people reflect anything about the behavior of the individuals themselves?

One way in which demographers have worked around this problem is through the use of samples of individual records from the Census. Files such as the public use microdata samples (PUMS) in the United States and Mexico, and the SARs (Samples of Anonymised Records) in the United Kingdom represent data for individuals, so that analysis can be run without aggregation, albeit at the cost of some geographic specificity, since privacy demands that you not be able to locate the person whose census record you are studying. Areal units typically are

larger (sometimes much larger) than census tracts, and so spatial questions can be answered only in very general terms.

The PUMS and SAR data are samples from the complete set of census data, and in that sense they are little different from survey data, except that the sample size is large enough to provide a reasonable amount of geographic detail. A sample of a thousand people may be quite representative of the entire voting population of the United States, but it is not large enough to tell us anything at the sub-national level. Larger samples are capable of being georeferenced, however, and the Demographic and Health Surveys (DHS) have been leaders in this direction. The design of the DHS has always been based on a cluster sample, in which a fairly large number of households in each village or neighborhood in a city will be included in the sample. This has permitted these data to be linked back to administrative units such as census tracts, so that respondents could be located within their neighborhood context. Since 1999, the DHS has supplied interviewers in several of its surveys with GPS units so that the location of each geographic cluster in which households are located can be recorded.

Administratively-derived data, such as birth and death certificates, often have addresses associated with them that permit location to be calculated by matching with street addresses, such as the TIGER file created by the US Census Bureau. Most commercial GIS software has the capability to accomplish this task.

The data from remotely-sensed images provide new opportunities to apply spatial analysis to demographic research. There is a long history in demography of using aerial photographs to aid in the estimation of population size (see, for example, Noin 1970). The general strategy is to count the number of households observed from the air, and then apply an average number of persons per household to estimate the total population. Although this does

not automatically utilize spatial analysis, new variations on the old theme do incorporate spatial analysis in order to increase the accuracy of the estimation process (Lo 1995), including the use of night-time imagery to measure population density (Sutton 1997).

More sophisticated uses of new multi-spectral imagery offer the prospect of enhancing our understanding of spatial processes for demographic purposes (Rindfuss and Stern 1998). Some of the applications include: (1) linking changes in vegetation to changes in population distribution and characteristics; (2) measuring the spread of urban areas, and measuring the environmental impact of that spread (see, for example Ridd 1995; Mesev, Longley, and Batty 1996; Ward, Phinn, and Murray 2000); (3) classifying human settlements in order to derive proxies for social structural variables (see Moran, this volume); (4) using night-time light images as a proxy for levels of rural development in less developed settings. Remote images can also be used to help physically to locate populations in rural areas that are referenced in census data as existing within the boundaries of a larger administrative unit. This provides a solution to the modifiable areal unit problem that plagues spatial research using data that are georeferenced to administrative units that are represented as polygons, rather than points. This use of the image helps to organize the data for spatial analysis, as I discuss now with reference to a study of the fertility transition in Egypt.

Illustration: Fertility Change in Egypt

In analyzing fertility change in Egypt, my colleagues and I (Weeks *et al.* 2000) have borrowed from Gadalla (1978), Namboodiri (1988), Entwisle (1989), Hill (1997), and Crenshaw and his associates (2000) the idea that an account of fertility decline must "nest fertility decision-making and micro-level behavior in their environmental contexts" (Crenshaw, Christenson, and Oakey 2000:371). The model that guides our research incorporates the assumptions that (a) the

built environment represents something tangible about the social environment; (b) the social environment influences the social and human capital variables that more directly influence the demand for children; (c) the reproductive behavior of some people within a neighborhood will influence the behavior of others, even net of the human capital opportunities that objectively exist in the neighborhood; and (d) these influences operate on reproductive levels through the mechanisms of the proximate determinants of fertility, such as age at marriage and the use of contraceptives within marriage, to determine fertility at the local level; but (e) changes in reproductive behavior at the local level may be influenced by changes in, and reciprocally influence changes in, other neighboring regions, resulting in spatial patterns of fertility transition; the consequences of which (f) ultimately determine the overall societal level of reproduction, thus creating the wider phenomenon of a fertility transition.

