Center for Demography and Ecology

University of Wisconsin-Madison

Growth Curve Modeling: An Alternative Approach to Estimating Patterns of Population Change

Katherine J. Curtis White

Jerald R. Herting

CDE Working Paper No. 2004-18

Growth Curve Modeling:
An Alternative Approach to Estimating Patterns of Population Change

Katherine J. Curtis White*
Center for Demography and Ecology
University of Wisconsin

Jerald R. Herting
Department of Sociology
University of Washington

July 2004

Word Count: 8,866

* Please direct all correspondence to Katherine J. Curtis White, Center for Demography and Ecology, 4412 Social Science Building, 1180 Observatory Drive, University of Wisconsin, Madison, WI 53706-1393, or kwhite@ssc.wisc.edu. This work was supported by the Shanahan Fellowship through Center for Studies in Demography and Ecology, University of Washington, and by Center Grant P30 HD05876 and Training Grant T32 HD07014 from the National Institute for Child Health and Human Development, National Institutes of Health through the Center for Demography and Ecology, University of Wisconsin.

ABSTRACT

Working from the formal model of population growth, $P_t = P_0e^{rt}$, we use population change in the U.S. Great Plains between 1900 and 1930 to demonstrate the efficiency of an alternative approach to estimating population change, r: growth curve modeling (GCM). Through this method, the researcher treats processes of change as a trajectory rather than a series of panels, resulting in a more efficacious and parsimonious analysis. Three general aspects of population growth can be addressed through linear and non-linear GCM techniques, namely (1) establishing a mean growth rate that represents the general pattern of change across all units, (2) estimating the degree of local level variation around the average rate, and (3) analyzing the relative contribution of factors influencing this variation. Our findings reveal that the non-linear cubic model most appropriately suits the early years of development on the Great Plains, indicating that the rate of population change varied over the period with rapid growth at the outset followed by a decreasing rate of growth in the middle decade and an accelerating rate during the later portion. In addition, we observed considerable variation around the average growth rate, and attribute much of this to county settlement date, urban proximity, and economic base, each to a varying extent across the different models.

INTRODUCTION

Much of traditional demographic work concerns estimating population change over periods of time. Contemporary demographers have taken their analyses a step further, often placing this estimation procedure within a causal framework to identify the factors influencing change. Working from the formal discussion of population growth, $P_t = P_0e^{rt}$ (see Shryock, Siegel, Stockwell (1976)), we offer an efficient and accessible method of analyzing population change over multiple time points that allows for variation in growth among entities (e.g., countries, counties, cities, etc.). We propose the adaptation and use of growth curve modeling (GCM). More specifically we turn the typical growth model into $P_{it} = P_0e^{r_it}$ or, in its additive form, $lnP_{it} = lnP_{i0} + r_it$, (where lnP_{it} is the natural log of the population at any time point t for each case, lnP_{i0} is the natural log of the initial population, r_i is the growth rate for location i, and t is observed years, or some other time interval). This formulation focuses on how growth varies across units and what characteristics influence variation in growth.

This paper is motivated by problems with other, perhaps more traditional techniques applied to research questions concerning change over multiple observation points. The more "traditional" approaches we refer to are panel-like models that typically focus on change between two time points, P_0 and P_t. The time points can include consecutive or interrupted observations such as population change between 1920 and 1930, or 1900 and 1930. In both instances, r_i does in fact represent the rate of change between P_0 and P_t. However, there are some limitations to this approach. First, the panel technique may oversimplify the actual process by considering r_i between the beginning and end points such as 1900 and 1930, while omitting rates of change between 1900 and 1910, 1910 and 1920, and 1920 and 1930, thus often failing to capture or characterize the information in a full time series. Second, it is redundant for questions concerning multiple time points such that one must estimate r_i for change between each time unit generally as some form of a difference (e.g., simple linear change). In attempting to capture the full series, proliferate parameters result through evaluation of numerous time segments (e.g., 1900-1910, then 1910-1920) and the general long-term trend remains unaddressed in favor of these individual segments.

In this typical approach, a global or average estimate is calculated to represent population change over a time period which includes multiple points. The rate of change is often considered linear and constant, in that each decade witnesses the same degree of change as adjacent decades. Such an assumption is not entirely adequate, as rarely will change occur at an identical rate across units and in a linear fashion over periods of time (Shryock et al. 1976). Studies of shorter time frames are typically not in egregious violation of this assumption. However, the utility of such an approach and the validity of its assumptions become increasingly unsuitable and unreliable as the time period of interest increases. Moreover, to capture change over a larger number of periods, the panel-like approach tends to provide a rather overwhelming number of parameters when estimating the influence of various factors on rates of change between each period.

We suggest that growth curve modeling techniques offer an alternative that fully considers rather than masks variation in patterns of population change over units and easily spans multiple observations on those units. In the following pages, we describe the basic principles behind the method and demonstrate the technique using county-level population change in the U.S. Great Plains over four time points between 1900 and 1930. In so doing, we illustrate an alternative approach to assessing population growth and estimating the factors driving such change.

THE GROWTH CURVE MODELING APPROACH

The growth curve analysis allows for estimation of (1) an average growth rate, (2) variation around the average rate, and (3) the relative contribution of initial conditions or levels, or change in other factors thought to influence this variation. Conceptually, growth curve modeling techniques allow us to consider population change as a trajectory rather than a simple transition between two time points.

In essence, a growth curve can be estimated to characterize the general trend in population change while considering population at each time point within the observed interval. This method is rooted in developmental psychological and educational research (see Bryk and Raudenbush (1992), Raudenbush and Bryk (2002) or McArdle and Epstein (1988)) and is often applied to individuals, where behaviors or characteristics are traced over the life-course (e.g., Huttenlocher et al. 1991). In the population context, the county or some other spatial unit easily replaces the individual as the unit of measure, and population growth replaces the individual behavior.

Through this approach, the data may be considered multilevel in that time is nested within county, thereby allowing for the use of multilevel modeling techniques.[1] Hox and Kreft (1994) provide a useful discussion of various multilevel models and their limitations, and software enabling the analysis of such models (see also Bryk and Raudenbush (1992), Raudenbush and Bryk (2002)). The multilevel approach in hierarchical linear modeling (HLM) addresses two

---

[1] Another variation is to apply latent model techniques to growth as demonstrated by McArdle and Epstein (1987) or Duncan et al. (1999); these latent models can provide some additional flexibility. In addition, latent mixture models that are currently being introduced in the deviance and criminology literature, where one estimates latent classes of growth, may also be employed (see Nagin and Trembly 1999). We use the multilevel format and HLM for ease of presentation and the greater familiarity of these models to sociologists, demographers, and ecologists who are acquainted with contextual models. Typically, models will produce the same or similar results regardless of the approach.

potential problems with the data, dependency and hierarchical nesting, and also allow us to consider modeling growth over multiple periods.

Hox and Kreft also offer a useful summary of the problems associated with using single level models for multilevel data. Only one stochastic error component is allowed in a single level model, and is assumed independent, normal, and homoscedastic. If the data are multilevel, the unmodeled group variation will produce a residual error term at the group level, assuming a nonzero covariance between the residual error terms of individuals comprising the group. Otherwise stated, it assumes independence between the observations. But because the data contain multiple observations of the same county, or some other location of interest, the data are actually dependent. Dependency, and the accompanying error structure, is not accounted for in a single level model. This violates the assumption upon which standard errors and variance are determined and, therefore, produces unreliable estimates of standard errors and variance. Such single level model results are unreliable for multilevel data structures.

The Average Growth Rate

Given the problems in using a single level model to estimate population change as a trajectory, we treat the data as multilevel. There are two levels and, therefore, two equations that are estimated through this method. The first level model is sometimes referred to as the repeated observation model while the second level model is often known as the personal- or individual-level model. In the current analysis, level 1 contains four repeated observations of counties, one for each census year between 1900 and 1930, and level 2 is the county-level model.

