Population Projections

M. V. George, Stanley K. Smith, David A. Swanson, and Jeff Tayman

"Population Projections," chapter 21 in Jacob Siegel and David Swanson (eds.), *The Methods and Materials of Demography.* San Diego: Elsevier Academic Press, 2004.

CHAPTER 21

Population Projections*

M. V. George, Stanley K. Smith, David A. Swanson, and Jeff Tayman

*Preliminary version of a chapter published in Jacob S. Siegel and David A. Swanson (editors), *The Methods and Materials of Demography*. San Diego: Elsevier Academic Press, 2004. This chapter updates and extends the chapter on population projections published in Henry S. Shryock and Jacob S. Siegel (editors), *The Methods and Materials of Demography*. Washington DC: U.S. Census Bureau, 1973.

INTRODUCTION

Estimates, Projections, and Forecasts

Demographers are frequently called upon to produce population information when census and related data are not available. Information about a present or past population is called an *estimate*. As discussed in the previous chapter, there are many ways to make population estimates. Some methods update information from the most recent census using ratio, regression, or component techniques. They often use data from sample surveys or administrative records. Others use various techniques of interpolation to develop estimates for dates between censuses. Some methods provide estimates only for the total population, whereas others provide estimates by age, sex, race, and a variety of other demographic and socioeconomic characteristics.

Demographers typically refer to information about the future as either a *projection* or a *forecast*. Although these two terms are often used interchangeably, they can be differentiated according to the expected likelihood of their outcomes. A *projection* may be defined as the numerical outcome of a particular set of assumptions regarding the future population. It is a conditional calculation showing what the future population would be if a particular set of assumptions were to hold true. Because a projection does not attempt to predict whether those assumptions actually will hold true, it can be incorrect only if a mathematical error is made in its calculation. Although a given projection can be judged by the merits of its assumptions in relation to the use to which it may be put, it can never be proven right or wrong by future events.

A *forecast* may be defined as the projection that is selected as the one most likely to provide an accurate prediction of the population. As such, it represents a specific viewpoint regarding the validity of the underlying data and assumptions. A forecast reflects a judgment and it can be proven right or wrong by future events (or, more realistically, it can be found to have a relatively small or large error). *Projection* is a more inclusive term than *forecast*: All forecasts are projections but not all projections are forecasts. Projections and forecasts sometimes refer solely to total population, but often include information on age, sex, race, and other characteristics as well.

Distinctions among the terms *estimate*, *projection*, and *forecast*, are not always clear-cut. When the data needed for population estimates are not available, techniques ordinarily used for population projections are sometimes used for calculations of current and past populations. A government statistical agency may view its calculations of future population as projections, but data users may interpret them as forecasts. In this chapter we use the term *estimate* to refer to a present or past population and *projection* to refer to a future population, regardless of their intended uses or the methodology employed. We use the term *forecast* for particular projections when discussing their accuracy.

Uses of Population Projections

Population projections can be used for a number of purposes. They provide a tool for analyzing the components of growth and the sensitivity of underlying assumptions. Projections can raise our understanding of the determinants of population change. For example, what impact would a 20% decline in birth rates have on a country's population size and age structure in 50 years? How would eliminating all deaths due to a particular cause affect the population growth rate? How many people would move into a local area if a new factory employing 1,000 people were opened?

Projections also can be used to provide information on possible future scenarios. Because we cannot "see" into the future, it is helpful to consider a range of scenarios based on different but reasonable assumptions. Alternative scenarios provide an indication of potential variations in future demographic trends, which facilitates planning for "worst-case" outcomes. Specific outcomes can be used to sound warnings about the perceived negative implications of particular trends and to call for actions directed toward preventing those outcomes from occurring.

Perhaps the most important use of population projections is in the role they can play as a rational basis for decision-making. Changes in population size and composition have many social, economic, environmental, and political implications; for this reason, population projections often serve as a basis for producing other projections (e.g., births, households, families, school enrollment, and labor force). Population projections help decision makers in both the public and private sectors make informed choices.

National population projections, for example, can be used to plan for future Social Security and Medicare obligations (Lee and Tuljapurkar, 1997; Miller, 2001). State projections can be used to determine future water demands (Texas Water Development Board, 1997) and welfare expenditures (Opitz and Nelson, 1996). Local projections can be used to determine the need for new public schools (Swanson et al., 1998) and to select sites for fire stations (Tayman, Parrott, and Carnevale, 1994). Business enterprises use forecasts to predict demands for their products (Thomas, 1994) and to anticipate the costs of current and retired employees (Kintner and Swanson, 1994). Population projections can be used to forecast the demand for housing (Mason, 1996), the number of people with disabilities (Michaud, George, and Loh, 1996), and the number of sentenced criminals (Oregon Office of Economic Analysis, 2000).

Population projections take advantage of the two strong points of demography described in Chapter 1: The accurate recording of demographic processes over a period of years; and the momentum that links demographic processes for one time period with those for another. Because the future is intimately tied to the past, projections based on past trends and relationships raise our understanding of the dynamics of population growth and often provide forecasts of future population change that are sufficiently accurate to support good decision-making. The diverse and increasingly influential roles played by population projections make them an important part of modern demographic analysis.

Population Projection Methods

Population projections may be prepared using either subjective or objective methods. Subjective methods are those in which data, techniques, and assumptions are not clearly identified;

consequently, other analysts cannot replicate them exactly. Objective methods are those for which data, techniques, and assumptions are clearly identified, such that other analysts can replicate them exactly. We do not cover subjective methods in this chapter, but it is important to note that even objective methods require choices regarding variables, data sources, projection techniques, and so forth. At some level, every projection method requires the application of judgment.

Following Smith et al. (2001), we classify objective methods into three broad categories: (1) Trend extrapolation; (2) Cohort-component; and (3) Structural models. *Trend extrapolation* methods are based on the continuation of observable historical trends. For methods of this type, future values of a variable are determined solely by its historical values. The *cohort-component* method divides the population into age-sex cohorts and accounts for the fertility, mortality, and migration behavior of each cohort. A variety of techniques can be used to project each of the three components of population growth. *Structural Models* rely on observed relationships between demographic and other variables (e.g., land uses, employment) and base population changes on projected changes in those other variables. The relationships in structural models are typically developed using regression analysis and variants thereof. In actual application, methods in these three categories are not always mutually exclusive. For example, applications of the cohort-component method often incorporate trend extrapolations of one type or another, and structural models are often used in conjunction with the cohort-component method.

Data Sources

Population projections are influenced not only by the methods and assumptions used in their production, but also by the historical data series upon which they are based. Census counts and post-censal estimates typically serve as the empirical foundation for the population upon which projections are based, while vital statistics and immigration data serve as the empirical foundation for births, deaths, and immigration. Other data used as a basis for population and related projections include social security enrollees, school enrollment, employment files, voter registration lists, change-of-address records, and property tax records. Data from sample surveys are sometimes used as well. Accurate and comprehensive data are essential for the production of useful projections.

Alternative Series

The magnitude, distribution, and composition of future populations are far from certain. To reflect this uncertainty, the producers of population projections often publish a number of alternative series rather than a single series. The production of alternative series based on different assumptions is common in "official" population projections throughout the world (Australian Bureau of Statistics, 2000; Bongaarts and Bulatao, 2000; George, 2001; and Mosert and van Tonder, 1987). Alternative series are sometimes based on different projection methods, but a more common approach is to apply different combinations of assumptions using a single method. In the cohort-component method, for example, alternative series are frequently based on different combinations of assumptions concerning mortality, fertility, and migration. The number of alternative series can vary considerably; recent applications by the U. S. Census Bureau have had as many as 30 (Spencer, 1989) and as few as two (Campbell, 1996). According to a 1988 survey on methodological issues of national projections in 31 countries, 23 computed more than one variant (Keilman, 1991). The most common practice is to produce two, three, or four alternative series.

Several interpretations can be given to the alternative series in a set of projections. One is that each gives a reasonable view of future population change and that no one series is preferable to any other. The U.S. Census Bureau gave this interpretation to its projections of state populations between the 1950s and early 1990s. Not only did the Census Bureau decline to designate a "most likely" series, but also it explicitly stated that *none* of the projections was intended as a forecast (Wetrogan, 1990). Another interpretation is that, although each alternative series is reasonable, one is preferable to all the others. This is the interpretation the Census Bureau gave its set of state projections in the mid-1990s (Campbell, 1996). Both interpretations are common, but the current tendency among the producers of population projections seems to be to designate one particular series as the most likely (i.e., as the forecast). However, the production of alternative series is not the only way to deal with uncertainty.

Geographic Areas

Population projections may be prepared for the world as a whole, for major regions of the world, for nations, and for a variety of subnational areas such as states, provinces, departments, cities, counties, census tracts, enumeration districts, postal areas, school districts, and individual blocks. Although many of the factors affecting the methodology and analysis of population projections are the same for all geographic areas, there are important differences as well. First, data are more readily available and more reliable for nations than for subnational areas and for large subnational areas than small subnational areas. Second, migration typically plays a greater role in population growth for subnational areas than for nations and for small subnational areas than for large subnational areas. Third, population growth rates are generally more variable for subnational areas than for nations and for small subnational areas than large subnational areas. Consequently, choices regarding data, techniques, and assumptions may be different for projections at one geographic than for projections at another.

Much of the research on the methodology and analysis of population projections has focused on projections at the national, regional, and global levels (Bongaarts and Bulatao, 2000; O'Neill et al., 2001; Lutz, Vaupel, and Ahlburg, 1999). However, some studies have dealt specifically with projections for subnational areas (Davis, 1995; Pittenger, 1976; Smith et al., 2001). Many of the issues we discuss in this chapter are common to all population projections, but several relate primarily to small areas.

Organization of this Chapter

We start by discussing the major producers of international, national, and subnational projections. Next, we provide a description of the methods and materials used in preparing three basic types of population projections: (1) trend extrapolation; (2) the cohort-component method; and (3) structural modeling. We briefly discuss methods for preparing related projections on such topics as school enrollment, employment, and households. We follow this with a review of issues that we believe should be considered when preparing or evaluating population projections, including a discussion of forecast accuracy. We close with several conclusions regarding the nature and utility of population projections.

Before proceeding, it is useful to define six terms that are frequently used in describing population projections. Although not universal, these terms are widely used and generally understood by those working in the field. They are: (1) base year; (2) launch year; (3) target year; (4) base period; (5) projection horizon; and (6) projection interval. The *base year* is the year of the earliest data used to make a projection, the *launch year* is the year of the most recent data used to make a projection, and the *target year* is the year for which the population is projected. The *base period* is the number of years between the base year and launch year, while the *projection horizon* is the number of years between the launch year and target year. The *projection interval* is the time increment for which projections are made (e.g., annually or every five years).

PRODUCERS OF POPULATION PROJECTIONS

International

Three major agencies produce population projections for the entire world, its major regions, and virtually all countries. These are: (1) the United Nations (UN); (2) the World Bank; and (3) the International Programs Center, which is part of the Population Division of the U.S. Census Bureau. The UN published its first comprehensive set of national, regional, and global population projections in 1958. It published its second set in 1966 and has published a new set every two years since 1978 (O'Neill et al., 2001: 207). UN (1998) projections provide information on the age and sex structure of the population and include several variants based on different combinations of assumptions. These projections incorporate information from the latest round of censuses in each country and use the latest vital statistics and international migration data.

The World Bank began producing national, regional, and global population projections in 1978 and has produced a new set every few years since that time. Some sets have included several alternative series, others only a single series. Until the mid-1990s these projections were published in various issues of the *World Development Report*. Since then, they have been produced only for internal use (O'Neill et al., 2001: 208).

The Population Division (International Programs Center) of the U.S. Census Bureau began producing national, regional, and global projections in 1985 and publishes updates approximately every other year (O'Neill et al., 2001: 208). These projections are available on-line in its "International Data Base," which covers 227 countries and the major regions of the world (U.S. Census Bureau, 2001). Currently, projections of total population are available in 10-year intervals through 2050 and projections by age and sex are available for 2000 and 2025.

Several other agencies also produce international projections. The Population Reference Bureau publishes projections for all countries of the world, using a combination of projections produced by other agencies and those produced internally. The International Institute for Applied Systems Analysis (IIASA) produced several sets of projections for the world and 13 of its regions during the 1990s. With the assistance of Statistics Netherlands, the Statistical Office of the European Communities (EUROSTAT) produces national population projections by age and sex (generally 3 scenarios) for the countries of the European Union and the countries of the European Free Trade Association every three to five years (Cruijsen, 1994; EUROSTAT, 1998). Academic demographic centers (e.g., Australian National University), private sector entities (e.g., The Futures Group), and other specialized institutions (e.g., U.S. National Research Council) also conduct research on population projections.

National Producers

Many agencies produce national-level projections for a single country. Typically, these agencies are parts of the national governments of the countries involved. The projections vary tremendously in terms of methodology, assumptions, quality of input data, frequency of production, length of projection horizon, and amount of detail provided.

The U.S. Census Bureau began producing projections of the U.S. population in the 1940s and has published updated projections a few times each decade ever since. Although there have been numerous changes in assumptions, application techniques, demographic detail, alternative series, and projection horizons, the Census Bureau has used some form of the cohort-component method for every set of its national projections (Long and McMillen, 1987). A recent set included nine principal alternative series through 2050, each with projections by single year of age, sex, race, and Hispanic origin (Day 1996a). The alternative series were based on combinations of different assumptions regarding fertility rates, mortality rates, and levels of net immigration. The most recent set of national projections, released in 2000, provide, for the first time, projections of population by nativity, and extend the time horizon to 100 years (Hollmann, Mulder, and Kallan, 2000).

Subnational Producers

A variety of government agencies, research institutes, and private businesses produce subnational population projections. In the United States, for example, the U.S. Census Bureau makes projections for states; most state governments (or their designees) make projections for counties in their states (and for the state as a whole); and many local and regional governments make projections for cities, census tracts, block groups, and other small areas. Private businesses make (or compile from other sources) projections for states, counties, subcounty areas, and a variety of customized geographic areas and demographic subgroups. Subnational projections have become increasingly common over the last few decades, especially for small areas. Similar trends have occurred in other countries, including Australia (Australian Bureau of Statistics, 2000), Canada (George, 2001), India (Indian Office of the Registrar General, 2001), Israel (Israeli Central Bureau of Statistics, 1987), New Zealand, (Statistics New Zealand, 2000), and virtually all countries in Europe (Kupiszewski and Rees, 1999).

METHODS

Trend Extrapolation

"Trend extrapolation" involves fitting mathematical models to historical data and using these models to project future population values. Relatively low costs and small data requirements make trend extrapolation methods useful, not only in demography, but in other fields as well (Armstrong, 2001). Although there are many different methods by which historical values can be modeled, it is convenient to organize these methods into three categories: (1) *Simple extrapolation methods*, which require data for only two dates and for which we discuss three approaches, linear change, geometric change, and exponential change; (2) *Complex* *extrapolation methods*, which require data for a number of dates and for which we discuss four approaches, linear trend, polynomial curve, logistic curve, and ARIMA time series; and (3) *Ratio extrapolation methods*, in which the population of a smaller area is expressed as a proportion of the population of its larger, "parent" area, and for which we discuss three approaches, constant-share, shift-share, and share-of-growth.

Although there are exceptions, trend extrapolation methods are used much more frequently for projections of total population than for projections of population subgroups (e.g., race or ethnic groups). We illustrate these methods using annual total population data for two counties, Island and Walla Walla, in the state of Washington, USA, for the period 1960 to 2000 (Washington State Office of Financial Management, 2000). This constitutes a longer base period than the 10 to 20 years that are generally sufficient for applying trend extrapolation methods; we employ this data set simply as a heuristic device.

The Washington data are shown in Table 21-1, while projections for 2005, 2010, and 2015 are shown in Table 21-2. Figure 21-1 shows the population change in both Island and Walla Walla counties from 1960 to 2000. Note that during this period, Island County grew rapidly while Walla Walla County grew very slowly. We return to this fact in our summary comments on trend extrapolation methods.

(TABLE 21-1 ABOUT HERE)

(FIGURE 21-1 ABOUT HERE)

Linear Change

This method assumes that in the future a population will change by the same amount over a given period (e.g., a year) as occurred during the base period. Average absolute change during the base period can be computed as:

$$\Delta = (P_l - P_b) / (y)$$

where Δ is the average absolute change, P_l is the population in the launch year, P_b is the population in the base year, and y is the number of years in the base period (i.e., the number of years between the base year, b, and the launch year, l). A projection using this method can be computed as:

$$P_t = P_l + [(z)(\Delta)]$$

where P_t is the population in the target year, P_l is the population in the launch year, and z is the number of years in the projection horizon (i.e., the number of years between the target year, t, and the launch year, l), and Δ is the average absolute change computed for the base period.

For Island County, a model expressing its average absolute change between 1960 and 2000 is computed as Δ = 1,364.05 = (74,200 – 19,638)/(40), and a projection for 2010 is computed as 87,840 (where 87,840 ≈ 74,200 + [(10)(1,364.05)]). For Walla Walla County, a model expressing average absolute change over the same forty year period is computed as Δ= 300.13 = (54,200 – 43,195)/(40), and a projection for 2010 as 57,201 (where 57,201 ≈ 54,200 + [(10)(300.13)]). Projections for both counties in 2005 and 2015 are found in Table 21-2.

Geometric Change

This method assumes that a population will change by the same percentage rate over a given increment of time in the future as during the base period. The average geometric rate of population change during the base period can be computed as:

$$r = [(P_l / P_b)^{(1/y)}] - 1$$

where r is the average geometric rate of change, P_l is the population in the launch year, P_b is the population in the base year, and y is the number of years in the base period. A projection using this method can be computed as:

$$P_t = (P_l) [(1 + r)^z]$$

where P_t is the population in the target year, P_l is the population in the launch year, r is the average geometric rate of change, and z is the number of years in the projection horizon.

For Island County, the annual rate of geometric change between 1960 and 2000 is computed as $r = 0.0338 = [(74{,}200 / 19{,}638)^{(1/40)}] - 1$, and a projection for 2010 as $103{,}459 \approx [(74{,}200) [(1 + 0.0338)^{10}]$. For Walla Walla County, its annual rate of geometric change for the same forty year period is computed as $0.0063 = [(54{,}200 / 42{,}195)^{(1/40)}] - 1$, and a projection for 2010 as $57{,}713 \approx [(54{,}200)(1 + 0.0063)^{10}]$. Projections for both counties in 2005 and 2015 using these models are provided in Table 21-2.

Exponential Change

The exponential change approach is closely related to the geometric, but it views change as occurring continuously rather than at discrete intervals. The exponential rate of population change during the base period can be computed as:

$$r = [\ln (P_l / P_b)] / (y)$$

where r is the average annual exponential rate of change, ln represents the natural logarithm, P_l is the population in the launch year, P_b is the population in the base year, and y is the number of years in the base period. A population projection using this method can be computed as:

$$P_t = (P_l)(e^{rz})$$

where P_t is the population in the target year, P_l is the population in the launch year, e is the base of the system of natural logarithms (approximately 2.71828), r is the average exponential rate of change computed for the base period, and z is the number of years in the projection horizon.