In this research we are interested in the extent to which the variation in fertility from one rural village (the local context) to another may be explained by a process of diffusion of behavior from some villages to others, net of the human capital variables such as education that may exist within the village. We lack direct evidence of such spatial diffusion, but can infer it from the spatial and temporal patterning of reproductive behavior. First we must show that proximity matters, and then we must show that changes occurred over time in a sequence that is consistent with spread or diffusion.

We illustrate this procedure using data for a rural governorate of Egypt (Menoufia) for 1976 and 1986. Menoufia is one of the 26 governorates that comprise the administrative regions of Egypt, roughly equivalent to states in the United States, although perhaps more analogous to Counties in the United Kingdom. For decades Menoufia has been one of the most rural and most densely populated rural areas of Egypt (Gadalla, McCarthy, and Kak 1987). It has been, and

remains, predominantly agricultural, and the high rate of population growth has increased the redundancy of the rural labor force and encouraged out-migration--to Cairo or to other Arab (especially oil-producing) nations.

The demographic data used for this study come from the 1976 and 1986 censuses of Egypt. Data were coded from the Arabic-language publications using the smallest geographic unit available in the Egyptian censuses--the shiakha or village. The "shiakha" literally refers to the area controlled by a sheikh, but in more practical terms it is the area serviced by a police post. The dependent variable is the level of fertility in each shiakha, which is an estimate of the total fertility rate, derived from age data in the census. It is measured as the net reproduction rate, which takes mortality into account, for a measure that represents the actual "supply" of children. See Arriaga (1994) for a review of the methods that we employed in these calculations. The human capital variables derived from the census included adult female illiteracy and female labor force participation rates, with a control for the percent of women who were currently married.

In attempting to model the diffusion of fertility and/or its antecedents (the human capital variables) we had to deal with an important methodological issue with regard to spatial analysis: the modifiable areal unit problem (MAUP) (Openshaw 1984). In almost every instance demographic information is gathered at some arbitrarily-defined geographic level such as a census tract or enumeration district. However, unless this area defines a small and heavily built area, its areal boundaries will include space in which people do not reside. This is especially true in agricultural areas where most space is devoted to crop, orchard, or pasture land. Thus, a rural village, even if densely populated within its own boundaries, may consume only a small portion of the administrative boundaries to which the demographic data are attached. We dealt with this

problem through the use of remotely sensed images. We employed remotely-sensed images to classify land cover by built/non-built use in order to undertake what is sometimes known as dasymetric mapping (Langford and Unwin 1994), in which information inside a zone is used to map the population density or distribution within that zone. The advantage of a classification of data from a remotely sensed image which spatially defines built areas is that it frees us from the "tyranny of an arbitrary imposed and fixed set of census geographies" (Openshaw and Rao 1995:425). The classification of remote images provides a zone-design tool that can help to resolve the modifiable areal unit problem. The details of this process are discussed in Weeks et al (2000). The results of the classification of the image allowed us to determine a unique location for each village in Menoufia, and this set of coordinates for each village was then used in the spatial analysis.

The first question of interest was whether, in fact, fertility exhibited a spatially dependent pattern in Menoufia in each of the years under study. We utilized the global spatial statistic, Moran's I, to test the null hypothesis of spatial independence. In 1976, the normalized random Z-score for Moran's I for the NRR was 6.82, indicating a statistically significant amount of spatial autocorrelation, thus leading us to reject the null hypothesis that fertility is spatially independent in Menoufia. In 1986, the NRR produced a normalized random Z-score for Moran's I of 5.50, again indicating a statistically significant level of spatial autocorrelation.

The $G_i^*(d)$ statistic (see Getis and Ord 1992; Getis 1995; Getis and Ord 1996) provides a more precise local test for spatial dependence. For any given cell in the grid, its statistically significant difference from spatial independence is given by the ratio of the $G_i^*(d)$ statistic relative to its expected value at a calculated critical distance d . We calculated the critical

distance as that distance at which a spatial filtering algorithm had removed the spatial autocorrelation (measured by Moran's I) from the variable (Scott 1999).