We formulate the log of the general growth model, level 1 is denoted as

$$P_{ti} = \pi_{0i} + \pi_{1i} T_t + \varepsilon_{ti} \qquad [1]$$

where π_{0i} indicates the initial population of county i when time is zero or 1900, π_{1i} represents the rate of growth in the population over the period for county i, T_t is the value of time, and ε_{ti} is the error for county i. Here, the effect of time is represented as a linear function.

The accompanying level 2 model is denoted as

$$\pi_{0i} = \beta_{00} + \Sigma\, \beta_{0q}\, X_{qi} + \rho_{0i} \qquad [2a]$$

$$\pi_{1i} = \beta_{10} + \Sigma\, \beta_{1q}\, X_{qi} + \rho_{1i} \qquad [2b].$$

This simply describes the county-level population change by considering both the initial population (π_{0i}) and the growth rate (π_{1i}) as varying over units (e.g., counties). This variation is a function of the characteristics of each county (the X_{qi}'s). In effect, we are both adding direct or main effects of county characteristics to explain population variation and interacting county measures with time, which generates the variation observed in the rate of growth.[2]

We can easily extend this to a non-linear context, as the equations are only slightly different. When using a quadratic form, the level 1 model is denoted as

$$Y_{ti} = \pi_{0i} + \pi_{1i}\, T_{ti} + \pi_{2i}\, T_{ti}^2 + \varepsilon_{ti} \qquad [3]$$

where $\pi_{2i}\, T_{ti}^2$ can be seen as representing the expected change in population growth in terms of deceleration or acceleration; it is the rate of change in the rate of change.

The second level model includes as many equations as necessary to capture the appropriate time function, such as quadratic or cubic. Other representations of the effect of time are also possible, for example a spline function. In the quadratic instance, the level 2 model is denoted as

$$\pi_{0i} = \beta_{00} + \Sigma\, \beta_{0q}\, X_{qi} + \rho_{0i} \qquad [4a]$$

$$\pi_{1i} = \beta_{10} + \Sigma\, \beta_{1q}\, X_{qi} + \rho_{1i} \qquad [4b]$$

$$\pi_{2i} = \beta_{20} + \Sigma\, \beta_{2q}\, X_{qi} + \rho_{2i} \qquad [4c]$$

[2] The equations show the same "explanatory" variables in each equation, although this need not be the case as variable inclusion is dependent upon the research question at hand.

where equation 4c reflects the effect of other characteristics on acceleration or deceleration depending on the estimated coefficient sign.

It is worth stressing that in the context of population change, we are typically less concerned with explaining variation in the initial *population* (π_{0i}). Instead, we are interested in explaining variation in the rate of *population growth* (π_{1i}). Yet both components, the intercept and time coefficients, are used to calculate the average trajectory such that the linear model produces estimates for π_{0i} and π_{1i} through equations 2a and 2b respectively. In the non-linear quadratic context, estimates for π_{0i}, π_{1i}, and π_{2i} are produced through equations 4a through 4c. These values are then substituted into equation 1 (or 3), and the value of time (T_{ti}) is varied to construct a population *trajectory*. The dependent variable is population and equation 1 reflects the predicted population. However, the growth rate is estimated in equation 2b, and included in the solution of equation 1 to produce a population trajectory that reflects the variation in growth rate over the observed interval across units.

Variation Around the Average Growth Rate

After establishing an average growth rate and the accompanying average population trajectory, researchers might also be interested in estimating the degree of local level variation around this average trend. Here, the researcher is concerned with whether and to what degree there is variation in population change. Focus is devoted to variation around the time coefficient (e.g., the variance of ρ_{1i} in equation 2b) as this reflects the extent of local level variation in the growth rate. In practice, one asks whether a baseline model in which there is only an average growth parameter and error presents significant variance. If so, then explanatory variables, the X_{qi}, may be added to explain the observed variation in growth.

The Contribution of Other Factors

Finally, growth curve modeling techniques also permit the researcher to estimate the impact of various factors on population change. In our substantive example we use initial conditions (i.e. characteristics at T_0) to estimate which factors at the beginning stage drive growth. There is no restriction to initial conditions and one can just as easily use change in other conditions to model growth.

Again, the focus remains on the growth rate rather than the initial population. Therefore, county characteristics ($\Sigma\ \beta_{1q}\ X_{1qi}$) are introduced in equation 2b (or 4b and $\Sigma\ \beta_{2q}\ X_{2qi}$ in 4c). The contribution to the variation around the average growth rate is noted by comparing the variance components of the unconditional model to those of the model conditional on other county characteristics. The relative influence of the county characteristics is revealed through coefficient comparisons and by solving equation 1 using varying values of particular county characteristics while holding all others at their mean value. The approach is similar to calculating a predicted probability in the logistic regression context, where population estimates are derived through a two-step process solving equations 1 and 2b (or 3 and 4b and 4c).

First, the mean effect of time or the mean growth rate (π_{1i}) is calculated by adding the coefficients for the intercept and the independent variables while holding the value of the independent variables at their mean ($\pi_{1i} = \beta_{10} + \Sigma\ \beta_{1q}\ X_{1qi} + r_{1i}$, where β_{10} is the intercept and $\Sigma\ \beta_{1q}\ X_{1qi}$ is the sum of the independent variables (β_{1q}) at their mean values (X_{1qi})). Second, the calculated mean growth rate (π_{1i}) is then used to estimate the population for each time point ($Y_{ti} = \pi_{0i} + \pi_{1i}\ T_{ti} + e_{ti}$, where π_{0i} is the mean initial status, and T_t is the time point such that 1900 is 0, 1910 is 1, and so forth).

AN ILLUSTRATION: COUNTY-LEVEL POPULATION CHANGE IN THE U.S. GREAT PLAINS

We have selected population change in the U.S. Great Plains between 1900 and 1930 to illustrate the utility of the growth curve model strategy. The region underwent considerable population change during this period, characteristic of early modernization and burgeoning industrialization. Several factors suspected of motivating population change are applied to this period of growth in the Great Plains: (1) environmental conditions, (2) economic base, (3) population characteristics, and (4) spatial factors. Using GCM techniques, we estimate the average growth rate, county-level variation around the average, and the relative contribution of the forces underlying the process of change.

The Suspected Driving Forces

In general, little sociological attention has been devoted to the influence of *environmental conditions* on population. These factors are typically viewed as secondary, control factors. However, contemporary demographic work is marked with a re-emerging interest in the relationship between population and the environment. Perspectives oriented toward the natural- versus the built-environment argue that places with greater resources will experience greater growth since people move to the most desirable places. For example, research focusing on recent migration in the Great Plains has found an increase in population growth in counties possessing natural amenities such as warm climates and varied topography (Cromartie 1998:27). For the U.S. in general, research has shown that migration among the young and elderly is motivated by non-economic factors such as tepid climates and the presence of water, both of which might be considered indicators of natural amenities (Heaton et al. 1981). While older populations may be attracted to these rural, high-amenity locations for quality of life aspects, the young are arguably drawn by the economic opportunities accompanying expanding recreation and retirement populations.

Yet, the influence of the natural environment is not limited to contemporary retirees and vacationing populations flocking to moderate climates. In fact, much of the historical work on the Great Plains regions accentuates the importance of precipitation, especially during the settlement era (Stegner 1954; Gregory 1989). Stegner states, "Water is the true wealth in a dry land; without it, land is worthless or nearly so" (1954:226). Precipitation is likely a key factor for the economic well being of this agriculturally dependent region, where greater precipitation is associated with greater agricultural productivity and, therefore, growing or stable populations.

A second body of research identifies *economic base* as a major contributor to population change. According to this perspective, people go where economic opportunities exist. This argument has been applied to the individual (Sjaastad 1962) as well as aggregate populations (Greenwood 1981). At the individual level, migration is viewed as an investment in human capital, where people migrate from places with relatively low wage structures and little economic opportunity to places with higher wages and greater economic opportunity. A push-pull framework is implicit in this approach where certain factors push people out of their place of origin while other factors pull them to their place of destination. A similar logic is applied to the population level. Within this framework, the economic base serves as both the push and pull factor (Greenwood 1981:3). In the Great Plains context, we might expect that the economic base is associated with population change, such that counties largely dependent on the agricultural industry are more likely to experience population growth during the early part of the century when the industry was expanding. Similarly, counties with high manufacturing employment are also anticipated to experience growth during this early period of settlement.