For Island County, the annual rate of exponential change from 1960 to 2000 is computed as 0.0332 = [ln (74,200 / 19,638)]/(40), and a projection for 2010 as 103,416 ≈ (74,200)($e^{0.03332*10}$). For Walla Walla County, its annual rate of exponential change for the same forty year period is computed as 0.0063 = [ln (54,200 / 42,195)]/(40), and a projection for 2010 as 5,774 ≈ (54,200)($e^{00.0063*10}$). Projections for both counties in 2005 and 2015 using these models are provided in Table 21-2.

Complex Extrapolation Methods

Unlike simple methods, complex extrapolation methods are constructed using base period data for more than two dates. Accordingly, they can deal better with non-linear population change. They also offer a quantitative basis for constructing measures of forecast uncertainty because statistical algorithms are used to estimate model parameters (Swanson and Beck, 1994). However, these features do not guarantee that complex extrapolation methods provide more accurate forecasts than simple extrapolation methods.

Typically, three basic steps are followed when applying complex extrapolation methods. The first is to assemble historical population data for different dates during the base period. For a model to be valid, the data must be based on consistently defined geographic boundaries for each time point. The second step is to estimate the parameters of the model selected to generate the projection. Typically, graphs and statistical measures are used to determine how well a given model is fits the data for base period while the choice of a particular model reflects judgment about the nature of future population change. The final step is to generate projections using the model(s) selected.

A critical issue in the use of complex extrapolation methods is the selection of time units. This is important because time can be measured in several different ways and the one selected for a given problem affects the scale of certain parameters estimated by the curve fitting process. Using the data in Table 21-1 as an example, consider two linear trend models, one using the original units for time (i.e., 1960 through 2000) and one using a logically equivalent alternative (i.e., 1 through 41, where 1 corresponds to 1960 and 41 corresponds to 2000). Both models will have the same fit with the historical data (e.g., the r^2 values will be the same), but the intercept will be different. Consistent time units must be used when estimating complex models and using them to project future population values.

Linear Models

Linear models are the simplest of the complex trend extrapolation methods. They assume that a population will change by the same numerical amount in the future as in the past. This assumption is identical to that underlying the simple linear method discussed earlier, but the model is computed differently:

$$Y_i = a + [(b)(X_i)]$$

where Y_i is a set of i observations of values of a "dependent variable," X_i is a set of i observations of an "independent variable," *a* is the constant term, and *b* is the slope of the line describing the "best fitting linear" relationship between X and Y, as found by, e.g., the method of least squares (NCSS, 1995: 1309-1310). In using this approach for purposes of fitting a population projection model, it is convenient to recast X and Y as time and population, respectively. That is, as:

$$P_i = a + [(b)(T_i)]$$

where P_i is the population for a set of time points (e.g., years) over the period i = b to l (b= base year and l = launch year); *a* and *b* are the estimated intercept and slope, respectively, and T_i is time over the period i = b to l.

For Island County, a linear trend model using the NCSS "linear trend growth model" routine (NCSS, 1995: 1371-1388) was estimated as $P_i = 12{,}639.64 + [(1{,}504.282)(T_i)]$, with an r^2 value of 0.986. The slope value implies that the population of Island County increased on average by 1,054 persons each year of the base period. With this model, a projected population for Island County in 2010 is 89,358 ≈ 12,639.64 + [1,504.282)(51)]. For Walla Walla County, the linear trend model was estimated as $P_i = 39{,}002.14 + [(358.0467)(T_i)]$, with an r^2 value of 0.928. With this same model, a projected 2010 population for Walla Walla County is 57,263 ≈ 39,002.14 + [358.0467)(51)]. Projections for both counties in 2005 and 2015 are shown in Table 21-2.

Polynomial Models

Polynomial models can be used for projections in which change is not constrained to be linear. The general formula for a polynomial curve is:

$$Y_i = a + (b_1)(X_i) + (b_2)(X_i^2) + (b_3)(X_i^3) + \ldots + (b_n)(X_i^n)$$

where Y_i is a set of i observations of values of a "dependent variable," X_i is a set of i observations of an "independent variable," and a represents the constant term and b_j the slope of the line describing the "best fitting" relationship between X_i^j and Y, holding constant the effects of X_i^k, (where $k \neq j$). Recasting X and Y as time, and population, respectively, we have:

$$P_i = a + b_1(T_i) + b_2(T_i^2) + b_3(T_i^3) + \ldots + b_n(T_i^n)$$

where P_i is the population for a set of time points over the period i = b to l (b= base year and l = launch year), a is the estimated intercept term, *bi* are the estimated partial slope coefficients, and T_i is time over the period i = b to l.

In contrast to linear trend models, polynomial models have more than one term reflecting the independent variable (time). Consequently, there are more parameters to estimate. The coefficients of a polynomial curve (a, b_1, b_2, …b_n) can be estimated using OLS regression techniques (NCSS, 1995: 1309-1310). These coefficients include both a measure of the linear trend (b_1) and measures of the nonlinear patterns (b_2, b_3,…, b_n).

To illustrate the use of a polynomial curve for population projections, we use a second-degree polynomial (sometimes called a *quadratic function*). This function includes time (the linear term) and time squared (also called the *parabolic term*) on the right-hand side of the equation:

$$P_i = a + (b_1)(T_i) + (b_2)(T_i^2)$$

where b_1 is the slope for the linear trend and b_2 is the slope for the nonlinear (parabolic) trend. A quadratic curve can produce a variety of growth scenarios, such as a population growing at an increasing rate, a population growing at a decreasing rate, a population declining at an increasing rate, or a population declining at a decreasing rate. Projections based on a quadratic curve can lead to very high (or low) projections for places that were growing (or declining) rapidly during the base period. Although any degree can be used, polynomials higher than a "second" or at most a "third" degree are seldom used for population projections. Non-linear trends in the historical data also can be projected using curves based on logarithmic or other transformations of the base data.

For Island County, a quadratic model using the NCSS "multiple regression" routine (NCSS, 1995: 155-188) was estimated as $P_i = 16{,}432.8 + (974.998)(T_i) + (12.602)(T_i^2)$, with $r^2 = 0.993$. With this same model, a projected population for Island County in 2010 is $98{,}936 \approx 16{,}432.8 + (974.998)(51) + (12.602)(51^2)$. For Walla Walla County, the quadratic model was estimated as $P_i = 40{,}983.2 + (81.618)(T_i) + (6.5816)(T_i^2)$, with $r^2 = 0.963$. With this model, a projected population for Walla Walla County in 2010 is $62{,}264 \approx 40{,}983.2 + (81.618)(51) + (6.5816)(51^2)$. Projections for both counties in 2005 and 2015 are shown in Table 21-2.

Logistic Models

Unlike the extrapolation methods considered so far, the logistic approach explicitly allows one to place an upper limit on the ultimate size of the population for a given area. It is designed to yield an S-shaped pattern representing an initial period of slow growth rates, followed by a period of increasing growths, and finally a period of declining growth rates that approach zero as a population approaches its upper limit. The logistic model is consistent with Malthusian and other theories of constrained population growth.

Keyfitz (1968: 215) provides the following formula for a 3-parameter logistic curve:

$$Y = a / [1 + (b(e^{-cX}))]$$

where Y is the population; X is the time period; *a* reflects the upper asymptote; *b* and *c* are parameters that define the shape of the logistic curve; and e is the base of the natural logarithm. Note that other formulas are available, some including more than three parameters (NCSS, 1995: 1375; Pielou, 1969: 19-32).

In using the logistic curve, one must sometimes determine the magnitude of the upper asymptote and the time required in reaching it. However, there are algorithms available that estimate these parameters within the context of the model, but like parameters in an ordinary regression model (e.g., the intercept term), the estimated parameters may not be consistent with a substantive interpretation (e.g., *a* represents an actual upper population limit).

For our purposes, Keyfitz's formula is re-written as:

$$P_t = (a)/ [1 + ((b)(e^{-ct}))]$$

where P_t is the population in the target year, *a*, *b*, and *c* are the estimated parameters, and t is time. This model is useful because it can be generated by the NCSS package without the user having to provide pre-determined population limits; rather, the NCSS algorithm uses the available historical information to generate the needed parameters.

For Island County, a logistic model using the NCSS 3-parameter logistic model routine (NCSS 1995: 1375) was estimated as $P_t = (118{,}272.7) / [1 + (6.061401(e^{-0.05843697t}))]$ after 11 iterations, with $r^2 = 0.995293$. With this model, the projected population for Island County in 2010 is $90{,}437 \approx (118{,}272.7) / [1 + (6.061401(e^{-0.05843697*51}))]$. For Walla Walla County, the logistic model was estimated as $P_t = (41{,}712{,}310.00) / [1 + (1{,}059.927(e^{-0.007814847t}))]$ after 235 iterations, with $r^2 = 0.940$. With this model, the projected population for Walla Walla County in 2010 is $58{,}542 \approx (41{,}712{,}310.00) / [1 + (1{,}059.927(e^{-0.007814847*51}))]$. (Projections for both counties in 2005 and 2015 using these models, respectively, are provided in Table 21-2.

ARIMA Time Series Models

ARIMA ("Autoregressive Integrated Moving Average") models have been used in the analysis and projection of populations as a whole and of their demographic attributes (Alho and Spencer, 1997; Carter and Lee, 1986; De Beer, 1993; Land, 1986; Lee, 1993; McDonald, 1979; McKnown and Rogers, 1989; Pflaumer, 1992; and Saboia, 1974). The procedures used in ARIMA models are complicated, making them difficult to implement and explain to data users. We suggest consulting standard texts before attempting to apply this method (Box and Jenkins, 1976; Hanke et al., 2001; Yaffee, 2000). It also may be useful to review the method of "moving averages," which forms part of the foundation of ARIMA (Hanke et al., 2001: 101-123). ARIMA models attempt to uncover the stochastic mechanisms that generate historical data series and use this information as a basis for developing projections. Three processes can describe the stochastic mechanism: (1) autoregressive; (2) differencing; and (3) a moving average. The autoregressive process has a memory in the sense that it is based on the correlation of each value of a variable with all preceding values. The impact of earlier values is assumed to diminish exponentially over time. The number of preceding values incorporated into the model determines its "order." For example, in a first-order autoregressive process, the current value is explicitly a function only of the immediately preceding value. However, the immediately preceding value is a function of the one before it, which is a function of the one before it, and so forth. Consequently, all preceding values influence current values, albeit with a declining impact. In a second-order autoregressive process, the current value is explicitly a function of the two immediately preceding values; again, all preceding values have an indirect impact.

A stationary time series is very important for the construction of a given ARIMA model. The differencing process is used to create a stationary time series (i.e., one with constant differences over time). When a time series is non-stationary, it can often be converted into a stationary time series by calculating differences between values. First differences are usually sufficient, but second differences are occasionally required (i.e., differences between differences). Logarithmic and square root transformations can also be used to convert non-stationary to stationary time series. The moving average is used to represent any event that has a substantial but short-lived impact on a time series pattern. The order of the moving average process defines the number of time periods affected by a given event.

The ARIMA method is usually written as ARIMA (p,d,q), where p is the order of the autoregression, d is the degree of differencing, and q is the order of the moving average. (An ARIMA model based on time intervals of less than one year may also require a seasonal component.) The first and most subjective step in developing an ARIMA model is to identify the values of p, d, and q. The d-value is determined first because a stationary series is required to properly identify the autoregressive and moving average processes. The value of d is usually 0 or 1, but occasionally 2. Like the d-value, the p- and q-values are also relatively small (0, 1, or—at most—2). The patterns of the autocorrelation function (ACF) and the partial autocorrelation function (PACF) and their standard errors are used to find the correct values for p and q (Box and Jenkins, 1976; Yaffee, 2000). For example, a first-order autoregressive model [ARIMA (1,0,0)] is characterized by an ACF that declines exponentially and quickly along with a PACF that has a statistically significant spike only at the first lag. Once p, d, and q are determined, maximum likelihood procedures are used to estimate the parameters of a given ARIMA model.

The final step involves assessing the suitability of a given model. An adequate ARIMA model will have random residuals, no significant values in the ACF and PACF, and the smallest possible values for p, d, or q. Only after an ARIMA model has passed this assessment, should it be used.

It is not unusual to repeat this sequence of steps several times before a suitable ARIMA model is found. In this process, it is best to start simple (e.g., ARIMA (0,1,0)), check the results, and then add additional changes in a systematic and incremental manner (e.g., ARIMA(1,1,0), if the model is not found to be suitable. If a suitable model is not found by the time one reaches, say, ARIMA (2,1,2), it is probably wise to abandon this approach.

One characteristic of an autoregressive model is that projections will eventually reach and maintain a constant numeric difference similar in value to the mean of the historical series (McCleary and Hay, 1980: 218). Consequently, population projections using ARIMA will often be similar to projections based on linear extrapolation methods (Pflaumer, 1992; Voss and Kale, 1985). The formulas used in computing projections from an ARIMA model depend on the specification of the values of p, d, and q, each of which is most easily determined using a computer package designed for estimating them.

Using the 1960 through 2000 population figures shown in Table 21-1, the NCSS (1995: 1427-1436) ARIMA routine was used to develop models for Island and Walla Walla counties. Because the NCSS algorithm does not use ordinary least squares, it employs "pseudo-r^2" as a measure of fit rather than "r^2."

For Island County, a first order-auto regressive model with one degree of differencing and no moving average (ARIMA(1,1,0)) was found to be sufficient after 21 iterations and estimated as $P_{t+1} = [((0.876942)(P_t - P_{t-1})) + P_t]$, with "pseudo-$r^2$" = 0.9982. With this same model, a projected population for Island County in 2010 is 78,889 ≈ [((0.876942)(78,646.5 – 78,370.4)) + 78,646.5]. For Walla Walla County, a first order-auto regressive model with one degree of differencing and no moving average (ARIMA(1,1,0)) also was found to be sufficient after three iterations, with a "pseudo-r^2" of 0.98505. In the case of Walla Walla County, it was estimated as $P_{t+1} = [((0.4573916)(P_t - P_{t-1})) + P_t]$. With this same model, a projected population for Walla Walla County in 2010 is 53,863 ≈ [((0.4573916)(53,863 – 53,863)) + 53,863]. Projections for both counties in 2005 and 2015 are shown in Table 21-2.

There are three important points to bear in mind when using ARIMA models. First, because a given ARIMA model can have a number of alternative configurations, those that have values higher than 1 for p, d, and q, can be difficult to operationalize manually. That is, it may be difficult to manually recreate projection results created by a given software package using the parameters shown. This may be of particular concern for some users because it is not easy to explain how the projections are calculated. Second, during the base period "projected values" from a given ARIMA model are generated using the actual historical values. However, once beyond the scope of the historical data, projected values themselves are used to generate subsequent projected values (There may be a period subsequent to launch date when a combination of actual and projected data may be used depending on the order and degree of differencing). Third, ARIMA (as well as the other complex extrapolation techniques covered here) can be used to place probabilistic confidence intervals around their forecasts.

Ratio Extrapolation Methods

Ratio extrapolation methods may be used where an area containing the population to be projected is part of a larger ("parent") area for which projections are available. They are often used where areas exist in a perfect hierarchical structure; that is, where geographic units at each level are mutually exclusive and exhaustive and can be aggregated to higher levels, culminating in one all-inclusive unit. As an example, consider census blocks in the U. S., which can be aggregated successively into block groups, census tracts, counties, states, and finally the country as a whole. Ratio methods also can be used where there is not a perfect hierarchy – for example, in a city that is part of a county in which the geographic area (population) covered by all cities is less than the total area (population) of the county as a whole. In this case, the parent area is not the county but the area represented by all of the cities. Ratio methods can be applied in situations where the area (population) of interest is linked to the area (population) through considerations other than geography (this is akin to the "targeting" approach described later in regard to projecting mortality and fertility) and, in addition, ratios can be formed via lagged relationships.

We discuss three commonly used ratio methods: (1) Constant Share; (2) Shift-Share: and (3) Share-of-Growth. As noted, all three require projections of a "parent" area in which the area of interest is located. We use Washington State as the "parent" area for applying these methods to Island and Walla Walla counties. Projections for Washington State are: 6,137,403, 6,545,786, and 6,987,273, for the years 2005, 2010, and 2015, respectively (Washington State Office of Financial Management, 2001).

Constant-Share

In this method, the smaller area's share of the larger area's population is held constant at a level observed during the base period. Typically it is as the share observed in the launch year. This *constant-share* method is expressed as:

$$P_{it} = (P_{il} / P_{jl})(P_{jt})$$

where P_{it} is the population projection for smaller area (i) in the target year; P_{il} is the population of the smaller area in the launch year; P_{jl} is the population of the parent area (j) in the launch year; and P_{jt} is the projection of the parent area in the target year.

The constant share method requires historical data from only one point in time; consequently, it is particularly useful for areas where changing geographic boundaries or poor records make it difficult or impossible to construct a reliable historical data series. Another attribute of this method is that projections for all of the smaller areas add exactly to the projection for the parent area. The main drawback of this method is that it assumes that all the smaller areas will grow at the same rate as the parent area. In many instances, this will not be a reasonable assumption.

Using the 2000 population figures shown in Table 21-1 for Island County and Washington State, a constant share model yielded a projected 2010 population for Island County of 83,692 ≈ (74,200/5,803,400)(6,545,786). Using the 2000 population figures shown in Table 21-1 for Walla Walla County and Washington State, a constant share model yielded a projected 2010 population for Walla Walla County of 61,133 ≈ (54,200/5,803,400)(6,545,786). Projections for 2005 and 2015 using these models, respectively, are shown in Table 21-2.

Shift-Share

Unlike the constant share method, the *shift-share* method is designed to deal with changes in population shares. Here, we describe one of several methods in which population shares are extrapolated linearly over time. This *shift-share* method is expressed as:

$$P_{it} = (P_{jt})[(P_{il} / P_{jl}) + ((z/y)((P_{il} / P_{jl}) – (P_{ib} / P_{jb})))]$$

where the smaller area is denoted by i; the parent area by j; z is the number of years in the projection horizon; y is the number of years in the base period; and b, l, and t refer to the base, launch, and target years, respectively.

There is a problem inherent in the shift-share method: it can lead to substantial population losses in areas that grew very slowly (or declined) during the base period, especially when the projections cover long-range horizons (e.g., 20 or 30 years). In fact, this method can even lead to negative numbers. This problem can be dealt with by incorporating constraints on projected population shares or on the projected rates of change in those shares. The method also can lead to absurdly high projections for areas that have been growing very rapidly. As with many extrapolation methods, shift-share must be used very cautiously for long-range projections, especially for places whose population shares have been declining (or increasing) rapidly.