Any cell with a $G^*_i(d)$ value that is statistically significant at the .05 level would cause us to reject the null hypothesis of no clustering and would assign that cell to a cluster of either high or low fertility, depending upon the sign of G^* . In 1976 the critical distance was 5km, indicating that on average villages that were clustered were most similar in fertility levels to those that were within a 5km radius. In 1986 the critical distance was 4km. In both 1976 and 1986 the villages clustered around low fertility also exhibited, as expected, lower than average proportions married, lower proportions of illiterate women, and lower adult sex ratios. In 1986, but not in 1976, the lower fertility clusters were associated with more populous villages (which we used as a proxy for level of urbanization). On each characteristic the villages clustered around high fertility exhibited the opposite patterns--higher proportions married, higher proportions of illiterate women, and higher adult sex ratios, again consistent with our theoretical expectations.

Panel A of Figure 2 shows the spatial clustering of high and low fertility in 1976, where villages in high clusters were those whose normalized z-scores for the $G^*_i(d)$ statistic were above 2, villages in low cluster were those whose normalized z-scores for the $G^*_i(d)$ statistic were below -2, and those villages with normalized z-scores between -2 and 2 were considered not to be clustered. A kriging function was applied to smooth the data for purposes of enhanced visualization of the results. In 1976 the clusters of low fertility were found in the north and northeast, while the clusters of high fertility were concentrated in the south and southwest of the governorate.

FIGURE 2 ABOUT HERE

Although the pattern of clustering shown in Figure 2 could be interpreted as being influenced by edge effects, the clustering in 1986, shown in Panel B of Figure 2, seems to belie that explanation. In 1986 the clustering of low fertility had moved toward the middle of the governorate, although still concentrated in the north, whereas the clustering of high fertility was more concentrated in the southern portion of the governorate. Although the southern portion of Menoufia is closest to Cairo, it is also the site of the Barrage--the dam that controls water from the Nile as it enters the Delta region. This is rich agricultural land with centuries, if not millennia, of rural tradition that almost certainly contributes to the maintenance of low levels of education, low levels of female labor force participation, and higher than average levels of fertility.

The data in Panel C of Figure 2 illustrate the change in the pattern of clustering between 1976 and 1986. Villages were categorized according to the combinations of clustering in the two time periods. Thus, the lightest shades of clustering are assigned to villages that were in low fertility clusters in both 1976 and 1986, while the next lighter shade indicates villages that moved from not being clustered in 1976 to being a in low fertility in 1986. The data thus show the concentration of lower fertility in the north, and the diffusion of lower fertility in that region. At the other extreme, the darkest shading is assigned to villages that were in high fertility clusters in both 1976 and 1986 and the next darker shading reflects villages that went from not being clustered in 1976 to being in a high fertility cluster in 1986. These villages are concentrated in the southern portion of the region. In general, the changes between 1976 and 1986 exhibit a spatial diffusion effect, with a spread of higher than average fertility to contiguous villages, and a spread of lower than average fertility to contiguous villages.

It is clear from Figure 2 that spatial variability in fertility exists in Menoufia. It does matter where you are--lower fertility is clustered in the north, and higher fertility is clustered in the south. How important is this spatial effect as a determinant of fertility levels? We used the technique of spatial filtering of variables in a regression model to try to answer this question. First we developed an OLS regression model that did not include a spatial component, echoing the typical such model in demographic analysis. Then, we filtered the statistically significant predictor variables to assess the importance of the spatial effect.

The basic non-spatial model is that the fertility level in a village is a function of female illiteracy, controlling for the sex ratio at the reproductive ages (as a control for the effect of out-migration), the percentage of adult women who are currently married (as a control for the effect of marital status on the measure of fertility that we calculated) and for total population size (as a control for urban-ness). Two of the predictor variables--female illiteracy and proportion married--have statistically significant levels of spatial autocorrelation, whereas the other two predictor variables do not. The two with spatial autocorrelation were then filtered to decompose the spatial component from the non-spatial (called the "filtered") component, using the method described in Getis (1995). The regression results produce an overall adjusted R^2 of 0.393, as summarized in Table 1. The filtered component of the proportion married has a standardized beta coefficient that is virtually the same as the spatial component of that variable, indicating that the spatial component explains about half of that variable's relationship to fertility levels. The spatial component of female illiteracy is slightly more important than the filtered component, as can be seen in Table 1. Summing the absolute values of the statistically significant standardized beta coefficients, we can determine that the spatial component accounts for 33 percent of the explained variation in the net reproduction rate in 1976.