Proponents of the *spatial perspective*, reflected in the metropolitan dominance theory (Hawley 1950), the central place model (Christaller 1966), and the size-function hierarchy

(Duncan et al. 1960), argue that proximity and/or accessibility to a metropolitan center drive population growth. Here, settlement systems can be viewed as territorial divisions of labor with units of systems integrated through social and economic functions (Johansen and Fuguitt 1984). Social and economic innovations and the accompanying growth responses flow down the hierarchy, where cities are at the top and settlements dispersed throughout the open country are located at the bottom.

Researchers highlight the importance of transportation in this perspective. In summary, transportation influences the location of the economic base, which in turn, affects the development of the urban hierarchy (McKenzie 1929; Stephan 1971). Transportation, according to Durkheim (1933), produces divisions in labor and territorial organization. Further, transportation allows businesses and residents to select alternative locations (McKenzie 1929).

Although this perspective recognizes the influence of industry or economic base, it emphasizes spatial patterns and their impact on population distribution. For example, in *Rural and Small Town America*, Fuguitt et al. (1989) state that the changing geographic location of economic activities in recent periods has had important impacts on population growth and distribution of rural areas within the U.S. Namely, the decentralization of the manufacturing and service industries has permitted smaller places to compete more effectively with their larger, urban counterparts (1989:426).

Researchers also suggest that the importance of spatial proximity and accessibility appears to vary over time (Johansen and Fuguitt 1984). The spatial perspective applies most appropriately when central place manufacturing and primary production are the predominant industries, and least appropriately in periods and places where communication and transportation expand and dominate. In essence, technological advancements made during the later part of the century have detached spatial distances from communication distances. This implies that the metropolitan dominance perspective does not apply to places in later stages of positive urban development where communication and transportation abound. However, the metropolitan dominance framework may be valid for places or periods where agriculture persists. Importantly, such places exist within the Great Plains, especially during the settlement years.

Regarding *population characteristics*,[3] one of the central elements of migration is its selective nature; certain people are more likely to migrate than others, and those who migrate are somehow different from those not migrating. Selection criteria range from educational and economic status to life cycle stage, to racial identity. Research has shown that people are more likely to move at certain stages in their life cycle, often reflecting family composition changes (Clark and Withers 1999; Rossi [1955] 1980; Mincer 1978). Each stage or composition change presumably influences the likelihood of migrating as well as the type of migration. For example, the probability of migrating generally declines with age, as married people are more likely to move shorter distances while non-married people are more likely to move longer distances (Greenwood 1981). Yet it seems likely that these effects may not be overwhelmingly influential during the settlement years of the Great Plains, when both individuals and families were moving great distances to make their mark on the frontier.

However, concentrations of racial and ethnic groups are likely to influence change in the Great Plains. Scattered throughout the Great Plains are towns originally settled by immigrant

---

[3] It is worth noting that while economic base and spatial factors might be considered primary factors or structural forces shaping population change, population characteristics might be more readily conceptualized as secondary influences. For example, while population characteristics are anticipated to have some impact on population change, other forces, such as economic base, likely shape the variation in population characteristics. A boom in mining may attract younger populations, which in turn increases the likelihood of continued population growth. In this example, age influences growth, yet it does so as a consequence of economic base.

communities from places such as Germany, Sweden, and Norway. Large concentrations of Native Americans inhabiting reservations are also found throughout the region. Such concentration leads to familial and community ties that likely prevent decline, and may actually attract similar populations thereby encouraging overall population growth.

The Great Plains Context

At the beginning of the 20th century, the Great Plains region was still organizing. In fact, the western half of the plains had only recently become formal states within the union, and New Mexico and Oklahoma did not gain statehood until 1907 and 1912 respectively. In contrast, the eastern portion of the region had a longer history, lending to greater organization and development. Therefore, a great deal of growth is expected to occur during the settlement era, between 1900 and 1930, with much of it concentrated in the western portion of the Plains.

The Great Plains experienced agricultural expansion aided by World War I and innovations in agricultural technology, impacting seed, soil, and equipment. The agricultural industry boomed in the earlier years, fueled by favorable environmental conditions and an expanding need for agricultural goods stimulated by the war and increased industrial wages among consumers on the eastern seaboard. World War I decimated the agricultural productivity of Europe and, therefore, created a global need for U.S. products, much of which were grown in the Great Plains. Following the war, young soldiers returned to the farm equipped with a renewed desire for the American Dream and low land prices and a demanding global market to assist their enthusiasm.

The railroads played a significant role during the late 19th century, as the railroad companies were rapidly developing new lines to connect the expansive continent and relying on hopeful settlers to foot the bill (Webb 1931). With the land companies, the railroads spread word of hope and happiness within and beyond U.S. borders through advertisements claiming that "your crops in this fertile soil will truly bring you contentment, peace and plenty; in one word—happiness…Come and pick your farm and thus pave the road to happiness and contentment" (Rawlings Land Company Advertisement (1925) in Lange and Taylor 1999:91). Railroad companies would often reimburse passengers with land exploration tickets if they purchased any land and would even provide free transportation when traveling back to settle the recently procured lands (Hedges 1926).

When considering the long and dusty alternative of a covered wagon, the railroads were a quick, clean, and relatively attractive and affordable mode of transportation. However, the entrance of the 20th century was accompanied by the advent of the automobile. Researchers have suggested an outward movement of businesses and populations from the city as roads were first developed and automobiles became more widespread (McKenzie 1929).

Yet the Great Plains was not the leading region in highway expansion, and railroads did not reach all corners of the vast prairie. It is likely that variation in population trajectories within the region is associated with the degree of accessibility. Simply stated, those places that were easier to reach were more likely to be settled than more isolated and difficult areas. The direction of the association might be positive through its role as a receiving agency, yet it might be negative as accessibility also serves to send populations to alternative destinations. Given that railroad lines connected already established metropolitan centers, it is also likely that proximity to a metropolitan center influenced population change during the settlement era.

In sum, among the various factors identified by previous works, settlement date, economic base, and spatial factors are anticipated to have the greatest impact on population change during the settlement years of the Great Plains. More recently organized counties are likely to experience greater growth simply because these are the places undergoing "settlement." Such areas are expected to have a large agricultural presence, and are not likely to house large urban centers. Rather, it is the previously unsettled counties that are likely to experience population growth during the early years of the 20th century—counties with the promise of freedom and contentment through agriculture and the accessibility necessary to pave the road to happiness.

Data

The sample for this study follows the U.S. Geological Survey definition of the Great Plains, including approximately 876 counties within 13 states aggregated into 742 county clusters.[4] The counties lie within Montana, Wyoming, Colorado, New Mexico, Texas, Oklahoma, Kansas, Nebraska, South Dakota, North Dakota, Minnesota, Iowa, and Missouri. We have selected a broader definition than some studies of the Great Plains in an attempt to remain as inclusive as possible. Minnesota, Iowa, and Missouri are often omitted from studies of this region due to the variation in precipitation, grass length, and altitude. However, while places within these states may receive more than 20 inches of rainfall, evaporation maintains semi-arid conditions among

[4] County borders change over time. And county borders tend to change for political reasons, mainly associated with population size. An advantage of selecting 1900 as the start date is that most Great Plains states achieved statehood or were organized territories by this time. Therefore, most counties were already established. However, the last border change for counties within the sample did not occur until 1960. Until this date, counties were born through fission, died through merges, and were renamed for some fairly interesting reasons (e.g., Calhoun County, Nebraska was changed to Saunders County due to the name's negative association with a ballot-box stuffing incident).