Using the 1960 and 2000 population figures shown in Table 21-1, a shift-share model yielded a projected 2010 population for Island County of 93,352 ≈ (6,545,786)[(74,200 / 5,803,400 + ((10/40)((74,200/ 5,803,400) – (19,638/2,853,214)))]. For Walla Walla County, a shift-share model yielded a projected 2010 population of 52,216 ≈ (6,545,786)[(54,200 / 5,803,400) + ((10/40)((54,200 / 5,803,400) – (42,195 / 2,853,214)))]. Projections for 2005 and 2015 are shown in Table 21-2.

Share-of-Growth

The third ratio method deals with shares of population *change* rather than population *size*. In this method, it is assumed that the smaller area's share of population change in the parent area will be the same over the projection horizon as it was during the base period. This *share-of-growth* method can be expressed as:

$$P_{it} = P_{il} + [((P_{il} - P_{ib}) / (P_{jl} - P_{jb}))(P_{jt} - P_{jl})]$$

where the components are defined as those in the shift-share method

In many instances, the share-of-growth method seems to provide more reasonable projections than either the constant or shift-share methods. However, it runs into problems when a growth rate in a smaller area has the opposite sign that that for the parent area. This can be dealt with using the "plus-minus" method described in Appendix C or by setting the share to zero and not letting it change.

Using the 1960 and 2000 population figures shown in Table 21-1, a shift-share model yielded a projected 2010 population for Island County of 87,930 ≈ 74,200 + [((74,200 – 19,638) / (5,803,400 – 2,853,214))(6,545,786 – 5,903,400)]. For Walla Walla County, a shift-share model yielded a projected 2010 population of 57,221 ≈ 54,200 + [((54,200 – 42,195) / (5,803,400- 2,853,214))(6,545,786 – 5,803,400)]. Projections for 2005 and 2015 are shown in Table 21-2.

(TABLE 21-2 ABOUT HERE)

Summary Comments on Extrapolation Methods

Both simple and complex trend extrapolation methods suffer from several shortcomings. They do not account for differences in demographic composition or for differences in the components of growth. They provide little or no information on the projected demographic characteristics of the population. Because they have no theoretical content, they cannot be related to theories of population growth, except perhaps the logistic model, which is consistent with a Malthusian view of population dynamics. Consequently, they have limited usefulness for analyzing the determinants of population growth or for simulating the effects of changes in particular variables or assumptions. In addition, they can lead to unrealistic or even absurd results, even over relatively short horizons. In spite of their shortcomings, trend extrapolation methods have a number of advantages over other projection methods. They have few data requirements, and, with the exception of the ARIMA and Polynomial models, are quick and easy to apply. They are particularly useful when data series are incomplete, time and budgets are highly constrained, and information on population characteristics is not needed. Perhaps most important, they often provide reasonably accurate forecasts over short projection horizons. There is no empirical evidence showing that more complex or sophisticated methods consistently produce more accurate forecasts than trend extrapolation methods.

As shown in Table 21-2, different methods sometimes produce dramatically different results. Island County, for example, grew rapidly between 1960 to 2000. Thus, it is not surprising that the range of projections is quite large, extending from 79,719 (logistic) to 122,166 (geometric). The case is quite different for Walla Walla County, which experienced a much smaller population increase between 1960 and 2000. Here, the range is much smaller than that for Island County: the largest projection for 2015 is provided by the quadratic method (66,194) and the smallest (50,978) by the shift-share method. If trend extrapolation methods are to be used, which one(s) should be chosen for these two counties? Alternatively, should an average of projections from several methods be calculated? Should the same methods be used in both counties? The use of trend extrapolation methods does not remove the need to exercise judgment. We return to this issue toward the end of this chapter.

COHORT-COMPONENT METHOD

The cohort-component method was introduced by Cannan (1895), subsequently used by Bowley (1924), and later re-discovered independently by Whelpton (1928). It is the most widely used method for producing national-level population projections. Although current applications are more detailed and sophisticated than the earliest applications, the basic framework of the method has changed little since the pioneering work by these three men.

The cohort-component method divides the launch-year population into age-sex groups (i.e., cohorts) and accounts separately for the fertility, mortality, and migration behavior of each cohort as it passes through the projection horizon. It is a flexible and powerful method that can be used to implement theoretical models or serve as an atheoretical accounting procedure. It can provide in-depth knowledge on population dynamics. Also the cohort-component method can accommodate a wide range of assumptions and can be used at any level of geography - from the world as a whole down to nations, states/provinces, counties, and subcounty areas.

For purposes of population projection, the division of the population into age groups was an important methodological advance (de Gans, 1999). It allows one to account for the differences in mortality, fertility, and migration rates among different age groups at a particular time, and to consider how rates change over time for individual cohorts. In implementing the cohort-component method, age groups are typically divided by sex and often further subdivided by race or ethnicity.

Cohort-component models typically use either single years or 5-year groups. The oldest age group is virtually always "open-ended," usually 75+ , 85+, or 90+. Age groups are typically divided by sex and are sometimes further subdivided by race, ethnicity, and other characteristics. Our discussion and examples focus on populations divided by age and sex, but the procedures we describe would be basically the same if the population were further subdivided by race, ethnicity, and other characteristics.

In the application of the method as originally developed, the first step in the projection process is to establish the launch year population and calculate the number of persons in it who survive to the end of the projection interval. This is done by applying age-sex-specific survival rates to each age-sex group in the launch year population. The second step is to calculate migration during the projection interval for each age-sex group. The application of migration rates provides a projection of the number of persons in each age-sex group moving into or out of an area during the projection interval (or, for models using net migration data, the net change due to migration). The third step is to calculate the number of births occurring during the projection interval. This is accomplished by applying age-specific birth rates to the female population in each age group. The final step in the process is to add the number of births (distinguishing between males and females) to the rest of the population. These calculations provide a projection of the population by age and sex at the end of the projection interval. This population then serves as the starting point for the following interval. The process is repeated until the final target year is reached. Figure 21-2 illustrates the steps in this process. Figure 21-2 is meant to be illustrative. Not each and every application of the cohort-component method follows the steps shown in it.

Illustrating this point, in the examples for national and subnational projection provided later for Canada, there are adjustments made to the basic steps shown in Figure 21-2 to deal with the effect of migration on births and deaths.

(FIGURE 21-2 ABOUT HERE)

Projecting Mortality

The mortality rates (or their functional equivalents) used in cohort-component projections can be projected in a number of ways. The simplest is to assume that age-specific rates will remain unchanged at current levels. For short horizons, this will often be a reasonable assumption. For longer horizons, however, this assumption may not be valid and methods that incorporate changing rates then become necessary. Such methods include a variety of extrapolation techniques, techniques tying mortality rates in one area or population to those in another, and structural models that base changes in mortality rates on changes in socioeconomic variables.

The use of extrapolation techniques assumes that the future will mirror the past in certain important ways. Although this is not always a valid assumption, it has often led to reasonably accurate forecasts. Extrapolation techniques have been widely used for mortality projections, sometimes following fairly simple procedures and other times applying more sophisticated procedures such as ARIMA time series models.

Extrapolation techniques are particularly appropriate when mortality trends are following a stable path. The critical question is whether trends will remain stable in the future. The answer to this question depends on one's views regarding the determinants of mortality trends (e.g., Ahlburg and Vaupel, 1990; Fries, 1989; Manton, Stallard, and Tolley, 1991; Olshansky, 1988).

A number of techniques tie mortality rates in one population to those in another. For example, the "targeting" approach is based on the idea that mortality rates in a given population will converge toward those observed in another population (i.e., the target). A target population is chosen which provides a set of mortality rates believed to be realistic for the population to be projected. This choice is based on similarities in socioeconomic, cultural, and behavioral characteristics; levels of medical technology; and primary causes of death (Olshansky, 1988: 500). It could be implemented using a "ratio" approach, as described earlier in this chapter in the section, "Trend Extrapolation."

One form of targeting is called "cause-delay." In this approach, the target population is a younger cohort in the same population rather than the same cohort in a different population. Cause-delay models focus on the implications of delaying (or completely eliminating) the occurrence of one or more causes of death (Manton, Patrick, and Stallard, 1980; Olshansky, 1987). The basic premise behind this approach is that changes in lifestyle and medical technology have delayed the occurrence of various types of deaths until progressively older ages. Consequently, as time goes by, each cohort faces lower mortality risks at each age than did the previous cohort. Another form of targeting is to link changes in rates for the area in question to changes projected for a different area through the use of ratios.

Life tables for most areas of the world are readily available and can be used as a base for mortality projections. Life tables are prepared both by the United Nations and by the vital statistics agencies in most countries. In the United States, for example, national life tables are published annually by the National Center for Health Statistics and a few times each decade by the Social Security Administration (Bell et al., 1992; National Center for Health Statistics, 1997). State life tables are typically constructed once every ten years in the United States, when decennial census data become available to serve as denominators for the mortality rates.

Projecting Fertility

In projecting births, one can use a period perspective, a cohort perspective, or a combination of the two. The period perspective focuses on births to women during a particular period of time (e.g., one year). The cohort perspective is longitudinal, focusing on the fertility patterns of a cohort of women as they pass through their childbearing years.

Although the cohort perspective is superior for some analytical purposes, it is difficult to implement when constructing population projections. Data on completed cohort fertility do not become available until after a cohort has passed through its childbearing years; for women under age 50, only partial data are available. Birth histories of past cohorts and the fertility expectations of current cohorts may be used as proxies for the missing data, but they do not necessarily provide reliable fertility forecasts for current and future cohorts. In addition, projections for subnational areas are complicated by the lack of relevant data and by the effects of migration, which may have a significant impact on the composition of an area’s population and its fertility behavior. Because of these problems and the complexity of the method, we believe that most practitioners will be better served by using a period perspective for the production of fertility projections.

Several approaches may be used when applying the period fertility perspective. One is to hold current age specific birth rates (ASBRs) constant throughout the projection horizon (Day, 1996; Treadway, 1997). These rates are often based on data for the most recent year, but can also be based on an average of data for several recent years. Holding rates constant can be justified not only by the expectation that they will not change, but also by the belief that increases in current rates are as likely as declines. Another approach is to extrapolate historical trends. This approach is particularly useful for countries in the midst of the demographic transition from high to low fertility rates, but can lead to problems for countries in which the transition has already been completed. If no long-run trends are clearly discernible, recent changes in fertility rates may simply reflect short-run fluctuations and extrapolating those changes into the future may create large forecast errors. Time series techniques are frequently used to develop nonlinear models for projecting ASBRs (Carter and Lee, 1992; Land, 1986; Lee, 1993; Lee and Tuljapurkar, 1994).

The "targeting" approach described earlier is another approach that can be used to project fertility rates (U.S. Bureau of the Census, 1979). Before applying this approach, however, one must decide whether the convergence of one set of rates toward another is a realistic assumption. Projected ASBRs also can be created by forming ratios of birth rates in one area to those in another and applying those ratios to the birth rates previously projected for the area of interest. Finally, structural models can be developed. Such models have occasionally been used for projecting fertility rates at the national level (e.g., Ahlburg, 1999; Sanderson, 1999), but have seldom been used at the subnational level (Isserman, 1985).

Even given our argument in favor of the period perspective, it is nonetheless important to keep the cohort perspective in mind when formulating assumptions about future fertility rates. For example, if recent changes have occurred in ASBRs, the cohort perspective may offer clues as to whether those changes reflect a shift in the long-run trend in ASBRs or simply a short-run change in the timing of births. It is in this sense that a combination of the period and cohort perspective may be useful. In addition, it is useful to note that there is increased use of the cohort approach in the European Economic Area, a development partially attributed to the availability of long series of historical data (Cruijsen, 1994).

Fertility is often the most problematic part of national population projections (Keyfitz, 1982; Ryder, 1990). Long (1989) found that differences in fertility assumptions accounted for more of the variation in long-run projections of the U.S. population than did differences in either mortality or immigration assumptions. However, immigration was considered the most problematic component for the 1993 round of Canadian population projections (George, Loh, and Verma, 1997). For subnational areas in many countries, fertility may be less important than migration in explaining differences in rates of population growth (Congdon, 1992; Smith and Ahmed, 1990).

Projecting Migration

Migration often has a substantial impact on population growth at the national and subnational levels. For projection purposes, we may view migration two different ways: (1) gross migration; and (2) net migration. As explained in Chapter 19, gross migration refers to the movement of people into and out of a given area; net migration refers to the difference between the two; that is, in-migration minus out-migration. For projecting population flows, each has its strengths and weaknesses. Gross migration models are “cleaner” from a theoretical and computational standpoint, but require more data and are more complicated to apply than net migration models (Smith and Swanson, 1998). Both approaches are widely used for population projections.

We discuss two basic techniques for projecting gross migration. The first is based on the application of out-migration rates and in-migration proportions for each area to be projected. (U.S. Census Bureau, 1966, 1972, 1979). We describe this technique using states as the unit of reference and migration data based on five-year time and age intervals. The same technique could be used for other geographic areas and different lengths of migration intervals, if the data were available. Under this technique, out-migration rates by age and sex are calculated for each state using out-migration data from the decennial census as the numerators and state populations by age and sex (five years earlier) as the denominators. These rates are then applied to the launch year population to provide a projection of the total "pool" of interstate out-migrants for all states. Migrants in this pool are allocated to each state according to the proportion of interstate migrants each state received during the base period (by age and sex).

The second technique for projecting gross migration uses "multi-regional models" (Rogers, 1985, 1995). In these models, migration is viewed as part of an integrated system of mortality, fertility, and origin-destination-specific population streams by age and sex (and sometimes by other characteristics as well). For example, interstate migration in a multi-regional model of the United States could be represented by a 51 by 51 matrix showing the number of people moving from each state to every other state (including the District of Columbia), by age and sex. Migration rates are calculated by dividing destination-specific gross migration streams by the population of each state of origin, giving each state 50 sets of age-sex-specific out-migration rates, one for each other state in the nation. Because these rates are based on the population at risk of migration, they reflect the probabilities of moving from one state to another during a given time period. A multi-regional approach has been used by Statistics Canada and is illustrated in the example provided later (see, e. g., Table 21-5).

Migration also can be projected using net rather than gross migration. Net migration can be projected using two approaches, either alone or in combination: (1) "top-down;" and (2) "bottom-up." The top-down approach distinguishes between the components of population growth (i.e., natural increase and net migration), but focuses on estimates of total net migration rather than separate estimates for each age-sex cohort. It requires two steps. First, projections of total net migration are made, based on recent levels, historical trends, structural models, or some other procedure. Second, these projections are made by age-sex categories, based on distributions observed in the past. We call this a "top-down" approach because projections for broad demographic categories are made first and subcategories are derived from them; here, individual age-sex groups are derived from projections of total net migration. This was the approach taken in the earliest sets of cohort-component projections made for states and regions in the United States (Thompson and Whelpton, 1933; U.S. Bureau of the Census, 1957). It is currently used for the international migration component of national population projections in the United States. Projections of the level of total net foreign immigration are based on historical data and expectations regarding future levels; they are done by age, sex, and race/ethnicity categories according to the distributions observed in recent historical data (Day, 1996).

The second approach to projecting net migration focuses on the development of separate net migration rates for each age-sex cohort in the population. Projections are based on the application of age-sex-specific net migration rates to the base population in this same detail. We call this a "bottom-up" approach because figures for the broad categories are derived from those for the subcategories; here, the total volume of net migration projected for an area is the sum of the individual values projected for each age-sex group.

Net migration models generally combine international and internal migration. When net migration is calculated as a residual, this is the simplest approach. Separate projections of immigration could be made, however, by subtracting immigration from total net migration in the base data, and developing separate assumptions regarding future net flows of foreign and domestic migrants.

One drawback of net migration models is that they do not base migration rates on the population at risk. As a consequence, they create inconsistencies in projections for a group of areas. Consider population projections for states. The application of constant net migration rates to states with rapidly growing populations leads to steadily increasing levels of net in-migration over time, but the application of constant rates to states with slowly growing (or declining) populations leads to slowly growing (or declining) levels of net out-migration. Because net internal migration must sum to zero over all states (that is, the total number of interstate in-migrants must equal the total number of interstate out-migrants), this creates an internal inconsistency within the set of state population projections. It can also lead to bias, as projections based on net migration rates tend to be too high for rapidly growing places and too low for slowly growing or declining places.

Some of the problems associated with net migration models can be reduced by changing the denominators used in constructing the migration rates. Net migration rates for rapidly growing areas can be based on the population of a larger geographic unit rather than of the area itself. For example, rates for rapidly growing states can be based on the national population rather than the state population. This change has been found to greatly reduce projected rates of increase for rapidly growing states (Smith, 1986). Alternatively, projections of net migration (or population) can be constrained or controlled in various ways to prevent unreasonably large increases or declines (Smith and Shahidullah, 1995).

Calculations for net migration and mortality can be combined to create a simplified version of the cohort-component method (Hamilton and Perry, 1962). In this method, cohort-change ratios (CCR) covering the time interval between the two most recent censuses are calculated for each age-sex cohort in the population. Also are known as *census survival rates*, they are expressed as:

$$_nCCR_x = {_nP_{x+y}} / {_nP_x}$$

where $_nP_{x+y}$ is the population aged x+y to x+y+n in the year of the most recent census; $_nP_x$ is the population age x to x+n in the second most recent census; x is the youngest age in an age interval; n is the number of years in an age interval; and y is the number of years between these two successive censuses. Cohort-change ratios also can be calculated for different race/ethnic groups. Projections can then be made by multiplying these ratios by the launch-year population in each age-sex group:

$$_nP_{x+y,\, t} = {_nCCR_x} \, ({_nP_{x,\, l}})$$

where $_nP_{x+y,t}$ is the population age x+y to x+y+n in year target year t.

In many circumstances, especially for small areas, unique events and special populations must be taken into account when developing migration assumptions. Unique events are those having a substantial but short-lived impact on an area's volume and patterns of migration – for example, an economic boom or bust may have occurred during the base period. In such cases, one will need to decide if these conditions are likely to continue into the future and, if not, how to make appropriate adjustments. Special populations are groups of people who are in an area because of an administrative or legislative action. These include refugees, college students, prison inmates, and military personnel. Changes in special populations result from a different set of causal factors than those affecting the rest of the population. If changes in special populations are substantial, it is important to account for them separately when implementing the cohort-component method (Smith et al., 2001: 239-277). Migration is considerably more susceptible than either fertility or mortality to changes in economic conditions, employment opportunities, housing patterns, transportation conditions, and neighborhood characteristics. Social and cultural conflicts, natural disasters, and government policies typically have more impact on migration rates than on mortality and fertility rates. Consequently, migration is generally more difficult to forecast accurately than either mortality or fertility, especially for small areas. In general, the smaller the subnational area, the greater the difficulty in developing accurate migration forecasts.