TABLE 1 ABOUT HERE

In 1986 the female illiteracy variable was a more important predictor of fertility than was the percent married, and neither the adult sex ratio nor the total population size was statistically significantly related. The spatial component was also more noticeable than in 1976 because all four predictor variables exhibited spatial autocorrelation and the residuals were also spatially autocorrelated. All four predictor variables were filtered in 1986 and the resulting regression model is shown in the lower panel of Table 1. Filtering raised the explained variation from .482 in the original model (not shown) to .513 in the spatial model, although female illiteracy and the proportion of women married remained as the only statistically significant variables in the equation. In 1986 the most important predictor was the filtered (non-spatial) component of female illiteracy, followed by the spatial component of the proportion married, then the filtered component of the proportion married, and the spatial component of female illiteracy. If we once again sum the standardized beta coefficients of the statistically significant variables, we find that the spatial component accounts for 42 percent of the explained variation--a higher fraction than in 1976.

The period from 1976 to 1986 was a period of overall relative stability in fertility levels in rural Egypt, yet it is obvious that at least by 1976 there were clear spatial patterns to fertility in Menoufia, and our analysis suggests that these spatial patterns were even more definitive in 1986 than they had been in 1976, with the southern portion of the governorate being more obviously the location of higher than average fertility in 1986 than had been true in 1976.

There is not space in this chapter, obviously, to detail all of the nuances of this type of analysis, but the important point to be made here is that there is a clearly established spatial component to fertility levels and fertility change in rural Egypt. From a theoretical perspective,

the spatial analysis helps us to quantify the roles that human capital and diffusion factors may be playing in the fertility transition in rural Egypt. From a research perspective, the spatial analysis helps to identify places where things are clearly different and where additional research ought to be focused. From a policy perspective, the spatial analysis helps planners and providers to know where programs of reproductive health and rural redevelopment are likely to have the greatest impact on fertility change.

Conclusion

Demographic research is moving inexorably from its long-standing pattern of spatial awareness to an increased appreciation for the value and utility of spatial analysis. In this chapter I have emphasized the role that spatial analysis can play in the testing of propositions that are central to demographic theory. I would be remiss to not also mention the importance of spatial analysis in models that link population growth and distribution to global issues such as land, water, and atmospheric degradation and change. The resources consumed locally are increasingly derived from non-local sources; and the polluting side-effects of resource consumption may occur locally, but their impact can spread well beyond that. Understanding these and other kinds of global population-environment interactions requires the application of large-scale GIS models, following the lead of organizations such as the International Institute for Applied Systems Analysis in Austria. Thus far, demographers have been only minimally involved in this kind of research (see Findlay and Hoy 2000), and the modeling that has been done has not involved intensive use of spatial analysis *per se*. This combination seems to signal an area ripe for potential growth in the level of sophistication with which the consequences of population growth and change can be researched.