Using a template developed by Horan et al. (1989), each county is converted into its 1900 form and given a unique county cluster code, producing 742 county clusters for analysis. This has the following implications: if county A and B are not involved in the formation of a new county there would be no change in their county code, and would each be assigned a separate county cluster code. But, if county A split to produce county B, they would share the same county cluster code. Similarly, if county B merged with county A, they would share the same county cluster code.

Some counties do not change their shape while others are dramatically different. For example, most of the counties in Iowa have not changed their boundaries since 1900, yet almost every county in Oklahoma has. In fact, in 1900, Oklahoma had not yet become a state and was largely considered "Indian Territory." The southern and northeastern parts of the state were divided between two large areas, while smaller county divisions were made in the northwestern part of the state.

the western counties within these states (Kraenzel 1955). And it is precisely this semi-arid quality that defines the Great Plains.

Data are drawn from several sources to construct a county-level data file for the years spanning 1900 through 1930. The main body of data is from the *Historical, Demographic, Economic, and Social Data: The United States, 1790-1970*, made available by the Inter-University Consortium for Political and Social Research (ICPSR) (ICPSR 1976) while data for the spatial factors was gathered by surveying Rand McNally Atlases (Rand McNally and Company 1911). Finally, environmental data was secured from The Vegetation/Ecosystem Modeling and Analysis Project (VEMAP) through Myron Gutmann (personal communication).

The Linear Growth Curve Model

We begin with estimating a linear growth curve model, which assumes that the rate of change is consistent across all time points within the observation period. Here, the dependent variable is the logged decennial county population between 1900 and 1930. According to the unconditional model, reported in Table 1, on average, counties within the Great Plains experienced considerable growth. The mean initial status and growth rate indicate that the predicted average county population was 8,714 in 1900, increasing to 10,245 in 1910, 12,045 in 1920 and 14,162 by 1930. These estimates are obtained by exponentiating the sum of the mean initial status (β_{00}) and the mean growth rate multiplied by time ($\beta_{10}T$). The GCM analysis also reveals a high and negative correlation between initial status and growth rate ($\tau = -0.782$), indicating that counties having a larger population in 1900 generally tended to change at a slower rate between 1900 and 1930 relative to counties beginning the period with a smaller population.

[Table 1 About Here]

The variance components for the growth rate (r_{10}) reveal whether there is variation about the mean trajectory. In the analysis of the Great Plains, the resulting estimates imply that there is a significant amount of variation around the mean trajectory; the estimated average growth rate is 0.162 while the estimated variation in growth is 0.085 for the linear model.

Thus far we have estimated an average growth rate for all counties and established that there is significant local level variation around this mean. The remaining portion of the analysis focuses on explaining this variation. More specifically, we turn our attention to estimating which county characteristics contribute to variation in population change.

The reader will note that the estimates for the average initial status (π_{1i}), reported in Model 1 of Table 2, is consistent with that reported in the unconditional model in Table 1. The initial status coefficient is stable simply because the influence of the various county characteristics is modeled for the growth rate and not the initial status. If no covariates are introduced to explain variation around the initial status, the estimate will remain unchanged.

[Table 2 About Here]

To review, those characteristics anticipated to have the greatest importance in shaping county population change are settlement date, economic base, and spatial factors. The relationship between settlement date and population change is more complex than originally anticipated. At first glance, it appears that when a county was settled has virtually no relationship with population change. In fact, settlement is associated with county growth at the bivariate level ($\beta = 0.002$, p-value < 0.001), but the relationship was accounted for by several factors each drawing on aspects associated with counties having later dates of settlement.

Geographically, settlement date is linked to region, where western counties were settled often considerably later than more eastern counties. New Mexico did not achieve statehood until 1907 while Iowa entered the Union in 1846. Ecologically, precipitation and temperature varies across the Great Plains region, mainly along geographic lines where the more western counties have less precipitation and more extreme temperatures. These geographic and ecological relationships are meaningful for more sociological attributes and processes. In terms of geography, spatial patterns of urban development are likely to be more mature in areas where urbanization has been granted additional time to progress. Counties in Iowa have had more time to build cities than those in New Mexico. However, New Mexico counties and other areas with later settlement dates had more room to grow at the beginning of the 20th century. In the Great Plains region, precipitation and temperature presumably have important consequences for economic base, especially those centered on agriculture (Kraenzel 1955; Webb 1931). Farmers in Iowa tend to receive greater amounts of rain and experience more tepid climates relative to their counterparts on the other side of the Plains. While counties with later settlement dates tend to have a smaller proportion of their total acreage devoted to farms, a higher proportion of their populations are employed in agriculture.

The notable influence of settlement date is explained through other county characteristics. Settlement date is associated with many attributes that are, in turn, also related to population change including temperature range, overall county acreage, initial population, and falling within the western versus the eastern region. It is precisely these factors that account for the direct effects of settlement date in the multivariate context. These findings suggest that the long-term positive effects of settlement date are indirect. For example, counties with later settlement dates have smaller initial populations, yet counties beginning the period with smaller initial populations are likely to end the era with larger populations. Settlement date also positively influences population growth through county size, such that counties with later settlement dates are more likely to have larger acreage, and larger counties are associated with higher levels of growth.

Settlement date is not the only factor contributing to variation in population change during this era. Spatial factors, including both urban proximity and railroad accessibility, are also substantially associated with population growth during the Settlement Period, at least according to the linear model. As anticipated, counties containing cities generally experienced growth rather than stability or decline. Counties containing an urban center and or neighboring an urban center differ from counties with no city or adjacency in their patterns of population growth. For instance, counties with no city or adjacency have an estimated population of 8,714 in 1900, growing to 10,090 by 1910, 11,684 by 1920, and 13,529 by 1930. In contrast, cities containing an urban center are estimated to grow at a much higher rate, such that by 1910 these counties have nearly 5,000 more residents, and by 1930 they have 30,000 more residents. Counties containing an urban center continued to grow at higher rates than the smallest and most distant counties throughout the period.

Sociologists have suggested that proximity is generally advantageous to population growth (Duncan et al. 1960; Hawley 1950; McKenzie, 1929). While the comparison of the most "urban" to the least "urban" counties supports this argument, a closer look at the estimated county populations according to proximity actually counters the proposition. For example, counties containing a moderate-sized city who also neighbor an urban center are estimated to have lower population growth relative to those that do not lie adjacent to a county with an urban center. In 1910, the estimated difference is just over 1,000, expanding to over 6,000 by 1930 with non-adjacent counties having the advantage. Further, there is no significant difference between the estimated populations for counties without cities regardless of adjacency. Containing a city rather than lying adjacent to one is more beneficial for county growth.

The second component of the spatial factors also suggests that growth between 1900 and 1930 was more likely occur in counties neither containing nor neighboring a county with a railroad. When coupling these findings with the urban proximity results, the data suggest that development at the beginning of the period did not guarantee continued growth. Rather, counties distant from urban centers or even railroads were likely to experience growth; counties with railroads are estimated to have populations of 14,365 in 1930 while those with no railroad or adjacency have an estimated 22,405 residents. Again, it appears that proximity did not promise or, in the case of railroads, encourage growth. Of course, these findings must be presented with a caveat re-emphasizing the period under study. This was a period of considerable growth in the Plains, and, it would seem from these results, the most underdeveloped counties were those most likely to experience the most rapid growth. We return to this idea in the discussion of the non-linear models.

Among the indicators tapping county economic base, only farm value (land, crops and buildings) appears to significantly contribute to county growth. Farm and manufacturing jobs, and farm acreage seem to have no significant impact. However, this lack of influence among the economic base indicators sets the stage for comparing the linear versus non-linear models, and the substantive variation accompanying the treatment of time. Importantly, these results are derived from a model where time is treated linearly; the rate of change remains consistent across each decade. As mentioned at the outset of the paper, this assumption is not always adequate, especially for studies covering larger periods of time given that change rarely occurs at an identical rate or in a linear fashion over extended periods (Shryock et al. 1976). Therefore, it is essential for the researcher to have a solid understanding of how time operates in the model and how the treatment of time influences study results.