Implementing the Cohort-component Method

Several issues must be considered when implementing the cohort-component method. To preserve the integrity of age cohorts as they progress through time, it is helpful to follow a basic principle: The number of years in the projection interval should be greater than or equal to the number of years in the age-groups. For example, five-year age-groups are well suited for making projections in five- or 10-year intervals, but are not well suited for making projections in one-year intervals. The logic is simple: the survivors of people aged 10-14 in 2005 will be 15-19 in 2010, but making projections of persons who will be 11-15 in 2006 is more complicated and the results less precise. Typically, the model is applied separately for each demographic subgroup. These strata are then combined to create other categories. The female stratum is typically projected first because a projection of females is needed to determine the projection of births. The procedures for applying the rates for components of change are the same for each subgroup. Cohort-component models are often constructed for five-year age groups, starting with 0-4 and ending with 75+ or 85+. The use of five-year age groups is common because projections of five-year groups in five-year intervals satisfy the needs of a wide range of data users and, in addition, can be interpolated both into single years of age and into individual years within a five-year projection interval using procedures described in Appendix C.

Single-year cohort-component models also are widely used, especially at the national level. Some county-level projections use this more detailed age breakdown. Cohort-component models with single years of age automatically provide annual projections and offer an obvious advantage over models built from more aggregated age groupings. They make it easier to provide projections for customized age groups (e.g., 5-17) required by data users in areas such as education, health care, and the criminal justice system. In addition, single-year models provide a more precise reflection of population aging; by focusing on single-year cohorts as they move through time, they pick up subtleties missed by five-year models.

Single-year models are considerably more time consuming and costly to construct and maintain than five-year models. A single-year model with 100+ as the terminal age group has 202 age-sex categories. In contrast, a five-year model with 85+ as the terminal age category has only 36 age-sex categories. For a 20-year projection horizon, a single-year model requires the application of 202 separate birth, death, and migration rates for each of 20 distinct time periods. A five-year model requires only 36 birth, death, and migration rates for four time periods. In spite of the widespread use of powerful microcomputers, issues of data management for single-year models are still imposing when three or four race/ethnic groups are added to the task.

In some circumstances migration data are available only in 10-year intervals (e.g., net migration between two decennial censuses). Strictly speaking, this would dictate the use of 10-year migration rates and 10-year projection intervals. However, a common practice is to transform 10-year migration rates into five-year rates by dividing by two and averaging the rates for two adjacent birth cohorts (Note that a given 5-year cohort appears in two 10-year cohorts, but in different 5-year time periods). Another approach is to use census data on migration in the five year-period preceding the census.

A final consideration before implementing the cohort-component method is the impact of data error and data consistency. Data problems tend to increase as the level of demographic detail increases and as population size declines. It is important to verify historical population data and, if necessary, to adjust the basic demographic rates before running the projection model. In some cases, it may be necessary to adjust for census enumeration and other forms of error as well.

Example of a National Projection

We illustrate the cohort-component method for national population projections using as an example Statistics Canada's "medium scenario" projection of the female population (Statistics Canada, 2001). Statistics Canada has been preparing population projections on a regular basis since 1969. They are given for single years of age and sex, each year, with a horizon of 25 years for the provinces and territories, and 50 years for Canada as a whole. Long-term projections are generally revised every five years, following the national census. The projections employ a regional cohort-component method (the term “region” represents Canada’s 10 provinces and 3 territories). The input data for the projections (population by age and sex, fertility, mortality, immigration, emigration, non-permanent residents, and internal migration) come from official population estimates published in Statistics Canada’s, *Annual Demographic Statistics*.

In order to produce consistent and comparable projections for Canada and its provinces simultaneously, a “hybrid bottom-up” projection model is used. In this model, assumptions on fertility, mortality, immigration, emigration, and non-permanent residents are developed at the national level and consistent provincial assumptions, incorporating internal migration projections, are derived from them. The model allows separate projections of each component at the provincial/territorial level, thereby taking into account regional differences. (George and Loh, 2000). It has been the general practice to include several alternate assumptions for fertility, mortality, and migration in preparing the projections. The combination of assumptions yields numerous projections from which a set of projections is selected for publication purposes representing plausible maximum, medium, and minimum population growth.

Other special features of the projection model include: (1) an adjustment of the base population for net census undercoverage; (2) the use of component parameters – fertility, mortality, emigration, and internal migration - based on population estimates, which also are adjusted for net census undercoverage; (3) the use of the “Pearson Type III curve” for projecting age-specific fertility rates; (4) the projection of mortality using the Lee-Carter model (Lee and Carter, 1992); (5) the use of age-specific emigration rates to project emigration; (6) the use of the Rogers-Castro multi-regional model (Rogers and Castro, 1978) to project inter-regional age-specific out-migration rates for migration; and (7) taking the indirect effects of migration (internal and international) on births and deaths into account by surviving the "population adjusted for migration," rather than the "launch" population, as is generally done in cohort component projections.

The launch population in our example is the official set of estimates of Canada's female population by age on July 1, 2000 (first column of Table 21-3). Life expectancy at birth (e_o), used to represent the mortality component, is based on: (1) the trend and pattern of life expectancy at birth in Canada; (2) the observed and projected mortality trends and patterns in other industrialized countries; and (3) consideration of medical progress and health-related factors which are expected to affect future mortality. Three assumptions are developed in regard to future mortality; they incorporate a greater increase in male life expectancy than female life expectancy and, hence, reductions in the gap between male and female life expectancy at birth. These assumptions are shown in Figure 21-4.

The Lee-Carter model used to distribute the projected gains in e_o by age (in the form of age-specific death rates) involves the following equation:

$$l_n(m_x) = a_x + b_x k_t$$

where $l_n(m_x)$ represents the logarithm of the central death rates at age x; a_x and $b_{x,}$ age specific constants; $k_{t,}$ time.

To ensure a smooth transition from the last year of observation to the first projection year, a_x is set equal to the logarithm of the 1996 age-specific death rates (m_x) for each sex, so that when k_t equals 0, the equation produces the 1996 central death rates at each age. The b_x series determines the rate of mortality change at each age. It is set to distribute the projected gains in e_o

by age, according to the age-specific rates of change observed over the 1971-1990 period for both sexes at the national level. The k_t values are calculated to yield the exact e_o values assumed for each sex. Life table values at ages above "zero" are calculated from projected age-specific death rates. The required schedule of survivorship probabilities at different ages for each sex (e. g., S_x values for females in Col. 2 of Table 21-4) is calculated from the L_x values of the life tables for Canada. Projected survival ratios by age for females (Col. 7 of Table 21-4) are applied to the corresponding female population adjusted for migration in Col. 6 of Table 21-4 to obtain the annual number of survivors, as shown in Table 21-4. The survivors of the births (155,990 in the table) are obtained by multiplying the total number of female births during 2000 to 2001 (156,690) by the survival ratios from birth to age "under one year." The female births in this table are obtained by multiplying total births (322,274) in Table 21-3 by the proportion of female births (0.4862).

For projecting *fertility*, a Pearson Type III curve was applied to the TFRs shown in Figure 21-4 to derive projected age-specific fertility rates. This required four parameters: (1) the total fertility rate (TFR); (2) the mean age of fertility; (3) the variance of the age specific fertility rates; and (4) the skewness of the age-specific fertility rates. The first parameter provides the level of fertility, while the other three provide a measure of the timing of births or age pattern of childbearing. The application of the model rests on an analysis of each of these four parameters, and the formulation of assumptions on their future course over the projection period (Verma, et al., 1994). A comparison of actual age-specific fertility rates for Canada in 1991 and those obtained from the Pearson Type III curve is shown in Figure 21-3.

(FIGURE 21-3 ABOUT HERE)

As described in Figure 21-4, three assumptions are developed for the first two parameters (TFR and mean age of fertility), and one assumption is developed for the variance and skewness of the age-specific fertility rates. Given their small impact, values for the latter two parameters are assumed to be constant over the projection period at the level of the three-year average for 1995, 1996, and 1997.

In generating the age-specific fertility rates using the parametric model, the low fertility assumption is combined with a high value for mean age of fertility, which is assumed to increase from 28.5 in 1997 to 31.0 by 2026; and the high fertility assumption is combined with a low value for mean age of fertility, which is assumed to increase from 28.5 in 1997 to 29.0 by 2026. For the medium fertility assumption, the mean age of fertility is assumed to increase from 28.5 in 1997 to 30.0 by 2026.

Table 21-3 shows the derivation of projected births for 2000-2001. Births at each age are calculated by multiplying the female population of each childbearing age (15-44) by the corresponding fertility rates. Total births are derived by summing the values for each age so obtained. Because the projected population refers to July 1, an adjustment is required to convert calendar year births to "census year" births, i.e., July 1 of year t to June 30 of year t+1. This adjustment is done by adding half of the births of year t and half those of year t+1 (see Table 21-3). The adjusted births between July 1 and June 30, 2001 are then distributed by sex using a sex ratio at birth of 105.33 boys to 99.67 girls. Total births are multiplied by 0.5138 to obtain male births; the male births are then subtracted from total births to obtain female births. Table 21-3 shows these calculations.

[TABLE 21-3 ABOUT HERE]

Statistics Canada deals with migration at the national level by projecting immigrants and emigrants as separate components. Net immigration accounted for 76% of the total population growth in Canada in 1999-2000. The impact of this component on growth is expected to increase substantially in the coming years, even if the current below-replacement fertility level remains constant.

Two approaches have been used for projecting immigration in Statistics Canada's past projections. In the first, migration assumptions were formulated based on the analysis of past trends, focusing on recent periods. The second approach was based on annual immigration planning levels by the government. The method chosen in the projection example presented here is a combination of these two approaches (George and Perreault, 1992). Given the increasing importance of this component and the wide fluctuations in immigration (e. g., 84,000 to 250,000 immigrants per year between 1985 and 1993), three assumptions (high, medium, and low) were formulated (See Figure 21-4). The age-sex composition of the projected numbers of immigrants was derived using an assumed age-sex distribution based on the average of “stock” (census) and “flow” (immigration) data (Verma and George, 1993).

Statistics Canada decomposes emigration into three elements: emigrants; net variation in persons temporarily abroad; and returning emigrants. The total numbers of emigrants are thus obtained by subtracting returning emigrants from the sum of emigrants and the net variation in persons temporarily abroad. Total emigration is projected by applying age-sex specific emigration rates to the projected population for each year. The required emigration rates were developed by calculating annual age-sex specific rates for the years 1997-1998 and 1998-1999 and averaging them. In the single emigration assumption, these rates are kept constant from 2000 onward, as shown in Figure 21-4.

The Non-permanent resident population (NPR) is a group that forms part of the initial population in year t. It consists of the following persons and their dependants: (1) student authorization holders; (2) employment authorization holders; (3) Ministers' permit holders; and (4) refugee status claimants. The size of the NPR population is expected to remain fairly stable. It is subject to natural increase but not to migration. Hence, only the effect of NPRs in year t on fertility and mortality (natural increase) is taken into account for projection purposes without actually "projecting" them to year t+1. Following are the steps in allowing for this component. First, before the t+1 years' projected population is produced, the number of NPRs disaggregated by age and sex is subtracted from equivalent age and sex groups in the total population (Column 1 of Table 21-4) in year t. Second, births and deaths of NPRs are then calculated separately for each year and are included in the totals for these components. Third, the stock of NPRs separated from the launch year population in year t is then added to the surviving permanent population in year t+1. The process is continued for each year until the end of the projection period (see, e.g., Table 21-4). A single assumption in terms of absolute numbers is developed for this component, as is shown in Figure 21-4. The projected NPR numbers are disaggregated by age and sex using an assumed distribution.

(FIGURE 21-4 ABOUT HERE)

Table 21-4 presents the various operations involved in projecting the female population of Canada, 2000 to 2001, by the cohort-component method. As stated earlier, it includes refinements that distinguish it from the illustrative procedure shown for the cohort-component method in Figure 21-2. As an example of these deviations, births are calculated separately, as shown in Table 21-3. The "under 1 year" population in col. 1 of Table 21-4 represents female births derived from the calculation of total births, as shown in Table 21-3 (The same procedure as shown in Table 21-4 is used to produce the projected population for males in P_{t+1}). The sum of female and male populations gives the total population for both sexes together in year t+1. The same process is continued for projecting the male, female and total population for each year until the end of the projection period (2026).

[TABLE 21-4 ABOUT HERE]

Example of a Subnational Projection

In the cohort-component method, the main difference between national and subnational projections is the addition of the component of internal migration. Although an assumption that future international migration will be negligible can be justified for many countries, internal migration plays a significant role in almost every country and at the subnational level is often the most important and complex component of population change. The example provided here is for the province of Ontario. The basic methodology used is the same as that used at the national level. As stated earlier, provincial projections of mortality, fertility, immigration, and emigration are tied to the national projections of these components.

The provincial assumptions of life expectancy at birth (e_o) are derived from the three national assumptions (Figure 21-4) by applying the 1995 and 1996 average provincial/national e_o ratios. The differences in e_o from one province to another are assumed to continue during the projection period. For example, the female provincial/national e_o ratio for Ontario was 1.00. This ratio was applied to the projected life expectancy value for Canada in 2026 to obtain 84.0, the life expectancy at birth projected for Ontario in 2026. The same approach with provincial/national ratios is used for other components to derive the corresponding provincial values from the values at the national level. The rest of the calculations involved in deriving the survival ratios shown in Table 21-5 are the same as for mortality projections at the national level.

The assumptions of fertility for Ontario are derived from the national assumptions as shown in Figure 21-4. In using the ratio method (as illustrated for mortality), average provincial/national ratios were calculated for the three most recent years and consideration was given to the extent to which Ontario (and each of the other regions) is "catching up" with the national fertility level. The calculation of fertility rates and births is made using the parametric approach described for Canada as a whole and as illustrated in Table 21-3.

With respect to immigration for provinces, the three assumed numbers at the national level were first distributed by province based on the basis of the average distribution of immigrants for each province for the most recent years (1997-1999). The provincial totals were then distributed by age and sex based on the basis of an assumed age-sex distribution. Emigration was projected by applying age-specific emigration rates to the projected population at risk for each province. The provincial emigration rates were derived from the single assumption of emigration at the national level. With regard to the non-permanent residents (NPR), the assumed number at the national level was distributed by province according to an average province/Canada ratio based on the distribution for the most recent years (see Col. 5 of Table 21-5).

The projections of internal migration for provinces are based on a multiregional migration model (as illustrated in Table 21-5). The application of this model at the provincial level requires detailed migration data as follows: (1) origin-destination-specific migration streams disaggregated by age and sex for each province at 1-year migration intervals for a substantial time interval; and (2) the corresponding base population to compute out-migration rates. Statistics Canada produces estimates of inter-provincial migration using administrative data files from three sources: Revenue Canada income tax files, Family Allowance files before 1993, and Child Tax Benefit Program files (which replaced Family Allowance files) since 1993 (Statistics Canada, 2002). The migration estimates are available (with age and sex detail) on an annual basis for each year since 1966-1967.

The application of a multiregional migration model requires projected age-sex specific out-migration rates and origin-destination proportions. The method has four basic steps. First, projected crude out-migration rates and origin-destination proportions are developed according to a selected migration scenario. Second, corresponding age-sex specific rates are derived from the extrapolated crude out-migration rates using the Rogers-Castro parametric model (Rogers and Castro, 1978; Bélanger, 1992). Third, these age-specific out-migration rates are applied to the corresponding provincial population to yield out-migrants by age and sex. Fourth, these out-migrants are distributed as in-migrants to other provincial destinations using the projected origin-destination proportions. (In this last step it is assumed that the destination proportions do not vary by age or sex.) The assumed rates and proportions are then assessed in terms of the reasonableness and acceptability of the resulting levels of net migration, taking account of local expertise and expert judgment. The application of the projected rates and proportions is illustrated by the following equations:

$$M_{xi} = (m_{xi})(P_{xi})$$

where:

M_{xi} = the total number of annual out-migrants from origin i by age and sex:

P_{xi} = the population of age and sex, x, at origin i; and

M_{xi} = the annual out-migration rates of persons by age and sex, x, from origin i.

The out-migrants from each area of origin is distributed by area of destination on the basis of in-migration proportions:

$$M_{xij} = (M_{xi})(P_{ij})$$

where:

M_{xij} = the number of annual out-migrants by age and sex, x, moving from area i to area j (origin-destination flows);

M_{xi} = the number of annual out-migrants by age and sex, x, from area i; and

P_{ij} = origin-destination proportions, from area i to area j where $\sum P_{ij} = 1$ for any i.

The in-migrants (I) and out-migrants (O) and net migrants (NM= I-O) are aggregated from the origin-destination flows for each region.

Three scenarios (assumptions) are developed to provide a range of net-migration for each province: “west;” “central;” and “medium.” The “west” scenario is based on the migrant data for the years 1992-1993; the “central” scenario is based on the migrant data for the years 1984 – 1987; and the “medium” scenario is the average of the “west” and “central”. The west scenario is considered relatively favorable for a certain group of provinces, while the central scenario is relatively favorable for the remaining provinces. The net-migration figures for Ontario presented in Col. 5 of Table 21-5 are taken from the medium scenario.

The multi-regional model described here is a sophisticated (and complex) method for projecting internal migration by age and sex (for further details, see Statistics Canada, 2001). Apart from the complexity of the method in terms of the data required and the projection process involved, the most cumbersome step is to obtain projected net-migration figures consistent with provincial inputs based on local knowledge or expert judgement. One way to simplify the process may be to automate the implementation of the net migration targets. This could hasten the crucial adjustment process required to improve the quality and acceptability of projection results.

[TABLE 21-5 ABOUT HERE]

Table 21-5 illustrates the various operations involved in projecting the female population of Ontario according to the "medium" assumption for one year, 2000-2001. The steps followed are the same as shown in Table 21-4, the only difference being the additional column of net internal migration (col. 5). (The same process is continued for each year to 2026.)

Summary Comments on the Cohort Component Method

The Cohort component method is widely used, relatively easy to explain, and practical. It permits the use of already available data and existing theoretical knowledge on the dynamics of population growth, and takes into account causal factors, at least at the level of basic components and compositional factors. It has the capability to produce consistent and comparable national and subnational projections that are easy to update on a regular basis. Much of the work required to use this method lies in the in-depth analysis and development of assumptions for each of the components of change. The cohort component method also has its shortcomings and limitations. One is that it does not explicitly incorporate socio-economic determinants of population change. For dealing with this issue, we now turn to a discussion of structural modeling.

STRUCTURAL MODELS

Demographers and others often face questions that cannot be answered using projection methods based solely on demographic factors - the demographic consequences of the closing of a large manufacturing plant, for example. Structural models come into play here because population projections developed by this method can account for factors including the economy, environment, land use, housing, and the transportation system. We describe two general categories of structural models—economic-demographic models and urban systems models. Economic-demographic models are typically used to project population and economic activities for larger geographic areas such as counties, labor market areas, states, and nations. Urban systems models focus on small geographic areas such as census tracts and block and typically include projections of population, economic activities, land use, and transportation patterns. In addition to their differences in geographic scale, these two types of models often provide alternative explanations of the causes and consequences of population change. Some structural models contain only a few equations and variables (Mills and Lubuele, 1995), while others contain huge systems of simultaneous equations with many variables and parameters (Data Resources Incorporated, 1998; Waddell, 2000). Our objective is to provide a general introduction and overview of the use of structural models for population projection. We do not provide details for building or implementing these kinds of models; such details can be found in Putman (1991), San Diego Association of Governments (1998, 1999), and Treyz (1993).