Figure 1. A Framework for Spatial Analysis in Demographic Research

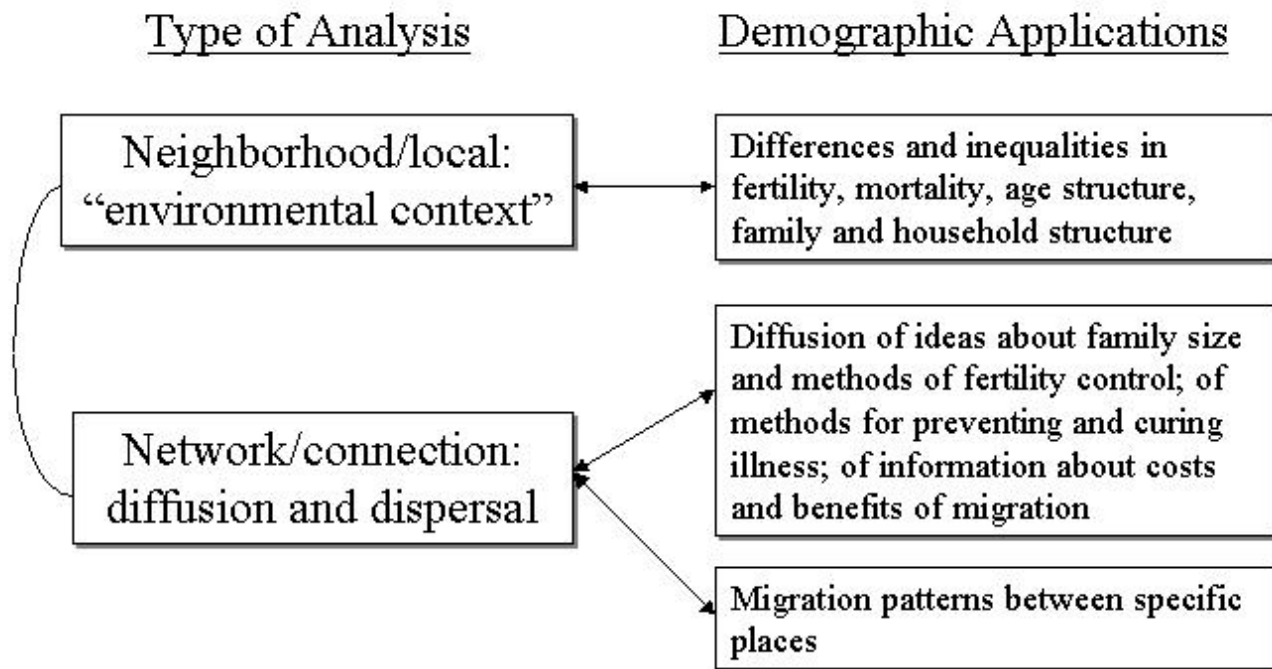
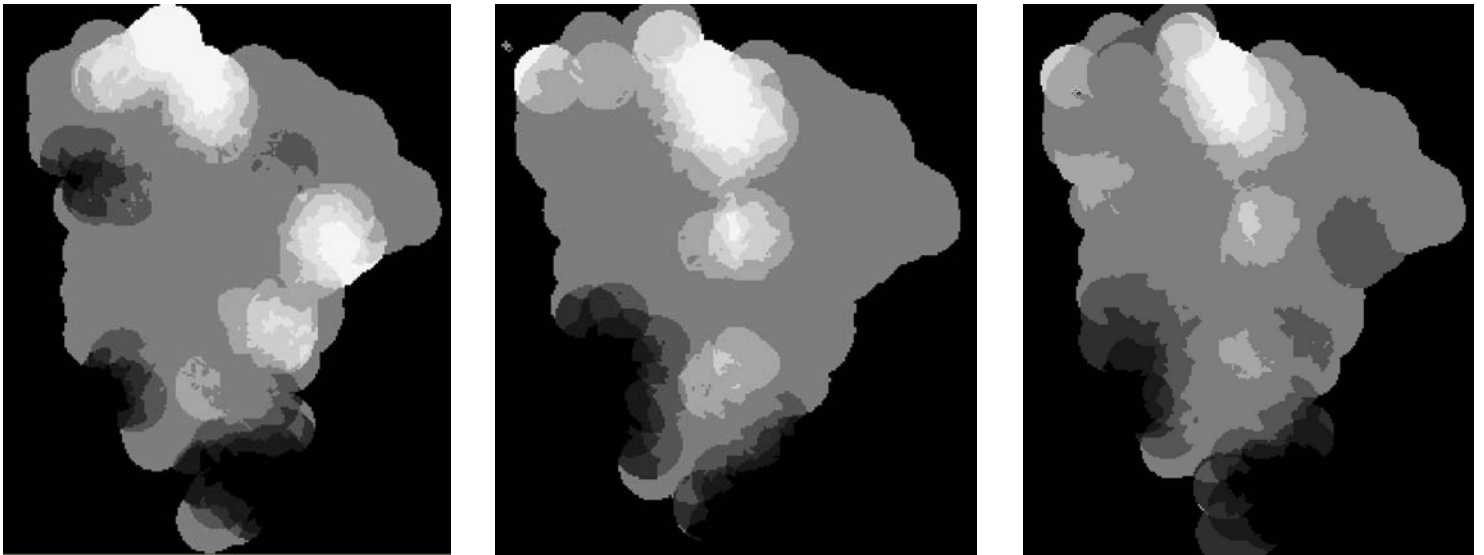