The Non-Linear Growth Curve Model

The non-linear approach allows for variation in the rate of change across time points within the period. For this example, we explore both the quadratic and cubic form. The deviance statistic reported in Table 1, as well as Table 2, indicates model fit. When comparing the score between the unconditional linear and quadratic models we note a reduction in the statistic and, therefore, an increase in model fit by including the quadratic function of time. Yet tests for higher order functions, the cubic form, even more dramatically enhanced model fit. It is useful to couple these statistics with a visual illustration of the observed pattern of growth relative to the predicted patterns for the various forms of time.

From Figure 1, we gather that the observed pattern of growth between 1900 and 1930 can be characterized as curvilinear. Between 1900 and 1910, the Plains experienced a dramatic increase that was followed by a decline, or negative growth rate. Some research has suggested that severe drought scattered at various points on the Plains during the mid-teens may have contributed to the marked change in the average growth rate (Ottoson et al. 1966). This mid-period population downturn was followed by a return to a positive growth rate in the final decade under analysis. Among the time forms, it appears that the cubic function most accurately reflects the general pattern of the average growth trend for the Great Plains between 1900 and 1930. The linear form, by definition, assumes a steady growth rate across the entire period. In contrast, the quadratic form permits a change in the value of the growth rate over time, and estimates a negative value in the later period, but no return to positive growth as the cubic form allows.

Despite the convincing evidence that the cubic function is the most appropriation time form, we estimate the quadratic function given the instructive nature of this analysis.

[Figure 1 About Here]

Time as a Quadratic Function

As in the linear model, the mean status and mean growth rate for the non-linear model indicate that, on average, counties grew over time, reported in Table 1. However, the additional time function (β_{20}) also suggests that counties grew at a decelerating rate. By accounting for the slowing rate of change, the county population estimates are slightly different than those obtained through the linear growth curve model. The county average began at a somewhat lower size, 8,562, but then outgrew the linear model at 10,429 in 1910, 12,699 in 1920, and 15,467 by 1930. This does not imply that the rate of change in the growth rate accelerated—recall that the coefficient is negative. Rather, if the influence of time remained constant throughout the period, the non-linear estimate of time would lead to lower predicted populations in the later years of the period. So while the influence of time is 0.215, this rate of growth was reduced at a rate of -0.018 as we move out in the time series; specifically the growth rate at any time point changes by the following equation: 0.215 + -0.018(T). This would imply that at T = 2 (1920) the average rate of growth is 0.179, and by T = 3 (1930) the estimate drops to 0.161. If we were to extend the equation to 1960, the estimated rate of growth is a mere 0.107.

The variance components and chi-square estimates suggest that there is considerable variation in the growth rate as well as the rate of deceleration. As counties tend to deviate in the slope of growth, they also vary in their rate of deceleration. Consistent with the linear analysis, the negative correlation between initial population and the growth rate (τ = -0.734) suggests that counties with larger populations in 1900 tended to change at a slower rate over the period. In addition, the positive correlation between initial status and the non-linear function of time (τ = 0.637) indicates that these larger counties also decelerated at a faster rate than smaller counties. Combined, these correlations suggest that counties beginning the period with smaller populations experienced slightly greater growth relative to those beginning with larger populations.

When turning to the analysis of county factors contributing to the observed variation, Model 2 in Table 2, we note that perhaps the largest discrepancy between the linear and quadratic models concerns economic base. Several economic indicators significantly contribute to variation in county growth in the non-linear quadratic model. As in the linear model, the rate of change is positively influenced by farm value, suggesting that counties with high farm value at the beginning of the period are likely to grow at higher rates over the period. The quadratic approach also reveals that the decelerating rate of change in the growth rate, or the change in growth, is negatively associated with farm value at the beginning of the period. This implies that counties with high values in farm land, buildings, and crops are likely to experience growth during the period, and at a lower decelerating rate relative to counties beginning the era with low farm value. Otherwise stated, counties with greater agricultural value benefited during the earlier part of the period, and continued to do so throughout the period.

In addition to farm value, the proportion of the population employed in manufacturing contributed to variation in county growth. And like farm value, manufacturing employment was positively associated with the rate of growth and negatively associated with the deceleration occurring over time. Again, this implies that high levels of manufacturing employment at the outset contributed positively to growth over the period.

Regarding the historical context, it is worth noting that the Depression presumably hit the agricultural sector earlier than the non-agricultural segment (Grant 2002). Despite this impact, counties with high farm value appear to be better off in terms of population growth relative to those with lower values, at least those counties possessing high farm value prior to the Depression. This coupled with the steady positive influence of manufacturing employment suggests that counties with high farm values and a manufacturing base were better insulated from the damaging impacts of the Depression relative to those with lower values or less manufacturing.

Another important substantive difference revealed in the non-linear model is the influence, or lack of influence, of urban proximity. In the linear model, counties lying adjacent to urban centers had lower estimated populations than non-adjacent counties. This pattern is even more pronounced in the quadratic results in two important ways: (1) they do not significantly vary from the most "rural" and distant counties (both adjacent counties with and without a moderate-sized city) and (2) they have even lower estimated populations than their non-adjacent counterparts (e.g., counties with moderate-sized cities neighboring urban centers are estimated to have more than 4,200 fewer residents than their non-adjacent counterparts by 1920 and 9,200 fewer by 1930).

The quadratic analysis also suggests that having an urban center is negatively associated with the decelerating change in the growth rate. This implies that while there is no impact of adjacency, having a city is positively associated with growth, and beginning the period with a large city is associated with lower rates of deceleration over the period. Scholars have argued that the Depression had a dampening effect on urban expansion, as the back to the land movement did not take hold until the early 1930s (Woofter 1936). Accordingly, the results provide no indication of a decline among counties with cities as the period under analysis pre-dates the population upheaval.

Time as a Cubic Function

While the results from the quadratic are more revealing than those obtained through the linear treatment of time, comparison of observed patterns of growth in Figure 1 and the unconditional model fit in Table 1 suggest that the cubic form most appropriately suits the period under analysis; the population experienced overall growth between 1900 and 1910, decline between 1910 and 1920, and a return to growth from 1920 to 1930. There is much substantive similarity between the quadratic and cubic analysis results, yet the cubic analysis further clarifies the varying contribution of the county characteristics over the Settlement Period and highlights some discrepancies.

Important differences, like that between the linear and quadratic analyses, emerge for the influence of economic base, Model 3 in Table 2. First, while the initial proportion of the population employed in manufacturing significantly and positively contributes to growth during the first part of the period, it does not during the remaining decades. This stands in direct contrast to the earlier models where the initial level of manufacturing employment was positively associated with growth throughout the period. Second, according to the linear and quadratic models, the original proportion of county acreage devoted to farmland was not associated with growth. However, the cubic model reveals that while there was no evidence of a relationship at the period's outset, one did emerge with the deceleration of the rate of change (about 1910-20) and persisted during the acceleration (about 1920-30); counties beginning the period with higher proportions of land engaged in farming had lower levels of deceleration and higher levels of acceleration in population growth. This association appears after the first major bout of growth on the Plains, suggesting that even while underdeveloped counties may have experienced considerable growth during the first part of the period, established farm counties were likely to continue growing throughout the remaining decades. Perhaps providing additional support to this argument is the consistent positive association between initial farm value and population growth. Recall that while the coefficient is negative for Time 2, it implies that higher farm values were negatively associated with the *deceleration* in growth.

There are also two noteworthy differences in the otherwise consistent storyline for the spatial factors. As in the previous models, having a city, whether large or modest, corresponds with population growth. However, proximity emerges as a contributor to growth among counties with no city by the end of the period. Despite this statistical significance, the estimated population for the respective urban proximity categories continues to indicate that counties not neighboring urban centers had higher population growth. For example, by 1930 there is more than a 14,000-person difference between adjacent and non-adjacent counties with a moderate-size city and a 1,000-person difference among counties with no city, both with a non-adjacent advantage.