Economic-Demographic Models

Economic-demographic models sometimes focus on the total population, but most often deal with one or more of the components of population change. Only a few applications have dealt with fertility and mortality; typically, these applications have focused on nations or regions of the world where fertility and mortality are the most important contributors to population growth (Ahlburg, 1999). Fertility and mortality models have also been proposed for subnational areas, but have rarely been implemented (Isserman, 1985). For population projections, internal migration has been the predominant concern of economic-demographic models; consequently, we confine our discussion to models for migration and total population.

Virtually all economic-demographic models of migration are based on a premise set forth more than a century ago that people move principally "to 'better' themselves in material respects" (Ravenstein, 1889: 286). Economic factors such as job change, unemployment, and wages or income are therefore used to project migration or population. The empirical evidence clearly shows that the strongest links are found with job change rather than other economic factors (Isserman et al., 1985; Kreig and Bohara, 1999). The fact that jobs attract people and people create jobs underlies most economic-demographic models in use today. Migration and population change are also influenced by non-economic factors such as climate, coastal location, life cycle changes, personal characteristics, and social networks (Astone and McLanahan, 1994; DaVanzo and Morrison, 1978; Fuguitt and Brown, 1990; Massey et al, 1987). A complete migration model including both economic and non-economic factors, however, is problematic for projecting migration or population because the independent variables themselves must be projected. Projections of these non-economic variables are rarely available, while projections of economic variables can be obtained from national, state or county-level economic models.

We describe three general approaches for designing and implementing economic-demographic models: (1) Econometric models, which use regression methods to project migration as a statistical function of the economy; (2) balancing models, which project migration as the difference between the projected supply and demand for labor; and (3) ratio-based models, which typically derive population projections directly from employment projections.

Econometric Models. The econometric approach uses equations that determine migration from one or more economic variables. Parameters for these equations are estimated from historical data using regression techniques. Projections are then made by solving the equation(s) using the projected values of the independent variable(s). The migration equation(s) are typically integrated into a large economic model that also provides projections of the economic factors.

The most widely used econometric models of migration are "recursive," whereby migration is influenced by the economy, but does not itself influence the economy. Recursive models cannot reflect the full range of interactions between migration and the economy, but nonetheless have proven successful for projecting migration (Clark and Hunter, 1992; Greenwood and Hunt, 1991; San Diego Association of Governments, 1999; Greenwood, 1975; Tabuchi, 1985). Recursive relationships have also been implemented in multiregional migration models (Campbell, 1996; Foot and Milne, 1989, Isserman, et al., 1985, Rogers and Williams, 1986). Non-recursive models attempt to capture the joint impacts of migration and the economy on each other. Although they are more complicated and require larger resources than recursive models, "non-recursive" models for projecting migration have been occasionally employed (Conway, 1990; Mills and Lubuele, 1995; Treyz et al., 1993).

Balancing Model. The concept behind the balancing model is simple. If labor supply exceeds labor demand, workers migrate out of the area; if labor demand exceeds labor supply, workers migrate into the area. Balancing models are typically less costly to implement and easier to use than econometric models because they do not require large-scale systems of equations, huge amounts of data, or the use of formal statistical procedures. However, they do require numerous computations and assumptions (see Murdock and Ellis, 1991, for an example). Labor demand is often represented by a measure of job opportunities typically projected using export-base models, input-output models, and extrapolation techniques (Greenberg, et al., 1978; Murdock et al., 1984). Labor supply is determined by applying labor force participation rates to a projected population derived from a cohort-component model that assumes zero net migration. The migration of workers is determined by the difference between projected labor supply and projected labor demand. As a final step, the migration of workers is transposed into a projection of all economic migrants, including other family members, through assumptions related to characteristics such as marital status and family size.

Population/Employment Ratio. The population/ employment (P/E) model projects total population directly; it does not consider any single component of change. Despite some drawbacks, the P/E model is the easiest and least expensive way to incorporate economic factors into a population projection. The simplest P/E model uses a single ratio representing total population to total employment, holds the ratio constant at its current value, and applies the ratio to a projection of employment. This approach is no longer used very often because P/E ratios are known to change over time and vary according to demographic subgroup (Murdock and Ellis, 1991).

For many years, the "OBERS" model, developed by the U.S. Bureau of Economic Analysis (BEA) in the mid-1960s, was arguably the most widely used P/E model. Population, employment, and earnings projections for states and metropolitan areas were developed from this model until the mid-1990s (U.S. Bureau of Economic Analysis, 1995), when budget cutbacks forced the BEA to stop preparing projections. The approach taken in OBERS divides the population into three age groups: pre-labor pool (less than 18), labor pool (18-64), and post-labor pool (65+). Projections of the labor pool population are directly related to changes in employment and the pre-labor pool population projections are tied directly to the projections of the labor pool population. Post-labor pool projections are independent of economic changes. A numerical example of the OBERS approach is found in Smith et al. (2001).

Urban Systems Models

Urban systems models are used throughout the world to project the distribution of residential and nonresidential activities within urban or metropolitan areas. They differ in several important ways from economic-demographic models. First, they are designed to be used for much smaller geographic areas. Second, they use different independent variables. Along with economic factors such as jobs and income, urban systems models include land use characteristics (e.g., zoning, environmental constraints, land value and land supply) and characteristics of the transportation system (e.g., travel times, cost, and distances). Third, they use geographic information system (GIS) technology, which plays an important, perhaps an essential, role in urban systems models (for a general discussion of GIS, see Appendix D). Fourth, urban systems models require considerably more information, time, and resources to implement than economic-demographic models. Finally, urban systems models address many issues (e.g., air quality, traffic congestion, loss of open space, and public transportation) that cannot be considered in most economic-demographic models.

Urban systems models vary considerably in their theoretical approaches, mathematical design, data requirements, and ease of implementation, but they typically consist of three major components—regional projections, land use and activity, and transportation. They are usually applied using five-year time intervals (see Figure 21-5). Population and economic projections are required for the region covered by the model (e.g., metropolitan or labor market area). These regional projections are often produced using the economic-demographic models previously discussed. The land use and activity component consists of a complex set of procedures for distributing the regional projections into zones within the region. Applications typically involve between 150 and 300 zones. These zones often comprise one or more census tracts, but land use and activity models have been developed for smaller geographic areas such as census blocks, gridcells, and assessors' parcels (San Diego Association of Governments, 1998; and Waddell, 2000). The transportation component projects characteristics of the transportation system such as traffic volumes and speeds on roadways and on public transportation lines.

(FIGURE 21-5 ABOUT HERE)

A fundamental characteristic of urban systems models is the iterative and explicit relationships between land use characteristics, activity location, and the transportation system as shown in Figure 21-5. The distribution of population in virtually all such models relies on the link between home (residential location) and workplace (employment location). These links are represented by travel probabilities between zones based on time, distance, or cost and commuting patterns (Putman, 1991). Residential location influences the spatial distribution of employment, particularly employment that serves a local population such as retail trade and services. As Figure 21-5 indicates, this relationship is implemented by assuming a lag between residential location and location of employment.

The transportation system both influences and land use characteristics play an important role in determining the location of population and other activities; and urban system models contain procedures to reconcile the demand for land with its available supply (San Diego Association of Governments, 1998; and Waddell, 2000).

Comments on Structural Models

Structural models - especially urban systems models - require more resources and are more difficult to implement than the other models discussed in this chapter. They often require extensive base data, sophisticated modeling skills, and complex statistical procedures and computer programs. Therefore, they are accessible only to a relatively narrow range of practitioners, although the Transportation, Economic, and Land Use System (TELUS) may help reduce the barriers to implementing this system (Pignataro and Epling, 2000). In addition, there is no evidence to suggest that structural models provide more accurate population forecasts than other methods and, given their small geographic scale, their forecast accuracy is not likely to be high in many applications. Yet, structural models are used more frequently today than ever before because of their ability to investigate and analyze a wide range of theoretical, planning, and policy questions (Boyce, 1988; Tayman, 1996b; Treyz, 1995). Decision making and planning often require the analysis of many interrelated factors for different geographic areas. For example, planners and policy makers may be required to meet the challenges posed by increasing traffic congestion, housing shortages, and deteriorating infrastructure. Structural models can make important contributions to the planning and decision-making process; they can, for example, provide warnings when proposed actions might lead to unintended or undesirable consequences (Schmidt, Barr, and Swanson, 1997; Tayman, 1996b). In some circumstances, preparing simulations and scenarios is more valuable than any specific projection or forecast.

RELATED PROJECTIONS

Projections of households, school enrollment, poverty, employment, health and other population-related characteristics are needed for many types of planning, budgeting, and analysis.

For simplicity, we refer to these as *socio-economic* projections. Because of the demand for socioeconomic projections and their close link to projections of basic demographic characteristics, it is not surprising that the former are often made on the basis of the latter and that public and private sector organizations around the world are involved in the production of socio-economic projections (CACI, 2000; Fullerton, 1999; Kintner et al., 1994; Siegel, 2002: 508-510; Snider, 1996; Tayman, 1996b).

Much of the previous discussion of population projections can be applied to socio-economic projections as well - terminology, data sources, methods, and evaluation criteria. The projection of socio-economic characteristics, however, has two important features that distinguish it from strictly demographic projections.

The first is that some socio-economic characteristics are directly affected by policy decisions (Opitz and Nelson, 1996). For example, projections of university enrollment are affected by changes in university entrance requirements, projections of prison populations are affected by changes in sentencing guidelines, and projections of housing demand are affected by changes in eligibility requirements for home mortgages. In some instances, then, knowledge regarding the details of public policy is essential to the production of projections of socio-economic characteristics.

The second is that projections of socio-economic characteristics involve *achieved* characteristics - those that can change over one's lifetime, such as marital status, income, educational attainment, occupation. As a result, projections of socio-economic characteristics involve a variety of assumptions in addition to those for projections of strictly demographic characteristics. At high levels of aggregation, achieved characteristics are often related to ascribed characteristics (those that are set at birth, such as age and sex) in clearly identified patterns. For example, school enrollment is closely linked to the age structure of the population. These patterns form a basis for projecting socio-economic characteristics.

Two fundamental approaches are frequently used to prepare such projections. The first approach is the "participation ratio method" (also known as the "participation rate method," "prevalence ratio method," and "incidence rate method"). In this approach, socio-economic characteristics are related to demographic characteristics through the use of ratios (Siegel, 2002: 509-511; Swanson and Klopfenstein, 1987; United Nations, 1999). Once such ratios are established, they can be projected in a number of ways, such as holding them constant at recent levels, extrapolating recent trends, tying them to ratios found in other places, or developing structural models that forecast changes in them. The second approach is the cohort-progression method. In this approach, projections are developed by "surviving" people with particular socio-economic characteristics. The Hamilton-Perry Method discussed earlier in this chapter is an example of this approach. Because these two approaches are used so frequently for projections of socio-economic characteristics, we discuss them in some detail.

It should be noted that virtually all socio-economic projections can be handled using structural models. In particular, there are many structural models designed to deal with economic activity, including employment according to industry and occupation, income, and other variables (Data Resources Incorporated, 1998; Treyz, 1993, 1995). In the case of the "REMI" model developed by Data Resources Incorporated (1998), detailed user manuals show how one can "call" for particular projection outputs such as the total employed according to age and sex. It is worth noting that many structural models such as REMI explicitly accommodate user judgment in the development of the projections.

Participation Ratio Method- In this approach, current and historical data are used to construct participation ratios – that is, proportions of the population (stratified by age, sex, and perhaps other demographic characteristics) that have the socio-economic characteristic of interest. These ratios are projected into the future using one or more of the techniques described previously. The projected ratios are then applied to population projections (stratified by age, sex, and other characteristics) for the geographic area(s) under consideration to obtain a set of socio-economic projections. The population projection must have sufficient demographic detail to match up conceptually and empirically with the denominator originally used to construct the participation ratio of interest.

The steps used in this approach can be summarized as follows:

(1) Launch Year Participation Ratio = P^{c}_{dt}/P_{dt}

(2) Projected Participation Ratio = (P^{c}_{dt+i}/P_{dt+i})

(3) Independently Projected Population = $(\mathbf{P_{dt+i}})$

(4) Projected Population with the Characteristic = $P^{c}_{dt+i} = (P^{c}_{dt+i}/P_{dt+i}) * (\mathbf{P_{dt+i}})$

where

P = population

c = socio-economic characteristic (e.g., number employed)

d = demographic data (e.g., age-sex)

t = launch date

t+ i = target date

As an example of this method, we show part of the results from a projection of "days of hospital care" prepared in the late 1980s by Kintner and Swanson (1994: 285), of male retirees of the General Motors Corporation. The projected number of male retirees aged 75-79 for 1990 was 5,719 (using the cohort-component approach). Multiplying this number by 4,263.7 "days of care" per 1,000 males in this age group (a value taken from the 1987 National Hospital Discharge Survey), Kintner and Swanson obtained a projected value of 24,384.1 days of care for this group (24,384.1 ≈ 5,719* 4.2637).

Cohort Progression Method- In this approach, participation ratios are constructed in the same manner as in the participation ratio method, but they are projected into the future on a cohort basis using information on changes in those ratios between two previous dates. The conventional form of this method uses ratios of the number of persons aged *a* with a particular socio-economic characteristic in year *t* to the number of persons aged *a-y* with that characteristic in year *t-y*. The initial projections are made by applying these ratios to the number of persons with the characteristic of interest in the launch year. This method can be represented as follows:

(1) Initial cohort progression ratio = $P^{c}_{d,a,t} / P^{c}_{d,a-y,t-y}$

(2) Projected cohort = $P^{c}_{d,a+y,t+y} = (P^{c}_{d,a,t} / P^{c}_{d,a-y,t-y}) * (P^{c}_{d,a,t})$

where the symbols have the same meaning as above.

The cohort progression method is applied recursively, as is done in any survivorship exercise. It is important to remember that cohort progression ratios represent net cohort change rather than gross change. This distinction is important because fundamental patterns may be masked without knowing the numbers "entering and exiting" a population (Fullerton, 1999). It also is worthwhile noting that the cohort progression method is used less often than the rate progression method, which employs absolute numbers rather than ratios. As an example of this method, consider a projection for 2003 of the number of persons aged 19 who are expected to have a driver's license in Fredonia. In 2001, 40,437 of the persons aged 18 held a driver's license and 41,073 of this same cohort (now 19 years of age) held one in 2002. The ratio of 2002 to 2001 is 1.01573 = 41073/40437. Multiplying this ratio by the number of 18 year-olds with a driver's license in 2002 (40,200) yields a projection for 2003 of 40,832 persons aged 19 who are expected to hold a driver's license (40,832 ≈ 1.01573*40,200).

Projecting Households and Families

Projections of households and families are required for many uses, particularly those that depend on information regarding future numbers of consumer units. For many goods and services, households and families are more effective units of demand than the individual because they are the basic units into which people are organized for purposes of consumption. In addition to demographic factors, the number of households depends on conditions affecting the supply and cost of housing, family income, and cultural norms regarding, among other things, the number of generations living together. Thus, procedures for projecting the number of households and families may vary according to whether they represent: (1) extensions of past trends; or (2) embodiments of various norms relating to the size and composition of households and assumptions regarding the supply and cost of housing, family income, and other such factors. Short of developing a structural model, assessment of potential changes in non-demographic factors over time will be important in choosing the assumptions for future participation ratios or cohort progression ratios.

Participation Ratio Method- A refined participation ratio would take into account demographic characteristics associated with household "headship." The general procedure consists of applying to the population, projected by age and sex, various estimating ratios which are related to marital, household, and family status, such as the proportion of the population in each marital category and the proportion of each age denominated as "household heads" ("headship ratios"). This method is often referred to as the "Headship Rate" Method. In it, the total number of households is derived by summing the number of heads obtained by age and sex. In a still more elaborate form of the participation ratio method, data on marital status, by age and sex, and data on various categories of family and household status, by age and sex of head, are combined, and the pertinent age-specific proportions and rates are projected or held constant as seems appropriate, to provide projections of married couples and of households and families by type. Depending on the data available and the procedure employed, one could obtain, for example, the number of households by age and sex of head; the number of households headed by families, by type of family (e.g., husband-wife, other male head, female head); the number of households headed by individuals, by sex; the number of (secondary) families who live with other (primary) families in the same household; the number of married couples and nuclear families, by whether they are living in their own households or in the households of others. In summary, probably the single most useful way to project households is to combine the projection of participation ratios with the results of a fully elaborated cohort-component population projection. This takes into account both changes in headship ratios and population composition.

Because of its operational simplicity, both in terms of the input required and the design of the model, and perhaps its reasonably good performance, the "headship ratio" approach has been adopted by the United Nations and several countries since it was first used in the United States as early as 1938 (United Nations, 1973: 31). It can take into account the latest population projections as its base, specifically, the projected adult population disaggregated by age-sex.

Cohort Progression Method - We have already alluded to the use of the cohort approach in projecting the percent ages in each marital category when making projections of marital status. The cohort approach may be extended to include the projection of the percentages of heads, when data on heads of households by age are available for a set of time points. This general procedure is directly applicable to series like percent married or percent heads, which are essentially cumulative by age.

Projecting School Enrollment

Projections of the number of children who will be enrolled in school are needed to formulate educational policies and plan educational programs and, specifically, to plan for needed schools, classrooms, and teachers. Because almost all children in the compulsory attendance ages for elementary school attend school, projections of the total population of elementary school age, in relation to the number expected to attend, are also useful in determining needs. In addition, projections of school enrollment ratios can be used in preparing projections of labor force participation ratios because the two sets of ratios are inversely related, especially at certain ages.

Projections of the educational attainment of a population are also needed for national planning. The present and prospective educational attainment of a population influences its development. Social and economic relationships in a population change in many ways as its educational level changes.

As was the case with the earlier discussion on household projections, this discussion covers the two principal methods, the participation ratio method and the cohort progression method. A simple component method can also be employed (i.e., allowing for new entrants, dropouts, deaths), but is not usually practical.

Participation Ratio Method - One way to project enrollment is to develop age-specific enrollment ratios (i.e., proportions of the population enrolled in school at each age) in combination with the projected population by age (Swanson, et al., 1998). The assumption relating to future age-specific enrollment ratios may be quite simple One may assume, for example, that current age-specific enrollment ratios will continue into the future. One could also use one of the extrapolation techniques to project age-specific enrollment ratios. Past trends in the ratios may be assumed to continue as observed or to continue in a modified fashion. In those instances where the enrollment ratios approach 100 percent and extrapolation using a growth rate could cause these enrollment ratios to exceed 100, the complements of the enrollment ratios, may be projected instead. Legal requirements and practices with respect to the ages or grades of school attendance have ordinarily been incorporated into projections of school-age population and school enrollment. The possibility of developing projections of enrollment which incorporate new norms representing more inclusive age and grade spans, and even hypothetical enrollment ratios for the base year, needs to be considered if provision is to be made in the future for upgrading the current level of school services.