Figure 2. Spatial Clustering of Fertility in Menoufia

A. 1976

B. 1986

C. Change 1976-86



Note: lighter shading indicates clusters around low fertility, whereas darker shading indicates clustering around high fertility. Clustering is based on G_i^ scores.*

Source: John R. Weeks, M. Saad Gadalla, Tarek Rashed, James Stanforth, and Allan G. Hill, 2000. "Spatial Variability in Fertility in Menoufia, Egypt, Assessed Through the Application of Remote Sensing and GIS Technologies," *Environment and Planning A*, (32):695-714: Figures 5, 6, and 7.

Table 1. Spatially Filtered OLS Regression Results

Variable	1976					
	Unstandard- ized coefficient	Standardized beta	t	Significance of t	Z (I)	
Dependent variable: NRR						6.82
Filtered Female illiteracy	.417	.149	2.699	.007	-0.99	
Filtered Proportion married	1.724	.329	5.907	.000	0.48	
Spatial female illiteracy	.640	.167	2.563	.011	10.92	
Spatial Proportion married	2.322	.331	5.024	.000	11.76	
Sex ratio at reproductive ages	.178	.060	1.223	.222	1.30	
Population size	-.000005	-.148	-3.111	.002	0.96	
R	.637					
Adjusted R ²	.393					
Z(I) for residuals	0.67					
Variable	1986					
	Unstandard- ized coefficient	Standardized beta	t	Significance of t	Z (I)	
Dependent variable: NRR						5.50
Filtered Female illiteracy	1.920	.480	9.496	.000	-.64	
Filtered Proportion married	1.057	.170	3.587	.000	-.36	
Filtered sex ratio at reproductive ages	.002	.006	0.136	.892	-.58	
Filtered population size	-.00000008	-.063	-.379	.705	.03	
Spatial female illiteracy	.913	.157	2.716	.007	26.80	
Spatial Proportion married	3.352	.317	4.755	.000	25.72	
Spatial sex ratio at reproductive ages	.640	.051	0.332	.332	28.96	
Spatial population size	-.00000004	-.021	-.131	.896	.22	
R	.717					
Adjusted R ²	.513					
Z(I) for residuals	1.32					

Source: John R. Weeks, M. Saad Gadalla, Tarek Rashed, James Stanforth, and Allan G. Hill, 2000. "Spatial Variability in Fertility in Menoufia, Egypt, Assessed Through the Application of Remote Sensing and GIS Technologies," *Environment and Planning A*, (32):695-714: Tables 5 and 7.

References

- Arriaga, Eduardo E. 1994. *Population Analysis With Microcomputers*. Washington, D.C.: U.S. Bureau of the Census.
- Bocquet-Appel, Jean-Pierre and Lucienne Jakobi. 1998. "Evidence for a spatial diffusion of contraception at the onset of the fertility transition in victorian Britain." *Population: An English Selection* 10:181-204.
- Brown, Lesley. 1993. "The New Shorter Oxford English Dictionary on Historical Principles.". Oxford: Clarendon Press.
- Burt, Ronald S. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge: Harvard University Press.
- . 1999. "The Social Capital of Opinion Leaders." *Annals of the American Academy of Political and Social Sciences* 566:37-54.
- Chou, Yue-Hong. 1997. *Exploring Spatial Analysis in Geographic Information Systems*. Santa Fe, NM: OnWard Press.
- Coale, Ansley and Susan Cotts Watkins. 1986. "The Decline of Fertility in Europe.". Princeton, New Jersey: Princeton University Press.
- Courbage, Yousef. 1994. "Demographic Change in the Arab World: The Impact of Migration, Education, and Taxes in Egypt and Morocco." *Middle East Report* 24:19-22.
- Crenshaw, Edward, Matthew Christenson, and Doyle Ray Oakey. 2000. "Demographic transition in ecological focus." *American Sociological Review* 65:371-391.
- Cressie, Noel A. C. 1993. *Statistics for Spatial Data: Revised Edition*. New York: John Wiley & Sons.