The second spatial difference concerns accessibility. In both the linear and quadratic models, counties without a railroad or adjacency were significantly more likely to experience growth relative to counties with a railroad. Yet in the cubic treatment of time, at no point is the presence or nearness to a railroad associated with population change, indicating that accessibility at the outset meant little for variation among county growth over the period.

SUMMARY AND DISCUSSION

In their handbook for demographic analysis, Shryock et al. (1976:213-214) state that calculating the average *amount* of change is straightforward, but estimating the *rate* of growth is more meaningful for comparisons. We would like to extend this claim, and posit that estimating the factors influencing *variation* in the rate of growth is even more meaningful for understanding disparities between comparison groups. Growth curve modeling offers an effective strategy to analyze the overall average growth rate, variation around this mean, and the factors driving the variation. In addition, the technique provides researchers with an approachable means of understanding the influence of time over extended observation periods.

Using population change in the Great Plains between 1900 and 1930, we have demonstrated the efficiency of the GCM techniques. This approach enables us to estimate a process covering multiple time points in a single analysis rather than pasting together results from multiple models to produce a story of population change over an extended time period. Time was treated as both a linear and non-linear function—quadratic and cubic—and comparisons reveal that the cubic form of time is most suitable to characterizing population change over this period within the Great Plains. Rather than maintaining a constant rate of growth, the rate of population change varied over the thirty year period, with rapid growth at the outset followed by a decreasing rate of growth and a return to acceleration during the middle and later portions.

Results also indicate that while the region experienced overall growth, this average trend was accompanied by notable variation. Not all counties shared the same growth rate. Moreover, the factors associated with this deviation slightly varied depending upon the time function. While settlement date (indirectly) influenced growth in all of the models, the extent and nature of the relationship between population growth and urban proximity and economic base varied. The linear model implied that urban proximity was positively associated with growth for the entire period. In this analysis, "urban" appeared to have a more significant relationship with growth than did "proximity." This difference was further highlighted in the quadratic model, such that adjacent counties not only had lower estimated populations than their non-adjacent counterparts, but the adjacent categories were not associated with growth at any point. In addition, only counties beginning the period with an urban center appeared to have a significant advantage over the most "rural" counties in this model, where having an urban center was positively associated with growth and negatively associated with the deceleration. This relationship was further clarified in the cubic model. Counties with an urban center at the outset were more likely have higher populations throughout the period due to the positive association with growth and acceleration and the negative relationship with mid-period deceleration. Non-adjacent counties with a moderate-size city shared in this pattern. The important difference in the cubic model from the others, aside from the more appropriate treatment of time and the accompanying relationships, is the emergence of adjacency as a significant contributor to growth late in the period. Although not substantively overwhelming, as there was only a 1,000-person advantage among adjacent counties for those counties without a city, it does entice one to seriously consider the point in a county's development when adjacency emerges as a meaningful contributor to population growth. These findings might suggest that such a relationship comes in later stages.

The second key discrepancy revealed when comparing the three time functions concerns economic base. There was very little influence in the linear model, with only initial farm value significantly contributing to variation in growth, yet more influence in the non-linear models. In the quadratic from, the proportion of the population employed in manufacturing accompanied farm value, both positively associated with growth and negatively associated with deceleration in the growth rate. Yet in the cubic treatment of time, farm value remained influential throughout the period, while initial manufacturing employment was only associated with growth at the outset and not during the remaining decades. Instead, the proportion of acreage devoted to farmland at the beginning of the era emerged during the middle and later points of the period, negatively associated with the deceleration and positively associated with the acceleration in the growth rate. Counties beginning the period with high farm acres and values appear to have had a greater advantage in terms of population growth.

Clearly, this method is more applicable to some research questions, and not at all appropriate for others. Simply stated, GCM is most useful for questions addressing change over multiple time points and not practical for those concerning only one or two observation points; more than two repeated observations are necessary. This technique is also suitable for questions concerning the extent of and factors contributing to variation in initial status, rates of change, or both. While the current example focuses on variation in the rate of change, this method can also be used to simultaneously estimate disparities in initial status as well as rate of change.

One drawback to this approach for demographers is the slightly more complex use of growth rate. Rather than calculating the growth rate for each county and then using this value as the dependent variable, GCM uses the observed population and estimates an average growth rate for all counties within the sample. Yet we believe that the interpretation of this mean growth rate, and the influence of the various factors on this rate, is no more complicated than interpreting the coefficients from other regression analyses. Moreover, this strategy offers greater insight into processes of change and in a much more succinct and efficient manner.

Perhaps the greatest limitation is that the model cannot simultaneously account for the influence of both time and space. While there are methods for spatial regression in the hierarchical context, the spatial unit is treated as the repeated observation rather than time. Such an approach will address spatial issues, but not questions of change over time. Some research has employed a two-stage procedure to estimate the effects of both factors (see Sampson et al. (1999) in addition to Fotheringham et al. (2000) for a description of potential methods as well as references to other studies addressing spatial dependence in the multilevel context). So although not inherent within GCM techniques, issues of spatial dependence can be addressed through supplemental procedures.

REFERENCES

Bryk, A. S. and S. W. Raudenbush (1992). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park: Sage Publications.

Christaller, W. (1966). *Central Places in Southern Germany*. Englewood Cliffs, NJ: Prentice Hall.

Clark, W. A. V. and S. Davies-Withers (1999). "Changing Jobs and Changing Houses: Mobility Outcomes of Employment Transitions." *Journal of Regional Science* 39(4): 653-673.

Cromartie, J. B. (1998). "Net Migration in the Great Plains Increasingly Linked to Natural Amenities and Suburbanization." *Rural Development Perspectives* 13(1): 27-34.

Duncan, T. E., Duncan S. C., Strycker, L. A., Li, F., and A. Alpert. (1999). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications.* Mahwah, NJ: Lawrence Erlbaum Associates.

Duncan, O. D., W. R. Scott, et al. (1960). *Metropolis and Region*. Baltimore: Johns Hopkins Press.

Durkheim, E. (1933). *The Division of Labor in Society*. Trans. G. Simpson. New York: Free Press.

Fotheringham, A. S., C. Brunsdon, and M. Charlton (2000). *Quantitative Geography: Perspectives on Spatial Data Analysis*. London: Sage Publications

Fuguitt, G. V., D. L. Brown, C. L. Beale. (1989). *Rural and Small Town America*. New York: Russell Sage Foundation.

Grant, M. J. (2002). *Down and Out on the Family Farm: Rural Rehabilitation in the Great Plains, 1929-1945*. Lincoln, NE: University of Nebraska Press.

Greenwood, M. J. (1981). *Migration and Economic Growth in the United States: National, Regional, and Metropolitan Perspectives*. New York: Academic Press.

Gregory, J. N. (1989). *American Exodus: The Dust Bowl Migration and Okie Culture in California*. New York: Oxford University Press.

Hawley, A. (1950). *Human Ecology: A Theory of Community Structure*. New York: Ronald Press Co.

Heaton, T. B., W. B. Clifford, and G. V. Fuguitt. (1981). "Temporal Shifts in the Determinants of Young and Elderly Migration in Nonmetropolitan Areas." *Social Forces* 60(1):41-60.

Hedges, J. B. (1926). "The Colonization Work of the Northern Pacific Railroad."

The Mississippi Valley Historical Review 13(3):311-342.

Horan, P. M., P. G. Hargis, and M. S. Killian (1989). "Longitudinal Research on Local Labor Markets: The County Longitudinal Template." *Research in Rural Sociology and Development* 4:99-121.

Hox, J. J. and I. G. G. Kreft (1994). "Multilevel Analysis Methods." *Sociological Methods and Research* 22(3):283-299.

Huttenlocher, J. E., W. Haight, A. S. Bryk, and M. Seltzer (1991). "Early Vocabulary Growth: Relation to Language Input and Gender." *Developmental Psychology* 27(2):236-249.