When projections of school enrollment are made by the participation ratio method, they rarely contain detail on grade or school level. To obtain such figures, one procedure is to prepare projections by age first, and then to distribute the projected total enrollment at each age by grade or school level on the basis of recent census, survey, or administrative data (holding the distribution constant or extrapolating it). Alternatively, projections of enrollment for broad school levels at each age may be calculated by the use of age-level enrollment rates (i.e., in relation to total population at each age) and total enrollment at each age may then be derived by summation.

Cohort Progression Method - This method is particularly useful for providing separate information on entries into and withdrawals from school by age. Here, one begins with a distribution of persons enrolled by age and carries this population forward by use of age-specific rates of net school accession and net school withdrawal. A set of net accession rates and net withdrawal rates may be derived by taking the relative difference between enrollment ratio at successive ages given in a census. A net accession rate is obtained for the ages where enrollment ratio are increasing and a net withdrawal rate is obtained where enrollment ratios are decreasing. Rates of net accession are applied to the population not-enrolled to derive new enrollees, who are added to the enrolled population, and rates of net withdrawal are applied to the enrolled population to derive dropouts and deaths, who are removed from the enrolled population.

The cohort progression method may also be used to develop enrollment projections for grades. In this procedure the number of enrolled persons classified by grade is carried forward to each subsequent calendar year by use of projected grade-retention rates or grade-progression rates, representing the proportion of children in a given grade who will advance to the next grade in the course of a year. A historical series of grade-retention ratios may be developed on the basis of survey data or data from the administrative records of the school system and then projected forward on an annual basis.

As with age-specific enrollment ratios, in projecting grade-retention rates, the regular methods of extrapolation can be used, but if the rates are very high (i.e., near 100 percent) and the trend is one of rapid increase, conventional methods can produce values in excess of 100 percent. In this case an asymptotic equation which prevents getting a projected value of 100 percent or more can be used, or the complements of the grade retention rates (i.e., grade-dropout rates) can be extrapolated on a geometric basis.

Projecting the Labor Force

Labor force projections are needed to indicate the number of jobs which the economy must make available and as a point of departure in making projections both of the economy as a whole (Fullerton, 1999) and of regional economies (Treyz, 1993, 1995). Labor force projections must be matched by projections of manpower requirements for effective national economic planning, however (Thompson, 1999). It is worthwhile noting that some labor force projections have tended to assume that labor force participation ratios are independent of the general state of the economy, although in fact the number of persons seeking work does depend on the demand for labor, the availability of jobs, and on the need for a job within the family (Rosenthal, 1999).

Participation Ratio Method - A simple age-specific ratio method may be employed: proportions of the population in the labor force by age and sex (i.e., labor force participation ratios or economic activity ratios) are assumed for future dates and applied to projections of the population of working age (e.g., 14 years and over) disaggregated by age and sex. In the simplest form of the method, these participation ratios are held constant at the level observed in the last census or some recent survey. Such projections take account only of expected shifts in the future size, age, and sex composition of the population. For many ages of males, say from 25 to 54, the best assumption may be to use current participation ratios because nearly all males are in the labor force and likely to remain so. The projections may, however, allow explicitly for expected changes in participation ratios. The difficult problem is to determine how the ratios might change for males at the fringe ages of economic activity, for females, and for race and ethnic groups that historically have lagged in their labor force participation. Accordingly, it is useful to take into account the relationship between worker ratios and various socioeconomic variables affecting labor force participation. Because the variables of marital status and the presence or absence of young children define quite different levels of economic activity for female workers, it is especially useful to consider ratios specific for these variables separately in developing projections for women. Average age at marriage and of childbearing and number of children previously born may also be taken into account directly or indirectly. Consideration may be given to the availability of public services and facilities, such as day-care centers and nurseries, which could affect the participation ratios for women with children.

Cohort Progression Method – A typical form of the cohort progression method consists of carrying forward the economically active population by age and sex to future dates by use of probabilities of net entry and probabilities of net withdrawal through death or retirement. The probabilities of net entry are applied to the inactive population to determine accessions for the first year, and the probabilities of net withdrawal, separately for retirement and death if possible, are applied to the active population to determine separations for the year. The estimated new entrants are then subtracted from the inactive population and added to the active population. Similarly, separations due to retirement are subtracted from the active population and added to the inactive population (see, e.g., Kintner and Swanson, 1994). Rates of net entry and of net withdrawal due to retirement may be derived from the relative change in activity rates at consecutive ages as given by census data. In the current data for the United States, until about age 35, these changes are positive and the rates are considered rates of net entry. At the older ages, the changes are negative and the rates are considered rates of net withdrawal. The rates can then be adjusted to exclude the effect of mortality. This variation of the cohort method has the virtue of providing, as valuable by-products, the annual number of net entrants into, and net withdrawals from, the labor force, disaggregated by age. Such information is of specific use in national economic planning, both for the utilization of new workers and for management of retirement and other programs for older ages.

Projecting Health Characteristics

Health and health care issues represent a major budget and planning issue for many organizations (Kintner, 1989, Kintner and Swanson, 1994, 1996; Pol and Thomas, 2001). It has been estimated, for example, that General Motors Corporation spends 30 percent of its annual budget on health care for its employees and their dependents (Kintner and Swanson, 1996). With huge amounts of resources at stake, it is little wonder that the projection of health characteristics is of wide interest. However, health characteristics are also of interest because of issues beyond budgets and planning. For example, by 2015, it has been projected that the population of 29 African countries will be 8.1 percent lower due to the impact of HIV/AIDS (UN, 1998: 31).

Participation Ratio Method- As described earlier in the beginning of this section, Kintner and Swanson (1994) used this approach in combination with the cohort-component method in projecting hospitalization levels for a retiree cohort. In the first step, hospital utilization ratios according to age and sex were developed from national surveys in the United States. These ratios were kept constant over the projection horizon, but they could have been modified using one of the trend extrapolation methods described earlier this chapter. Next, the numbers of retirees was found by age and sex using a cohort-component approach. In the final step, the projected numbers in each age-sex group were multiplied by the corresponding hospital utilization ratios and specific cohort utilization levels were found annually by taking a weighted sum of the products, were the weights are the number of surviving retirees in a given age-sex groups for a given year.

Cohort Progression Method - The projection of HIV/AIDS done by the UN represents a variation on the cohort progression method (UN, 1998). In the first step, models are used to estimate the annual incidence of new infections. Second, estimates of deaths due to AIDS are estimated using assumptions about the progression from HIV infection to AIDS and from AIDS to death. Third, these deaths are added to deaths expected in the absence of AIDS and revised life tables are calculated. Finally, the revised life tables are used in a cohort-component projection.

ADDITIONAL CONSIDERATIONS

The preceding sections of this chapter covered the “nuts and bolts” of producing population projections. Knowledge of these methods and materials is essential to the projection process, but does not resolve all the issues related to constructing and evaluating population projections and using them for planning and analysis. In order to maximize the usefulness of population and related projections, a number of additional issues must be considered. Among the most important are the following.

Providing Necessary Detail refers to customizing projections made for a specific use or data user, so as to fit exactly the purposes for which they will be used (e.g., annual projections covering a 10-year horizon of the number of children living in a school district, by single years of age). In many instances, however, projections are made without reference to a particular use or data user. For these "general purpose" projections, it is much more difficult to determine which geographic areas, target years, and characteristics to include.

The needs of the largest number of data users can be met by making projections for a wide variety of geographic areas, characteristics, and target years. Using these building blocks, data users can put together projections that cover the specific areas, characteristics, and time periods they need. The greater the amount of detail produced, however, the greater the amount of time and resources needed to construct the projections. Consequently, the producers of population and related projections typically provide only a limited amount of detail.

National government agencies generally make projections at the national level and for major subnational units such as states, provinces, or departments, often classified by age, and sex. Related projections may include households, the labor force, and poverty status. National-level projections often extend 50 or even 100 years into the future, while projections for subnational areas typically cover shorter horizons (e.g., 20-30 years). The level of detail included in national projections is often determined by the statistical needs of national government agencies.

Subnational (e.g., state, provincial, municipal) governmental agencies (or their designees) often make projections for states/provinces and smaller geographic areas (e.g., counties, cities, census tracts). The amount of socioeconomic and demographic detail included in these projections varies tremendously from one place to another; again, it is frequently determined by the statistical needs of government agencies. Some private companies make highly detailed projections for very small areas (e.g., block groups), but typically cover very short horizons (e.g., five years).

Face validity is the extent to which a projection is based on appropriate methods, incorporates high-quality data, uses reasonable assumptions, and accounts for relevant factors. The appropriateness of a method depends primarily on the purposes for which a projection will be used and the type of data available. Many methods are appropriate for projections of total population, but projections of age groups usually require some type of cohort approach and projections of economic-demographic interactions require a structural model. Other types of projections require data and techniques specific to the nature of those projections (e.g., labor force participation ratios, school enrollment ratios). Data quality is determined by the length of the data series as well as its completeness, reliability, and timeliness.

Although "reasonableness" (of assumptions) is a subjective concept, assumptions can be judged according to the extent to which they fit current conditions, relevant theory, and changes in factors known to affect population change. These factors include population structure, socioeconomic characteristics, and mortality, fertility, and migration. For small areas, other factors may play an important role as well: the size of the area, constraints on growth (e.g., flood plains, environmentally protected areas), location (e.g., distances from major employers and shopping centers), transportation characteristics (e.g., access to highways and railways), land-use policies (e.g., zoning and regulatory restrictions), and special populations, such as persons residing in prisons, college dormitories, and military barracks (Murdock et al., 1991). Information on factors like these has been called "domain knowledge" (Ahlburg, 2001; Armstrong, 2001: 778).

Plausibility is the extent to which a projection is consistent with historical trends, the assumptions inherent in the model, and projections for other areas. Plausibility is closely related to face validity, but focuses on the outcomes of the projection process rather than the inputs. If a projection is not based on valid data, appropriate methods, and reasonable assumptions, it is not likely to provide plausible results.

Like face validity, plausibility is a subjective concept but can be tested using a variety of internal and external evaluations. Internal evaluations address questions like: Are the projected trends consistent with those observed in the past, prevailing demographic conditions, and demographic theories? Consistency tests may be conducted by examining selected age groups (e.g., less than one year, ages of school attendance, labor force, retirement ages) and comparing projected demographic indices (e.g., growth rates, survival ratios, birth rates) with those observed over the past few years. External evaluations compare projections with those produced for similar areas or those produced in other countries at a similar stage of development.

Production Cost is an important consideration. Labor is the primary factor of production for most types of projections, so that the cost of labor is the primary cost of production. A great deal of time must be spent considering assumptions and relevant details; collecting, verifying, correcting, and adjusting input data; putting together projection models; and evaluating the plausibility of projection results. Other costs (e.g., computer hardware and software, purchases of proprietary data) are typically small in comparison. Costs increase with the level of methodological complexity and analytical sophistication required; the amount of socioeconomic and demographic detail included; the number of geographic units covered; and the extent to which domain knowledge is incorporated in the methodology. In many instances, trade-offs will have to be made between the scope of projections and cost.

Timeliness has several dimensions. One refers to the release of input data. Some data are available only once per decade, others are available annually or even monthly. Some are available shortly after their reference dates, others only after a lag of several years. The more frequently and quickly data become available, the greater their usefulness for producing projections. Another dimension is the amount of time needed to construct a set of projections. Other things being equal, the more quickly projections become available to the data user, the greater their potential usefulness. A third dimension of timeliness is the frequency with which projections are updated. Because shifting trends often reflect short-run deviations rather than fundamental long-run changes, the availability of recent projections may be more important when dealing with short horizons than long horizons. There is no uniform practice among national agencies: Some update their projections annually, some every other year, and some at irregular intervals. A survey of 30 industrialized countries in 1988 showed that 15 countries updated their population projections only at intervals of four years or longer (Cruijsen and Keilman, 1992: 23). In the United States, national projections are updated roughly once every three years; in the United Kingdom, every two years; and in Australia, every four years. Statistics Canada revises long-run projections for Canada and its provinces and territories following every quinquennial census and prepares short-run (five-year) updates every year (George, 2001).

Ease of application is determined by the amount of time and the level of expertise needed to collect, verify, and adjust input data, develop a projection model, and generate the desired projections. This issue is particularly important for those with limited training or expertise in constructing projections or who face severe time or budget constraints. Several computer software packages are available for applying the cohort-component method at the national level (Bongaarts and Bulatao, 2000: Appendix A), but there are few (if any) software packages that incorporate a variety of methods and account for the unique features of population growth and demographic change in small areas. This concept also refers to the extent to which data sources, assumptions, and projection techniques can be described clearly to data users. Some data users are interested only in the projections themselves, not in how they were produced; for them, this issue is irrelevant. Others, however, can truly evaluate (and properly use) a set of projections only if they understand precisely how they were made. For them, a clearer and more comprehensive description of the methodology adds considerable value to the projections.

Political Considerations refer to the context in which projections are made. All projections are influenced by the context in which they are produced and by the perspectives of those who produce them; i.e., all projections are judgmental in the sense that they reflect a variety of choices made during their preparation. The outcomes of cohort-component models are determined by assumptions regarding future mortality, fertility, and migration rates; structural models are affected by choices of variables and functional forms, labor force and school enrollment projections are influenced by assumptions as to trends in labor force participation and school enrollment ratios. Even projections from simple trend extrapolation are affected by choices of data, techniques, and length of base period. Judgment is sometimes influenced by political (i.e., non-technical) considerations. As noted by Moen (1984), population growth is deeply embedded in politics. A national government may want to show that the elderly population is growing rapidly in order to support its initiative to overhaul the current retirement system. A state government may want to show that poverty rates are declining to illustrate the effectiveness of its economic policies. A business group may want to show the need for additional public investment in infrastructure. If such political concerns outweigh technical considerations, the credibility of the projections may be compromised.

Data users must become aware of the political context in which projections were made. Who made the projections? Did they have a personal stake in the results? What roles were the projections expected to play? What data, techniques, and assumptions were applied? Political considerations do not uniformly compromise the objectivity of projections; in some instances, in fact, they may substantially improve their quality (Tayman, 1996a). Learning the answers to questions like these, however, will help data users evaluate projection results.

Forecast Accuracy is for many analysts and data users, the most important issue (Carbone and Armstrong, 1982; Mentzer and Kahn, 1995; Yokum and Armstrong, 1995). Without such information, the usefulness of projections for many purposes is limited. Fortunately, a substantial amount of information on population forecast accuracy is available. Forecast error can be defined as the difference between a projected number and an actual number (Smith, et al., 2001: 302). For population projections, census counts are typically used as proxies for "actual" numbers; population estimates are sometimes used as well. Although it is widely recognized that census counts and population estimates also contain errors, such errors have relatively little impact on the accuracy of long-range projections and are seldom accounted for in evaluations of forecast accuracy.

Many measures of forecast accuracy can be used. In demography, the most common are the mean absolute percent error (MAPE), mean percent error (MPE), median (or median absolute) percent error, root mean squared error, root mean squared percent error, and various transformations of these measures or their underlying data (Ahlburg, 2001; Siegel, 2002: 471-484; Smith, et al., 2001). Some measures (e.g., MAPE) refer to precision, or how close a projection is to an actual value regardless of the direction of the difference; others (e.g., MPE) refer to bias, or the tendency for projections to be too high or too low.

Some generalizations on forecast accuracy are as follows.

(1) Precision tends to increase as population size increases (Bongaarts and Bulatao, 2000; Isserman, 1977; Murdock et al., 1984; White, 1954). Large places typically have smaller MAPEs than small places, but once a certain population size has been reached further increases in size generally do not lead to further increases in precision (Smith, 1987; Tayman, 1996b). Bias appears to have no consistent relationship with population size, as the tendency for projections to be too high or too low is about the same for small places as large places.

(2) Precision tends to be greatest for places with positive but moderate population growth rates and to deteriorate as growth rates deviate in either direction from these levels (Isserman, 1977; Murdock et al., 1984; Smith, 1987; Tayman, 1996b). MAPEs are particularly large for places that have been either growing or declining rapidly. Bias is also affected by growth rates, as projections for places that have been losing population tend to be too low and projections for places that have been growing rapidly tend to be too high (Smith, 1987; Smith and Shahidullah, 1995).

(3) Precision tends to decline as the length of the projection horizon increases. This result has been found not only for population projections (Bongaarts and Bulatao, 2000; Keilman, 1990; Keyfitz, 1981; Stoto, 1983), but for forecasts in other fields as well (Ascher, 1981; Batchelor and Dua, 1990; Schnaars, 1986). MAPEs have often been found to grow about linearly with the projection horizon, at least for several decades (Ascher, 1981; Schmitt and Crosetti, 1951; Smith and Sincich, 1991). The length of the projection horizon, however,

appears to have no impact on the tendency for projections to be too high or too low (Smith and Sincich, 1991).

(4) The lower the current fertility rate and the higher the prevailing life expectancy, the higher the precision, ceteris paribus (George and Nault, 1991).

(5) Forecast accuracy is not the same for all launch years (Keilman, 1990; Keyfitz, 1982; Long, 1995). Although measures of precision often show some degree of stability over time, measures of bias vary dramatically from one launch year to another (Isserman, 1977; Kale et al., 1981; Smith and Sincich, 1988). The study of past forecast errors may therefore tell us something about the likely level of precision of current projections, but it can tell us little or nothing about whether those projections are likely to be too high or too low.

(6) The choice of projection method has no consistent impact on forecast accuracy. No single method uniformly produces more accurate population projections than all other methods. In particular, complex methods are no more likely to provide accurate forecasts of total population than simpler methods (Stoto, 1983; Long, 1995; Pflaumer, 1992; Smith and Sincich, 1992; White, 1954). Similar results have been found for other types of forecasts as well (Mahmoud, 1984; Pant and Starbuck, 1990; Schnaars, 1986). Causal models have been found to provide more accurate population forecasts than non-causal models in a few instances (Sanderson, 1999), but most studies have found no consistent differences between causal and non-causal models (Kale et al., 1981; Murdock et al., 1984; Smith and Sincich, 1992).

(7) Incorporating expert opinion on the selection of methods and on assumptions for particular methods can contribute to higher precision (Ahlburg and Lutz, 1998).

(8) Combining projections based on different methods, data sets, or combinations of assumptions often leads to greater forecast accuracy than can be achieved by individual

projections. This result has been found for forecasts of variables as diverse as gross national product, corporate earnings, stock prices, electricity demand, psychiatric conditions, rainfall, and sunspot cycles (Clemen, 1989; Mahmoud, 1984; Schnaars, 1986). Raising the number of projections in the combination generally improves forecast accuracy, but by diminishing increments. Combining projections can be accomplished by using simple averages or various types of weighted averages. It improves forecasting performance because each individual projection provides unique information and the errors tend to offset each other to some degree. Combining reduces the risk of making large errors. Although it has seldom been used for population projections, there is some evidence that its use may now be increasing (Ahlburg, 1999; Smith and Shahidullah, 1995).