- Duncan, Otis Dudley. 1959. "Human Ecology and Population Studies." in *The Study of Population: An Inventory and Appraisal*, edited by P. M. Hauser and O. D. Duncan. Chicago: University of Chicago Press.
- Entwisle, Barbara, John Casterline, and Hussein A.-A. Sayed. 1989. "Villages as contexts for contraceptive behavior in rural Egypt." *American Sociological Review* 54:1019-1034.
- Entwisle, Barbara, Ronald R. Rindfuss, Stephen J. Walsh, Tom P. Evans, and Sara R. Curran. 1997. "Geographic information systems, spatial network analysis, and contraceptive choice." *Demography* 34:171-188.
- Fargues, Philippe. 1997. "State Policies and the Birth Rate in Egypt: From Socialism to Liberalism." *Population and Development Review* 23:115-138.
- Findlay, Allan M. and Caroline Hoy. 2000. "Global Population Issues: Towards a Geographical Research Agenda." *Applied Geography* 20:207-219.
- Gadalla, Saad. 1978. *Is There Hope? Fertility and Family Planning in a Rural Community in Egypt*. Chapel Hill, NC: Carolina Population Center, University of North Carolina.
- Gadalla, Saad, James McCarthy, and Neeraj Kak. 1987. "The determinants of fertility in rural Egypt: A study of Menoufia and Beni-Suef." *Journal of Biosocial Science* 19:195-207.
- Getis, Arthur. 1995. "Spatial Filtering in a Regression Framework: Examples Using Data on Urban Crime, Regional Inequality, and Government Expenditures." in *New Directions in Spatial Econometrics*, edited by L. Anselin and R. Florax. Berlin: Spriger-Verlag.
- Getis, Arthur and J. Keith Ord. 1996. "Local Spatial Statistics: an Overview." in *Spatial Analysis: Modelling in a GIS Environment*, edited by P. Longley and M. Batty. Cambridge, England: GeoInformation International.

- Getis, Arthur and J.K. Ord. 1992. "The Analysis of Spatial Association by Use of Distance Statistics." *Geographical Analysis* 24:189-206.
- Hill, Allan G. 1997. "'Truth lies in the eye of the beholder': The nature of evidence in demography and anthropology." Pp. 223-247 in *Anthropological Demography: Toward a New Synthesis*, edited by D. Kertzer and T. Fricke. Chicago: University of Chicago Press.
- Kohler, Hans-Peter, Jere R. Behrman, and Susan C. Watkins. 2001. "The Density of Social Networks and Fertility Decisions: Evidence from South Nyanza District, Kenya." *Demography* 38:43-58.
- Langford, Ian H., Alastair H. Leyland, Jon Rasbash, and Harvey Goldstein. 1999. "Multilevel Modelling of the Geographical Distributions of Diseases." *Journal of the Royal Statistical Society, Series C: Applied Statistics* 48:253-68.
- Langford, M. and D.J. Unwin. 1994. "Generating and mapping population density surfaces within a geographical information system." *Cartographic Journal* 31:21-25.
- Leasure, J. William. 1962. "Factors Involved in the Decline of Fertility in Spain: 1900-1950." Doctoral Thesis, Economics, Princeton University, Princeton.
- LeClere, Felicia B., Richard G. Rogers, and Kimberly Peters. 1998. "Neighborhood Social Context and Racial Differences in Women's Heart Disease Mortality." *Journal of Health and Social Behavior* 39:91-107.
- Lo, C.P. 1995. "Automated population and dwelling unit estimation from high-resolution satellite images: a GIS approach." *International Journal of Remote Sensing* 16:17-34.
- Mesev, Victor, Paul Longley, and Michael Batty. 1996. "RS-GIS: Spatial Distributions from Remote Imagery." in *Spatial Analysis: Modelling in a GIS Environment*, edited by P. Longley and M. Batty. Cambridge, UK: Geoinformation International.