Inter-university Consortium for Political and Social Research. 1976. *Historical, Demographic, Economic, and Social Data: The United States, 1790-1970* [Computer file]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor].

Johansen, H. E. and G. V. Fuguitt (1984). *The Changing Rural Village in America: Demographic and Economic Trends Since 1950*. Cambridge: Ballinger Publications Co.

Kraenzel, C. F. (1955). *The Great Plains in Transition*. Norman, OK: University of Oklahoma Press.

Lange, D. and P. Taylor (1999). *An American Exodus: A Record of Human Erosion*. Paris: Jean Michel Place.

McArdle, J. J. and D. Epstein. (1987). "Latent growth curves within developmental structural equation models." *Child Development* 58:110-133.

McKenzie, R. D. (1929). "Ecological Succession in the Puget Sound Region." *American Sociological Society* 23.

Mincer, J. (1978). "Family Migration Decisions." *Journal of Political Economy* 286: 749-773.

Nagin, D. S. and R. E. Tremblay. (1999). "Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency." *Child Development* 70(5): 1181-1196.

Ottoson, Howard W., Eleanor M. Birch, Philip A. Henderson, and A.H. Anderson. 1966. *Land and People in the Northern Plains Transition Area*. Lincoln, NE: University of Nebraska Press.

Raudenbush, S. W. and A. S. Bryk. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods, 2nd Edition*. Newbury Park: Sage Publications.

Rand McNally and Company. (1911). Rand McNally and Company's Commercial Atlas of America. Chicago: Rand McNally and Company.

Rossi, P. H. ([1955] 1980). *Why Families Move*. London: Sage Publications.

Sampson, R. J., J. D. Morenoff, and F. Earls (1999). “Beyond Social Capital: Spatial Dynamics of Collective Efficacy for Children.” *American Sociological Review* 64(5):633-660.

Shryock, H. S., Siegel, J. S., and E. G. Stockwell. (1976). *The Methods and Materials of Demography.* San Diego: Academic Press.

Sjaastad, J. A. (1962). “The Costs and Returns of Human Migration.” *Journal of Political Economy* 70: s80-s93.

Stegner, W. E. (1954). *Beyond the Hundredth Meridian: John Wesley Powell and the Second Opening of the West*. Boston: Houghton Mifflin.

Stephan, G. E. (1971). “Variation in County Size: A Theory of Segmental Growth.” *American Sociological Review* 36:451-461.

Webb, W. P. (1931). *The Great Plains*. Boston: Ginn and Company.

Woofter, T. J., Jr. (1936). “Rural Relief and the Back-to-the-Farm Movement.” *Social Forces* 14(3):382-388.

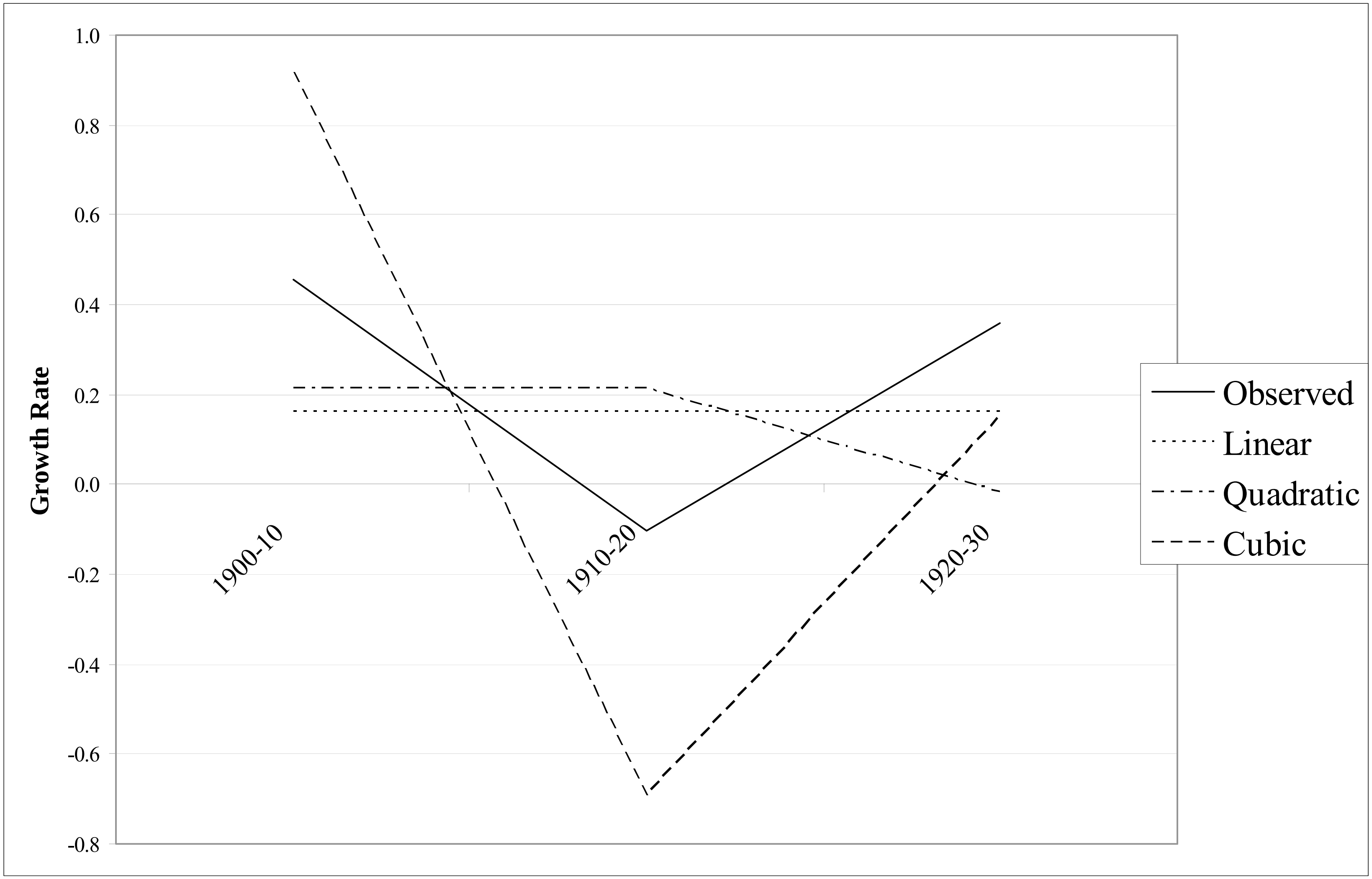

Figure 1. Observed and Estimated Growth Rates among Great Plains Counties between 1900 and 1930 (N = 742)

Table 1. Unconditional Linear and Non-Linear Models of Growth in Great Plains County Population between 1900 and 1930 (N = 742)

Linear Model				
Fixed Effect	*β*	*SE*	*p-value*	
Mean Initial Status, β_{00}	9.073	0.050	0.000	
Mean Growth Rate (Time), β_{10}	0.162	0.012	0.000	
Random Effect	*Variance Component*	*df*	X^2	*p-value*
Initial Status, r_{00}	1.761	741	20238.53	0.000
Growth Rate (Time), r_{10}	0.085	741	4032.76	0.000
Correlation between Change and Initial Status	-0.782			
Deviance Statistic	5303.48			
Quadratic Model				
Fixed Effect	*β*	*SE*	*p-value*	
Mean Initial Status, β_{00}	9.055	0.052	0.000	
Mean Growth Rate, β_{10}	0.215	0.025	0.000	
Mean Growth Rate Squared ($Time^2$), β_{20}	-0.018	0.006	0.002	
Random Effect	*Variance Component*	*df*	X^2	*p-value*
Initial Status, r_{00}	1.917	741	19538.57	0.000
Growth Rate, r_{10}	0.302	741	1764.35	0.000
Growth Rate Squared ($Time^2$), r_{20}	0.007	741	843.60	0.005
Correlation between Change and Initial Status	-0.734			
Correlation between $Change^2$ *and Initial Status*	0.637			
Deviance Statistic	5087.96			
Cubic Model				
Fixed Effect	*β*	*SE*	*p-value*	
Mean Initial Status, β_{00}	9.010	0.053	0.000	
Mean Growth Rate, β_{10}	0.918	0.049	0.000	
Mean Growth Rate Squared ($Time^2$), β_{20}	-0.691	0.032	0.000	
Mean Growth Rate Cubed ($Time^3$), β_{30}	0.150	0.007	0.000	
Random Effect	*Variance Component*	*df*	X^2	*p-value*
Initial Status, r_{00}	1.946	741	10199.40	0.000
Growth Rate, r_{10}	1.055	741	1823.03	0.000
Growth Rate Squared ($Time^2$), r_{20}	0.362	741	1424.35	0.000
Growth Rate Cubed ($Time^3$), r_{30}	0.015	741	1345.74	0.000
Correlation between Change and Initial Status	-0.882			
Correlation between $Change^2$ *and Initial Status*	0.948			
Correlation between $Change^3$ *and Initial Status*	-0.968			
Deviance Statistic	4005.09			