Accounting for Uncertainty has been the subject of many studies conducted over the last 50 years. They have found roughly similar results regarding the precision of population forecasts. Table 21-6 shows "typical" MAPEs for a variety of projection horizons and geographic levels. (These hypothetical errors are based on the assumption that there are no errors in the launch year populations.) Although errors for individual places will vary—with some having much larger errors and others much smaller errors than those shown here—this table provides rough but reasonable estimates of the average levels of precision that might be expected for current forecasts of states, counties, and census tracts. Errors for nations vary according to population size, with MAPEs for large countries generally being smaller than those shown here for states and MAPEs for small countries generally falling somewhere between those shown for states and counties (Bongaarts and Bulatao, 2000: 38-44).

(TABLE 21-6 ABOUT HERE)

Table 21-6 illustrates the uncertainty associated with population projections, especially for small areas and long horizons. Given the widespread use of population projections for decision making in both the public and private sectors—and the high stakes often associated with those decisions—it is essential that data users have some understanding of the uncertainty inherent in population projections. Summaries of previous forecast errors are helpful, but they do not provide information regarding the uncertainty of a specific current set of projections. Such information can be provided in several ways. One approach is to construct several alternative series based on different methods or different specifications of a particular method. The most common practice is to produce several sets of cohort-component projections based on different combinations of assumptions (Campbell, 1996; Day, 1996; Statistics Canada, 2001). The primary benefit of producing alternative series is that they show the populations stemming from different but reasonable models, techniques, and combinations of assumptions. The primary limitation is that they do not provide an explicit measure of uncertainty. The question, "How likely is it that the future population will fall within the range suggested by two alternative series?" cannot be answered using this approach.

Another approach is to construct prediction intervals to accompany a particular population forecast. Prediction intervals can be based on specific models of population growth (e.g., time series models), empirical analyses of past forecast errors, or the subjective judgment of population experts (Bongaarts and Bulatao, 2000: 200-204; Smith et al., 2001: 334-339). The primary advantage of this approach is that it provides an explicit probability statement to accompany a particular population forecast. Disadvantages include: (1) Model-based prediction intervals are data-intensive, difficult to produce, and subject to a variety of specification errors; (2)

Empirically-based prediction intervals require the collection of a large amount of historical data and are dependent on the assumption that future error distributions will be similar to past error distributions; and (3) Experts are often overconfident and tend to underestimate the uncertainty inherent in their forecasts (Bongaarts and Bulatao, 2000: 203). Alternative series and prediction intervals both provide useful information regarding the uncertainty of future population growth. Data users can incorporate this information into their deliberations and make better decisions than would be possible if they had no knowledge of the likely range of future errors. Using formal or informal "loss functions", they can assess the gains or losses associated with different forecast errors (Bongaarts and Bulatao, 2000: 188). In some instances (e.g., when it is more costly to anticipate too little growth than too much growth), the best choice may be to base decisions on low or high projections rather than the ones deemed most likely to provide an accurate forecast of future growth.

CONCLUDING NOTE

The production of population and related projections involves many choices regarding data, techniques, and assumptions. Making the best choices requires taking account of the purposes for which the projections will be used and the constraints under which they are produced. Simple methods require less time, data, and expertise than more complex methods, but provide less demographic and socioeconomic detail and offer fewer opportunities to analyze the determinants and consequences of population growth. Investigating input data series and correcting errors can improve data quality but is time consuming. Evaluating the consistency of projections with historical trends can uncover data and modeling errors but delays the release of projection results. Adjusting for special populations or unique events can improve forecast accuracy but requires specialized domain knowledge. Providing a range of projections increases the amount of information available to data users, but requires more production time and may open the door to political misuse of the results (e.g., choosing the scenario that favors a particular political position regardless of its technical merits). Only after balancing the costs and benefits of all aspects of the projection process can the analyst make optimal choices for any particular project.

These choices require the application of professional judgment. No matter how objective and rigorous the projection methodology, many subjective elements remain. Consequently, it is imperative that analysts provide a clear, comprehensive explanation of the projection methodology. Otherwise, data users cannot truly evaluate projection validity and plausibility.

We believe that it is generally best to use the simplest method(s) consistent with the purposes for which the projections will be used. This allows scarce resources to be directed toward activities that often have a substantial positive impact on the quality of the projections (e.g., evaluating and correcting input data, accounting for domain knowledge) rather than activities that have little or no impact (e.g., establishing an extensive database or developing an unnecessarily complicated model). Devoting resources to the collection of domain knowledge is particularly important for small-area projections, where unique events and special circumstances often play a major role in population change. In many instances, averages based on a variety of methods or sets of assumptions are likely to be preferable to projections based on a single model.

All forecasts are subject to error. Given the errors shown in Table 21-6, why should the analyst even bother making population and related projections? Why should data users pay any attention to them? There are several reasons for doing so. First, the projection process itself is educational, teaching a great deal about the components of population growth and the determinants of changes in related variables. Second, projections are helpful in analyzing the impact of alternative scenarios or combinations of assumptions on population growth and demographic change, regardless of the accuracy of specific forecasts. Finally, there is really no alternative to making projections: Ignoring potential change is generally not the best way to plan for the future.

Accuracy is an important characteristic, but it is not the only criterion upon which projections can (or should) be judged. In the final analysis, projections can best be judged according to their "utility," or the improvements they bring to the quality of information used in decision making (Tayman and Swanson, 1996). If these benefits are greater than the costs of production, then projections are worthwhile. Despite their shortcomings as forecasts, population and related projections can play an extremely important role in many types of planning and analysis.

Table 21-2 Projection Results for Island and Walla Walla Counties Using Different Extrapolation Methods, 2005 to 2015

	ISLAND COUNTY					WALLA WALLA COUNTY			
Method/Year	**2000**	**2005**	**2010**	**2015**		**2000**	**2005**	**2010**	**2015**
SIMPLE									
Linear	74200	81020	87841	94661		54200	55701	57201	58702
Geometric	74200	87617	103459	122166		54200	55929	57713	59554
Exponential	74200	87599	103416	122091		54200	55934	57724	59572
COMPLEX									
Linear Trend	74200	81837	89358	96879		54200	55472	57263	59053
Quadratic	74200	87949	98936	110936		54200	58664	62264	66194
Logistic	74200	83749	90437	96172		54200	56302	58542	60872
ARIMA	74200	77287	78889	79719		54200	53870	53863	53863
RATIO									
Constant	74200	78430	83692	89337		54200	57319	61133	65257
Shift	74200	82999	93352	104803		54200	53139	52216	50978
Share	74200	80371	87930	96095		54200	55559	57221	59017
RANGE									
Absolute	N/A	10662	24565	42447		N/A	5525	10048	15216
Percent	N/A	13.86	31.14	53.26		N/A	10.40	19.24	29.85

Table 21.3 Example for Calculation of Projected Births, Canada, 2001 (t+1

Age of mothers	Female Population July	Fertility Rates,	Number of births, 2000 (Calendar year)	Female Population July	Fertility Rates,	Number of births, 2001 (Calendar year)	Projected births (Census July 1, 2000 to June 30,
	(1)	(2)	(3)=(1)x(2)	(4)	(5)	(6)=(4)x(5)	(7)={(3)+(6)}/
15	200341	0.00548	1098	200725	0.00535	1074	1086
16	200917	0.00861	1730	202053	0.00842	1701	1716
17	200783	0.01302	2614	202684	0.01276	2586	2600
18	201164	0.01894	3810	202440	0.01861	3767	3789
19	204943	0.02656	5443	203345	0.02613	5313	5378
20	204680	0.03587	7342	206800	0.03536	7312	7327
21	202626	0.04670	9463	206939	0.04613	9546	9505
22	200593	0.05863	11761	204380	0.05802	11858	11810
23	203550	0.07101	14454	202822	0.07039	14277	14366
24	205067	0.08299	17019	205567	0.08241	16941	16980
25	205918	0.09364	19282	207255	0.09314	19304	19293
26	201925	0.10203	20602	208037	0.10166	21149	20876
27	204162	0.10740	21927	204000	0.10720	21869	21898
28	208697	0.10926	22802	206089	0.10925	22515	22659
29	218615	0.10746	23492	210782	0.10763	22686	23089
30	219846	0.10222	22473	220496	0.10255	22612	22543
31	219208	0.09406	20619	221512	0.09452	20937	20778
32	220974	0.08377	18511	220918	0.08431	18626	18569
33	226851	0.07222	16383	222732	0.07281	16217	16300
34	239975	0.06030	14470	228403	0.06088	13905	14188
35	257817	0.04877	12574	241593	0.04932	11915	12245
36	266928	0.03822	10202	258987	0.03872	10028	10115
37	271402	0.02904	7882	267849	0.02946	7891	7887
38	267494	0.02139	5722	272404	0.02173	5919	5821
39	271218	0.01528	4144	268253	0.01555	4171	4158
40	269080	0.01059	2850	271947	0.01080	2937	2894
41	264843	0.00713	1888	269641	0.00727	1960	1924
42	262199	0.00465	1219	265334	0.00476	1263	1241
43	257968	0.00295	761	262484	0.00302	793	777
44	250841	0.00182	457	258151	0.00186	480	469
Total	6830625	1.48002	322995	6824622	1.48001	321552	**322274**

Table 21.4 A Cohort-Component Example for Population Projections of the Female Population of Canada, 2000-2001 (medium projection)

Age	Pop. at t (2000): "Launch Year"	Immigrants	Emigrants	Net International Migration	Non-Permanent Residents (NPR)	Pop. Adjusted for Migration	Sx (Survival ratio)	Survivors at t+1	Pop. at t+1 (2001): First year of Projection
	(1)	(2)	(3)	(4)=(2)-(3)	(5)	(6)=(1)+(4)-(5)	(7)	(8)=(6)x(7)	(9)=(8)+(5)
-1	156690	0	0	0	0	156690	0.995536		
0	158410	1947	169	1778	66	160122	0.999294	155990	156056
1	167170	1513	206	1307	230	168247	0.999673	160009	160239
2	170843	1476	245	1231	371	171703	0.999767	168192	168563
3	176343	1500	290	1210	403	177150	0.999814	171663	172066
4								177117	177117
	..								
	..								
14	199459	1593	437	1156	1145	199470	0.999769		
15	200341	1588	406	1182	1300	200223	0.999740	199424	200724
16	200917	1520	378	1142	1878	200181	0.999710	200171	202049
17	200783	1496	332	1164	2557	199390	0.999684	200123	202680
18	201164	1715	266	1449	3109	199504	0.999659	199327	202436
19								199436	199436
	..								
	..								
49	217997	686	378	308	464	217841	0.997645		
50	213697	634	334	300	381	213616	0.997411	217328	217709
51	210634	568	301	267	385	210516	0.997150	213063	213448
52	210510	514	270	244	299	210455	0.996873	209916	210215
53	210399	513	234	279	274	210404	0.996545	209797	210071
54								209677	209677
	..								
	..								
95	7165	1	0	1	2	7164	0.826354		
96	5535	2	0	2	1	5536	0.810513	5920	5921
97	3871	1	0	1	0	3872	0.793388	4487	4487
98	2655	0	0	0	0	2655	0.774765	3072	3072
99	1832	1	0	1	0	1832	0.697146	2057	2057
100+	3317					3317	0.620460	3335	3335
Total	15522683	110128	32679	77449	98890	15501242		15553417	15652307.2

Note: The population aged 0 at time t+1 is obtained by first applying the proportion of female births (0.4862) to the total births (322,274), then the survival ratio from birth to age 0 for female births, and finally adding net migration.

Table 21.5 A Cohort-Component Example for Population Projections of the Female Population of Ontario, Canada, 2000-2001 (medium projection)

Age	Pop. at t (2000): "Launch Year"	Immigrants	Emigrants	Net International Migration	Net Internal Migration	Non-Permanent Residents (NPR)	Pop. Adjusted for Migration	Sx (Survival ratio)	Survivors at t+1	Pop. at t+1 (2001): First year of Projection
	(1)	(2)	(3)	(4)=(2)-(3)	(5)	(6)	(7)=(1)+(4)+(5)-(6)	(8)	(9)=(7)x(8)	(10)=(9)+(6)
-1	61330	0	0	0	0	0	61299	0.99567		
0	61525	886	88	798	219	31	62421	0.99935	61034	61065
1	65168	798	107	691	202	121	65877	0.99972	62380	62501
2	66361	804	127	677	185	184	67014	0.99980	65859	66043
3	68452	813	150	663	169	209	69014	0.99984	67001	67210
4									69003	69003
	..									
	..									
14	75848	844	220	624	123	571	75272	0.99980		
15	75750	857	204	653	159	572	76023	0.99978	75257	75829
16	74834	827	188	639	207	754	75808	0.99976	76006	76760
17	74561	820	166	654	249	1053	74627	0.99974	75790	76843
18	73734	929	134	795	296	1247	74217	0.99973	74608	75855
19									74197	74197
	..									
	..									
49	80975	343	188	155	82	232	82701	0.99775		
50	79518	342	167	175	76	174	81038	0.99752	82515	82689
51	78365	308	150	158	74	194	79575	0.99726	80837	81031
52	79434	269	137	132	67	139	78458	0.99698	79357	79496
53	80595	275	121	154	62	141	79492	0.99666	78221	78362
54									79226	79226
	..									
	..									
95	2660	1	0	1	1	2	2661	0.81887		
96	2054	1	0	1	1	1	2056	0.80265	2179	2180
97	1435	0	0	0	0	0	1435	0.78516	1650	1650
98	984	0	0	0	0	0	984	0.76634	1127	1127
99	679	0	0	0	1	0	680	0.74612	754	754
100+	1235						1236	0.66318	1327	1327
Total	5912532	59240	16717	42523	9941	42071	5922925		5944513	5986584

Note: The population aged 0 at time t+1 is obtained by first applying the proportion of female births (0.4862) to the total births (126,142), then the survival ratio from birth to age 0 for female births, and finally adding net migration.

Table 21-6 "Typical" Mean Absolute Percent Errors,

by Level of Geography and Length of Horizon

	Length of Horizon (years)					
Level of Geography	5	10	15	20	25	30
State	3	6	9	12	15	18
County	6	12	18	24	30	36
Census tract	9	18	27	36	45	54

Source: Smith, Tayman, and Swanson, 2001, Table 13.7.

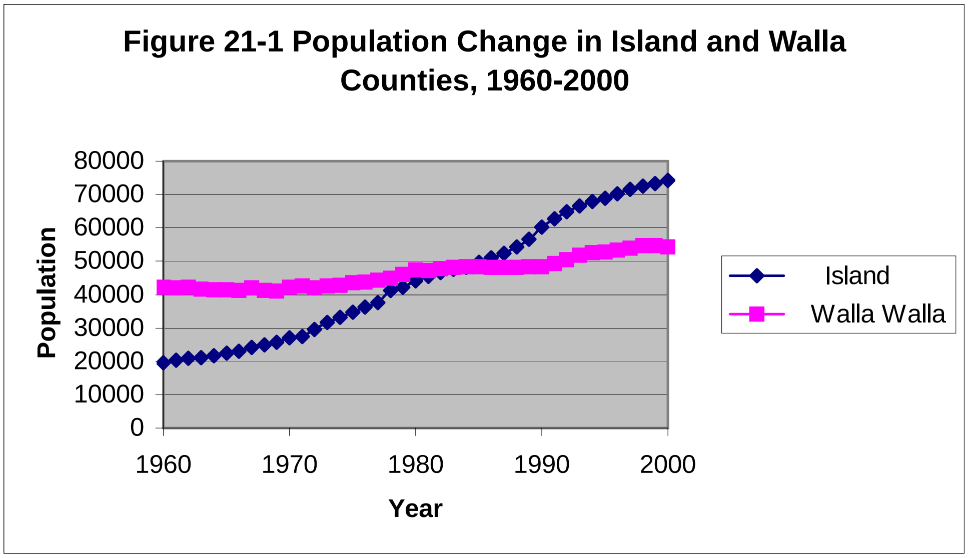
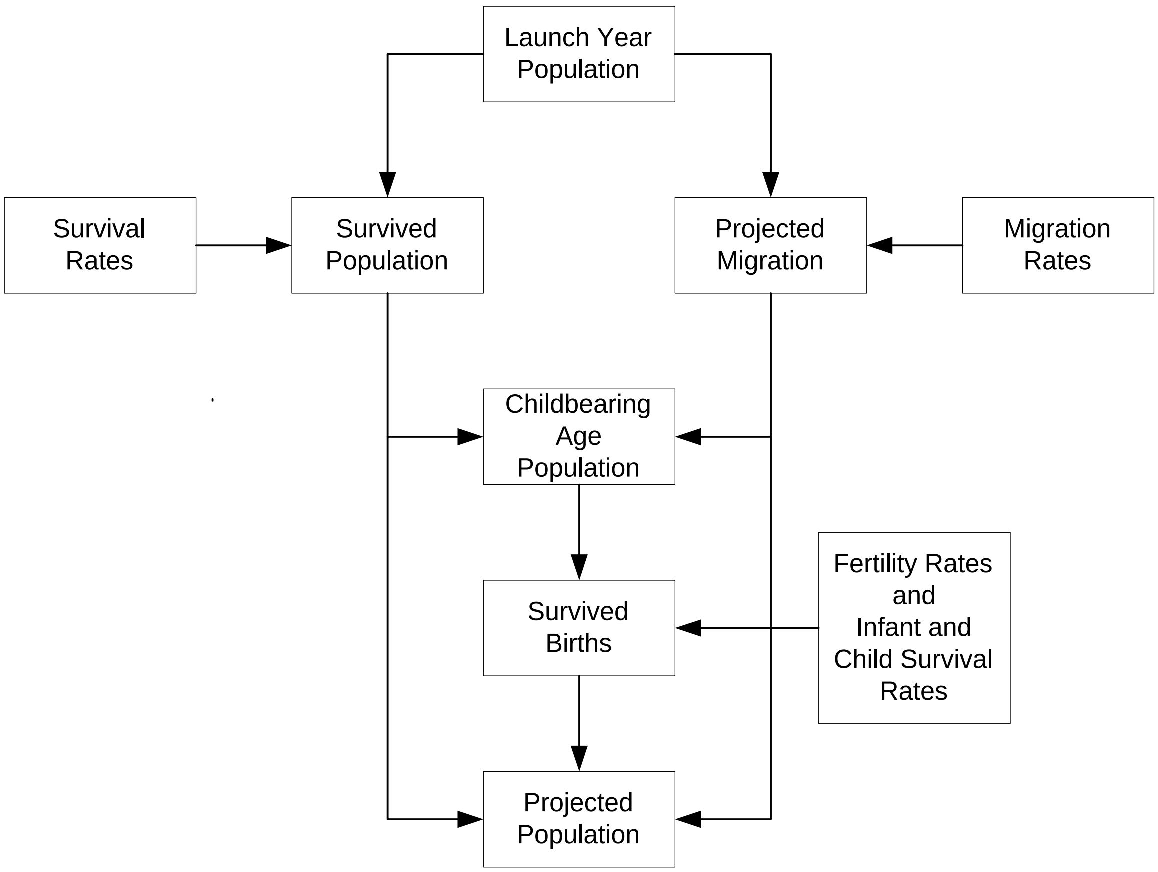