- Micklin, Michael and David F. Sly. 1998. "The Ecological Complex: A Conceptual Elaboration." in *Continuities in Human Ecology*, edited by M. Micklin and D. L. Poston. New York: Plenum Press.
- Montgomery, Mark R. and John B. Casterline. 1993. "The diffusion of fertility control in Taiwan: evidence from pooled cross-section time-series models." *Population Studies* 47:457-479.
- Namboodiri, Krishnan. 1988. "Ecological demography: its place in sociology." *American Sociological Review* 53:619-633.
- Noin, Daniel. 1970. *La Population Rurale du Maroc*. Paris: Presses Universitaires de France.
- Openshaw, S. 1984. *The modifiable area unit problem, CATMOG 38*. Norwich: Geoabstracts.
- Openshaw, S. and L. Rao. 1995. "Algorithms for reengineering 1991 Census geography." *Environment and Planning A* 27:425-446.
- Peak, Christopher and John R. Weeks. Year. "Does Community Context Influence Reproductive Outcomes of Mexican Origin Women in San Diego, California?" Presented at Annual Meeting of the Population Association of America, Washington, DC.
- Reid, Alice. 1997. "Locality or Class? Spatial and Social Differences in Infant and Child Mortality in England and Wales, 1895-1911." in *The Decline of Infant and Child Mortality. The European Experience: 1750-1990*, edited by C. A. Corsini and P. P. Viazzo. Dordrecht, Netherlands: Martinus Nijhoff.
- Ridd, M. 1995. "Exploring a V-I-S (Vegetation-Imperious Surface-Soil) Model or Urban Ecosystem Analysis Through Remote Sensing: Comparative Anatomy of Cities." *International Journal of Remote Sensing* 16:2165-2185.

- Rindfuss, Ronald R. and Paul C. Stern. 1998. "Linking Remote Sensing and Social Science: The Need and the Challenges." in *People and Pixels: Linking Remote Sensing and Social Science*, edited by D. Liverman, E. F. Moran, R. R. Rindfuss, and P. C. Stern. Washington, DC: National Academy Press.
- Robert, Stephanie A. 1998. "Community-Level Socioeconomic Status Effects on Adult Health." *Journal of Health and Social Behavior* 39:18-37.
- Rosero-Bixby, Luis and John B. Casterline. 1994. "Interaction diffusion and fertility transition in Costa Rica." *Social Forces* 73:435-62.
- Star, Jeffrey and John Estes. 1990. *Geographic Information Systems: An Introduction*. Englewood Cliffs, NJ: Prentice Hall.
- Sutton, Paul. 1997. "Modeling population density with night-time satellite imagery and GIS." *Computing, Environment and Urban Systems* 21:227-244.
- Tickamyer, Ann R. 2000. "Space Matters! Spatial Inequality in Future Sociology." *Contemporary Sociology* 29:805-812.
- Tobler, Waldo. 1970. "A Computer Movie Simulating Urban Growth in the Detroit Region." *Economic Geography* 26:234-40.
- Tolnay, Stewart E. 1995. "The spatial diffusion of fertility: a cross-sectional analysis of counties in the American South, 1940." *American Sociological Review* 60:299-308.
- van de Walle, Etienne. 1965. "An approach to the study of fertility in Nigeria." *Population Studies* 19:5-16.
- Wallace, R. and M.T. Fullilove. 1991. "AIDS Deaths in the Bronx 1983-1988: Spatiotemporal Analysis From a Sociodemographic Perspective." *Environment and Planning A* 23:701-23.

- Ward, D., S.R. Phinn, and A.T. Murray. 2000. "Monitoring Growth in Rapidly Urbanizing Areas Using Remotely Sensed Data." *The Professional Geographer* 52:371-385.
- Watkins, Susan Cotts. 1991. *From Provinces Into Nations: Demographic Integration in Western Europe, 1870-1960*. Princeton, New Jersey: Princeton University Press.
- Weeks, John, M. Saad Gadalla, and Allan G. Hill. Year. "The Environmental Context of Fertility in Egypt: Evidence from Demographic and Health Surveys." Presented at Annual Meeting of the Association of American Geographers, New York.
- Weeks, John R., M. Saad Gadalla, Tarek Rashed, James Stanforth, and Allan G. Hill. 2000. "Spatial Variability in Fertility in Menoufia, Egypt, Assessed Through the Application of Remote Sensing and GIS Technologies." *Environment and Planning A* 32:695-714.
- Wilson, Chris and Robert Woods. 1991. "Fertility in England: a long-term perspective." *Population Studies* 45:399-415.