Table 2. Influence of County Characteristics in the Linear and Non-Linear Models of Growth in County Population between 1900 and 1930 (N = 742)

	Linear Model 1			Quadratic Model 2				
	β (Time)	*SE*	*df*	*β (Time)*	*SE*	*β (Time2)*	*SE*	*df*
Mean Initial Status, π_{0i}								
Intercept, β_{00}	9 07 ***	0 05	741	9 06 ***	0 05	9 06 ***	0 05	741
Mean Growth Rate, π_{1i}								
Intercept, β_{10}	3 97 **	1 27	723	1 10	3 08	1 10	3 08	723
Mean Growth Rate Squared, π_{2i}								
Intercept, β_{20}						0 76	0 92	723
Mean Growth Rate Cubed, π_{3i}								
Intercept, β_{30}								
Environmental Conditions								
Average Monthly Precipitation	-0 001	0 009		-0 01	0 02	0 004	0 01	
Temperature Range	-0 01 ***	0 002		-0 01 *	0 004	0 001	0 001	
Economic Base								
Proportion Farm Acres, per 100k	0 04	0 04		-0 04	0 07	0 03	0 02	
Farm Value, per $1k	0 0003 **	0 0001		0 001 *	0 001	-0 0003 *	0 0002	
Proportion Farm Jobs	0 03	0 14		0 07	0 15	-0 01	0 04	
Proportion Manufacturing Jobs	0 03	0 01		0 11 ***	0 03	-0 03 ***	0 01	
Population Characteristics								
Proportion Age 21 and Older	-0 13	0 18		-0 62 **	0 21	0 16 *	0 06	
Proportion Non-White	0 15	0 15		0 43	0 26	-0 09	0 08	
Proportion Foreign-Born	0 07	0 08		-0 26	0 22	0 11	0 06	
Proportion of German Ancestry	-0 30	0 23		0 60	0 42	-0 29 *	0 12	
Presence of Indian Community	0 08 *	0 03		0 23 **	0 09	-0 05 *	0 03	
Spatial Factors								
Presence of City of 25k or More	0 38 ***	0 06		0 62 ***	0 12	-0 07 *	0 03	
Presence of City of 10k or More, and Adjacent to City of 25k or More	0 11 ***	0 03		0 14	0 21	-0 01	0 06	
Presence of City of 10k or More, and Not Adjacent to City of 25k or More	0 20 ***	0 03		0 27 *	0 13	-0 02	0 04	
No City, and Adjacent to a City	0 02	0 02		-0 07	0 06	0 03	0 02	
No City, and Not Adjacent to a City	-	-		-	-	-	-	
Presence of Railroad	-	-		-	-	-	-	
Adjacent to Railroad	-0 05	0 03		0 01	0 05	-0 02	0 02	
Not Adjacent to Railroad	0 15 *	0 06		0 19 *	0 09	-0 01	0 03	
Controls								
Date Settled	-0 0005	0 001		0 001	0 002	-0 0005	0 0005	
Initial Population (natural log)	-0 30 ***	0 01		-0 24 ***	0 02	0 0001	0 006	
Acres, per 100k	0 004 **	0 001		0 01 ***	0 003	-0 002 **	0 001	
West Region	0 02	0 05		0 08	0 10	-0 02	0 03	
Deviance Statistic	5012 26			4953 27				

Note: *** p < 0 001, ** p < 0 01, * p < 0 05

Table 2. (continued)

	Cubic Model 3						
	β (Time)	*SE*	*β (Time2)*	*SE*	*β (Time3)*	*SE*	*df*
Mean Initial Status, π_{0i}							
Intercept, β_{00}	9 01 ***	0 05	9 01 ***	0 05	9 01 ***	0 05	741
Mean Growth Rate, π_{1i}							
Intercept, β_{10}	-10 93 *	4 51	-10 93 *	4 51	-10 93 *	4 51	720
Mean Growth Rate Squared, π_{2i}							
Intercept, β_{20}			6 36	3 41	6 36	3 41	720
Mean Growth Rate Cubed, π_{3i}							
Intercept, β_{30}					-0 87	0 76	720
Environmental Conditions							
Average Monthly Precipitation	-0 01	0 03	0 01	0 03	-0 001	0 01	
Temperature Range	-0 02 **	0 01	0 01 *	0 004	-0 002 *	0 001	
Economic Base							
Proportion Farm Acres, per 100k	0 17	0 10	-0 17 *	0 08	0 04 **	0 02	
Farm Value, per $1k	0 003 ***	0 001	-0 002 **	0 001	0 0003 **	0 0001	
Proportion Farm Jobs	0 21	0 21	-0 16	0 16	0 03	0 04	
Proportion Manufacturing Jobs	0 11 *	0 04	-0 02	0 03	-0 002	0 01	
Population Characteristics							
Proportion Age 21 and Older	-0 85 **	0 31	0 34	0 23	-0 04	0 05	
Proportion Non-White	0 74	0 38	-0 34	0 29	0 05	0 06	
Proportion Foreign-Born	-0 65 *	0 32	0 50 *	0 24	-0 09	0 05	
Proportion of German Ancestry	-0 06	0 61	0 28	0 46	-0 12	0 10	
Presence of Indian Community	0 20	0 13	-0 04	0 09	-0 001	0 02	
Spatial Factors							
Presence of City of 25k or More	1 27 ***	0 17	-0 71 ***	0 13	0 14 ***	0 03	
Presence of City of 10k or More, and Adjacent to City of 25k or More	0 57	0 31	-0 43	0 24	0 09	0 05	
Presence of City of 10k or More, and Not Adjacent to City of 25k or More	0 58 **	0 19	-0 32 *	0 14	0 07 *	0 03	
No City, and Adjacent to a City	0 07	0 09	-0 11	0 06	0 03 *	0 01	
No City, and Not Adjacent to a City							
Presence of Railroad							
Adjacent to Railroad	0 11	0 07	-0 10	0 06	0 02	0 01	
Not Adjacent to Railroad	0 07	0 12	0 10	0 09	-0 03	0 02	
Controls							
Date Settled	0 002	0 002	-0 002	0 002	0 0003	0 0004	
Initial Population (natural log)	0 93 ***	0 03	-0 46 ***	0 02	0 06 ***	0 005	
Acres, per 100k	0 02 ***	0 004	-0 01 ***	0 003	0 002 **	0 001	
West Region	-0 60 ***	0 15	0 65 ***	0 11	-0 15 ***	0 02	
Deviance Statistic	3489 44						

Note: *** p < 0 001, ** p < 0 01, * p < 0 05

Center for Demography and Ecology
University of Wisconsin
1180 Observatory Drive Rm. 4412
Madison, WI 53706-1393
U.S.A.
608/262-2182
FAX 608/262-8400
comments to: kwhite@ssc.wisc.edu
requests to: cdepubs@ssc.wisc.edu