Figure 21-1 Population Change in Island and Walla Counties, 1960-2000
80000
70000
60000
50000
40000
30000
20000
10000
0
Population
1960
1970
1980
1990
2000
Year
Island
Walla Walla

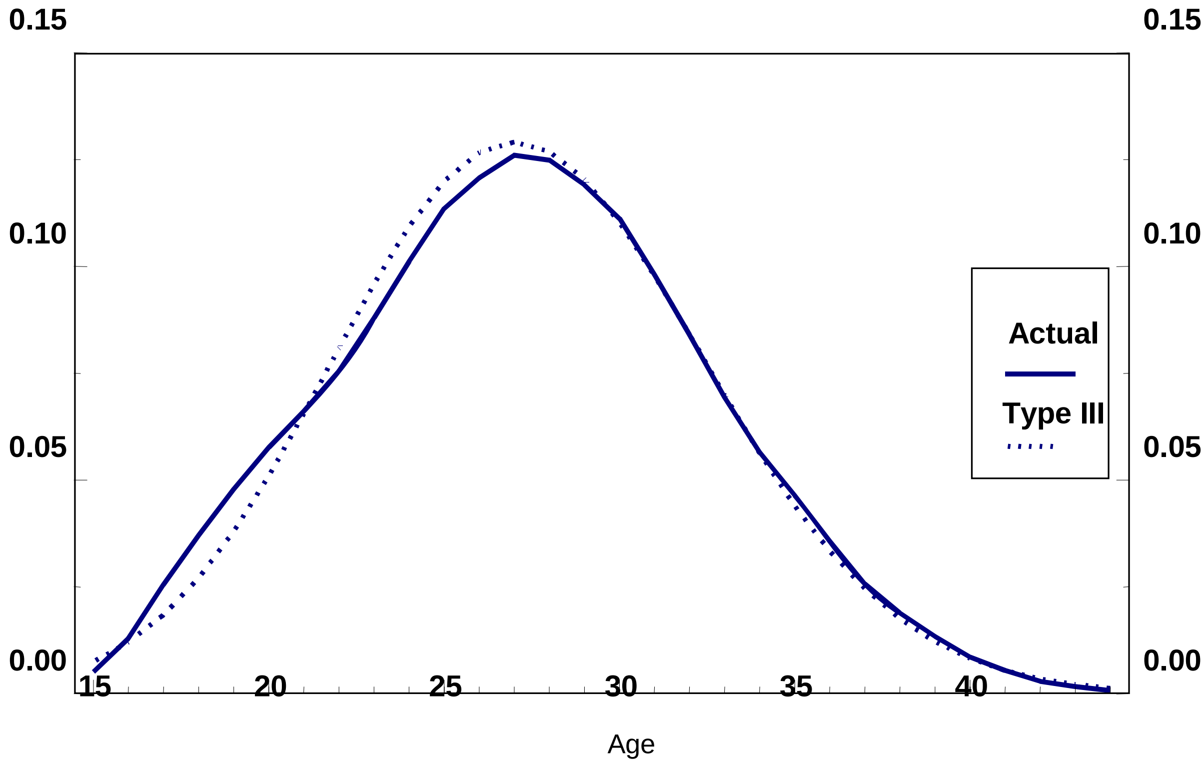
Figure 21-2 Overview of the Cohort Component Method



Smith et al. (2001: 47)

Figure 21-3 Comparison of Actual and Pearson Type III Distribution of Age-Specific Fertility Rates, Canada, 1991

Number of children per woman



Source: Statistics Canada, Demography Division, Population Projections Section

Figure 21-4. National Component Assumptions, Canada

Component	**No. of Assumptions**		**Assumptions**	
1. Fertility	3 TFR by 2026	High - 1.8	Medium - 1.48	Low - 1.3
2. Mortality	3 M/F e_o by 2026	High 81.5 85.0	Medium 80.0 84.0	Low 78.5 83.0
3. Immigration	3 Level by 2026	High - 270,000	Medium - 225,000	Low - 180,000
4. Total emigration	1	2-year average of age-sex specific emigration rates (1997-1998 and 1998-1999)		
5. Non-Permanent Residents	1	Constant number at 240,000		

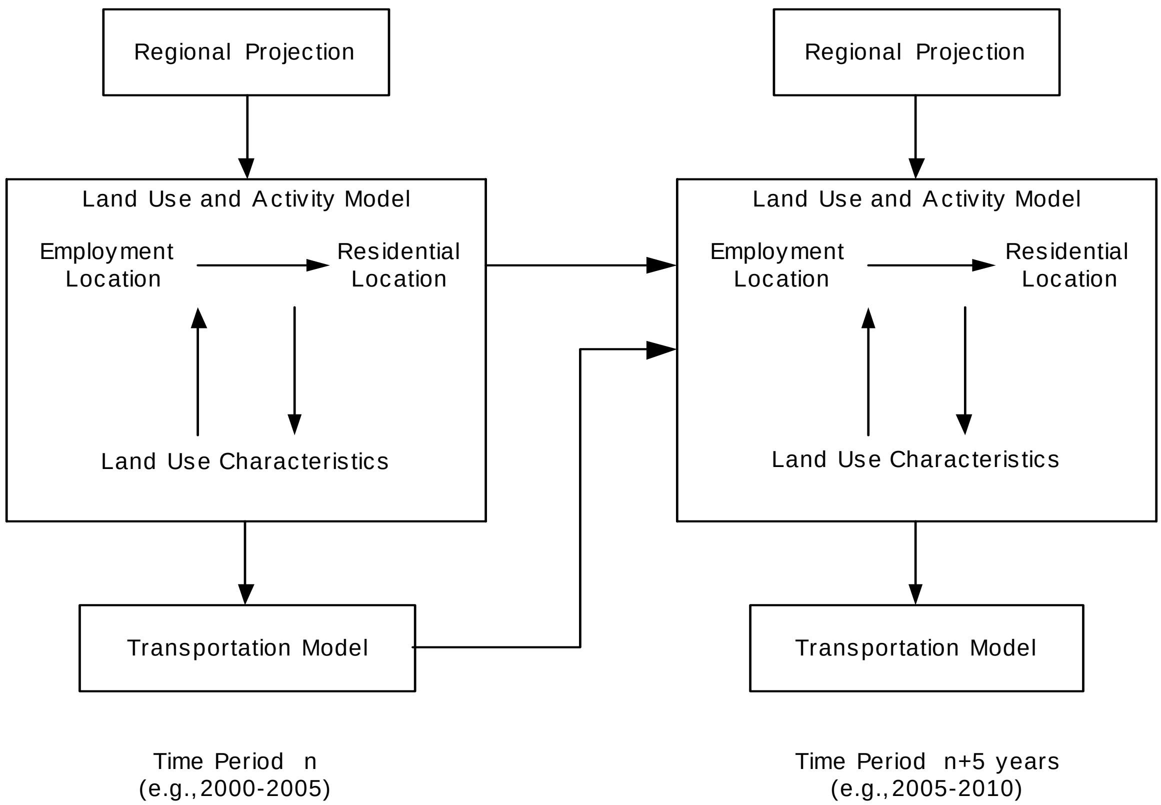


Figure 21-5 Urban Systems Model Components

Source: Smith et al. (2001: 219)

REFERENCES

Ahlburg, D. 1999. Using economic information and combining to improve forecast accuracy in demography. Unpublished paper. Rochester, MN: Industrial Relations Center, University of Minnesota.

Ahlburg, D. 2001. Population forecasting. pp. 557-575 in J. S. Armstrong, *Principles of forecasting: A handbook for researchers and practitioners*. Norwell, MA: Kluwer Academic Publishers.

Ahlburg, D. and W. Lutz. 1998. Introduction: the need to rethink approaches to population forecasts. *Population Development Review* 24: 1-14.

Ahlburg, D. and J. Vaupel. 1990. Alternative projections of the U.S. population. *Demography* 27: 639-652.

Alho, J., and B. Spencer. 1997. The practical specification of the expected error of population forecasts. *Journal of Official Statistics* 13: 203-225.

Armstrong, J. 2001. *Principles of forecasting: A handbook for researchers and practitioners*. Norwell, MA: Kluwer Academic Publishers.

Ascher, W. 1981. The forecasting potential of complex models. *Policy Sciences* 13: 247-267.

Astone, N., and S. McLanahan. 1994. Family structure, residential mobility and school report: A research note. *Demography* 31: 575-584.

Australian Bureau of Statistics. 2000. *Population Projections, Australia, 1999 to 2101*. Canberra, Australia: Australian Bureau of Statistics

Batchelor, R. and P. Dua. 1990. Forecaster ideology, forecasting technique, and the accuracy of economic forecasts. *International Journal of Forecasting* 6: 3-10.

Bélanger, A. 1992. Estimating Parameterized Migration Age Schedules for Canada and Provinces. Document de travail, Section des projections demographiques, Division de la demographie, Statistique Canada, Ottawa, Canada.

Bell, F., A. Wade, and S. Goss. 1992. *Life tables for the United States social security area 1900-2080*. Washington, DC: Social Security Administration.

Bongaarts, J. and R. Bulatao (Eds.). 2000. *Beyond Six Billion: Forecasting the World's Population*. Washington, DC: National Academy Press.

Bowley, A. 1924. Births and population in Great Britain. *The Economic Journal* 334: 188-192.

Boyce, D. 1988. Renaissance of large scale models. *Papers of the Regional Science Association* 65: 1-10.

Box, G. and G. Jenkins. 1976. *Time series analysis: Forecasting and control*. San Francisco, CA: Holden-Day.

CACI, Inc. 2000. *Sourcebook for ZipCode Demographics*. Chantilly, VA: CACI, Inc.

Campbell, P. 1996. Population projections for states by age, sex, race, and Hispanic Origin: 1995 to 2050. Report PPL-47. Washington, DC: U.S. Bureau of the Census.

Cannan, E. 1895. The probability of a cessation of the growth of population in England and Wales during the next century. *The Economic Journal* 5: 506-515.

Carbone, R. and J. Armstrong. 1982. Evaluation of extrapolative forecasting methods: Results of a survey of academicians and practitioners. *Journal of Forecasting* 1: 215-217.

Carter, L. and R. Lee. 1986. Joint forecasts of U.S. marital fertility, nuptiality, births, and marriages using time series models. *Journal of the American Statistical Association* 81: 902-911.

Carter, L. and R. Lee. 1992. Modeling and forecasting U.S. sex differentials in mortality.

International Journal of Forecasting 8: 393-411.

Clark, D. and W. Hunter. 1992. The impact of economic opportunity, amenities and fiscal factors on age-specific migration rates. *Journal of Regional Science* 32: 349-365.

Cleland, J. 1996. Demographic data in less developed countries, 1946-1996. *Population Studies* 50: 433-450.

Clemen, R. 1989. Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting* 5: 559-583.

Congdon, P. 1992. Multiregional demographic projections in practice: A metropolitan example. *Regional Studies* 26: 177-191.

Conway, R. 1990. The Washington projection and simulation model. *International Regional Science Review* 13: 141-165.

Cruijsen, H. 1994. National Population Projections in the European Economic Area - A Field in Motion. Joint ECE-Eurostat Work session on Demographic Projections, Luxembourg, (May). ECE Working Paper No. 4. Geneva, Switzerland: U.N. Economic Commission for Europe.

Cruijsen, H. and N. Keilman. 1992. A comparative analysis of the forecasting process. Pp. 3-25 in N. Keilman and H. Cruijsen (Eds.). *National Population forecasting in industrialized countries*. Amsterdam: Swets and Zeitliinger.

Data Resources Incorporated. 1998. *Review of the U.S. Economy: Long-range focus, Winter 97-98*. Lexington, MA: Data Resources Incorporated.

DaVanzo, J., and Morrison, P. 1978. *Dynamics of return migration: Descriptive findings from a longitudinal study. Report P-5913*. Santa Monica: The Rand Corporation.

Davis, H. 1995. *Demographic projection techniques for regions and smaller areas.*

Vancouver, Canada: UBC Press.

Day, J. 1996. Population projections of the United States by age, sex, race, and Hispanic origin: 1995 to 2050. Current Population Reports, P25, No. 1130. Washington, DC: U.S. Bureau of the Census.

de Beer, J. 1993. Forecast intervals of net migration: The case of the Netherlands. *Journal of Forecasting* 12: 585-599.

de Gans, H. A. 1999. *Population Forecasting 1895-1945: The Transition to Modernity*. Dordecht: Kluwer Academic Publishers.

EUROSTAT. 1998. Long-term Mortality Scenarios for the Countries of the European Economic Area (by W. Van Hoorn and J. De Beer). EUROSTAT Working Papers, E.no.8.

Foot, D., and W. Milne. 1989. Multiregional estimation of gross internal migration flows. *International Regional Science Review* 12: 29-43.

Fries, J. 1989. The compression of morbidity: Near or far? *The Milbank Quarterly* 67: 208-232.

Fuguitt, G., and D. Brown. 1990. Residential preferences and population redistribution. *Demography* 27: 589-600.

Fullerton, H. 1999. Labor force projections to 2008: steady growth and changing composition. *Monthly Labor Review* 122 (November): 19-32.

George, M.V. 2001. "Population Forecasting in Canada: Conceptual and Methodological Development." *Canadian Studies in Population* 28: 111-154.

George, M.V. 1999. On the use and users of demographic projections in Canada. Joint ECE-Eurostat Work session on demographic projections, Perugia, Italy, (May). ECE Working Paper no. 15. Geneva, Switzerland: UN Economic Commission for Europe.

George, M.V. and S. Loh. 2000. Top down versus bottom up: Statistics Canada's hybrid bottom-up case of national and provincial population projections. Paper presented at the annual meeting of the population association of America, Los Angeles, CA.

George, M.V., S. Loh, and R.B.P. Verma. 1997. Impact of Varying the Component Assumptions on Projected Total Population and Age Structure in Canada. *Canadian Population Studies* 24: 67-86.

George, M.V., and F. Nault. 1991. The Accuracy of Statistics Canada's Demographic Projections. Paper presented at the annual meeting of the population association of America, Washington D.C.

George, M.V., and J. Perreault. 1992. Methods of external migration projections and forecasts. pp. 87-103 in N. Keilman and H. Cruijsen (Eds.) *National Population forecasting in industrialized countries.* Amsterdam: Swets and Zeitlinger.

Greenberg, M., D. Krueckeberg, and C. Michaelson. 1978. *Local population and employment projection techniques*. New Brunswick, NJ: Rutgers University, Center for Urban Policy and Research.

Greenwood, M. 1975. Simultaneity bias in migration models: An empirical investigation. *Demography* 12: 519-536.

Greenwood, M., and G. Hunt. 1991. Forecasting state and local population growth with limited data: The use of employment migration relationships and trends in vital rates. *Environment and Planning A* 23: 987-1005.

Hamilton, C. and J. Perry. 1962. A short method for projecting population by age from one decennial census to another. *Social Forces* 41: 163-170.

Hanke, J. A., G. Reitsch, and D. Wichern. 2001. *Business forecasting, 7th Edition*. Saddle River, NJ: Prentice Hall.

Hollmann, F. W., T. J. Mulder, and J.E. Kallan. 2000. Methodology and Assumptions for the Population Projections of the United States: 1999-2100. *Population Division Working Paper No. 38*. U.S. Census Bureau, Washington, D.C.

Indian Office of the Registrar General. 2001. *Chapter 4, Population Projections, Provisional Population Totals, Census of India 2001, Series 1.* Dehli: Indian Office of the Registrar General

Israeli Central Bureau of Statistics. 1987. *Projections of Population in Judea, Samaria, and the Gaza Area up to 2002, based on the Population in 1982.* Jerusalem, Israel: Israeli Central Bureau of Statistics.

Isserman, A. 1977. The accuracy of population projections for subcounty areas. *Journal of the American Institute of Planners* 43: 247-259.

Isserman, A. 1985. Economic-demographic modeling with endogenously determined birth and migration rates: theory and prospects *Environment and Planning A* 17: 25-45.

Isserman, A., D. Plane, P. Rogerson, and P. Beaumont. 1985. Forecasting interstate migration with limited data: A demographic-economic approach. *Journal of the American Statistical Association* 80: 277-285.

Kale, B., P. Voss, C. Palit, and H. Krebs. 1981. On the question of errors in population projections. Paper presented at the meeting of the Population Association of America, Washington DC.

Keilman, N. 1990. *Uncertainty in National Population Forecasting: Issues, Backgrounds, Analysis, Recommendations*. Amsterdam: Swets and Zeitlinger.

Keilman, N. 1991. National projection methods in developed countries. pp. 465-486 in W. *Lutz (Ed.) Future* demographic trends in Europe and North America: What can we assume today? San Diego, CA: Academic Press.

Keyfitz, N. 1968. *An introduction to the mathematics of population*. Reading, MA:

Addison-Wesley.

Keyfitz, N. 1981. The limits of population forecasting. *Population and Development Review* 7: 579-593.

Keyfitz, N. 1982. Can knowledge improve forecasts? *Population and Development Review* 8: 729-751.

Kintner, H. 1989. Demographic change in a corporate health benefits population. *American Journal of Public Health* 79: 1655-1656.

Kintner, H. and D. A. Swanson. 1994. Estimating vital rates from corporate databases: How long will GM's salaried retirees live? pp. 265-295 in H. Kintner, T. Merrick, P. Morrison, and P. Voss (Eds.), *Demographics: A Casebook for business and government*. Boulder, CO: Westview Press.

Kintner, H. and D. A. Swanson. 1996. Ties that bind: A case study of the link between employers, families, and health benefits. *Population Research and Policy Review* 15: 509-526.

Kintner, H. , T. Merrick, P. Morrison, and P. Voss (eds.). 1994. *Demographics: A Casebook for business and government.* Boulder, CO: Westview Press.

Kupiszewski, M. and P. Rees. 1999. *Lessons for the Projection of Internal Migration from Studies in Ten European Countries. Working Paper No. 40.* Conference of European Statisticians, Joint ECE-EUROSTAT Work Session on Demographic Projections, Perugia, Italy, May.

Land, K. 1986. Methods for national population forecasts: A review. *Journal of the American Statistical Association* 81: 888-901.

Lee, R. 1993. Modeling and forecasting the time series of U.S. fertility: Age distribution, range, and ultimate level. *International Journal of Forecasting* 9: 187-212.

Lee, R. and L. Carter. 1992. Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association* 87: 659-675.

Lee, R. and S. Tuljapurkar. 1994. Stochastic population forecasts for the United States: Beyond high, medium, and low. *Journal of the American Statistical Association* 89: 1175-1189.

Lee, R. and S. Tuljapurkar. 1997. Death and taxes: Longer life, consumption and social security. *Demography* 34: 67-81.

Long, J. 1989. The relative effects of fertility, mortality, and immigration on projected population age structure. Paper presented at the meeting of the Population Association of America. Baltimore, MD.

Long, J. 1995. Complexity, accuracy, and utility of official population projections. *Mathematical Population Studies* 5: 203-216.

Long, J. and D. McMillen. 1987. A survey of Census Bureau projection methods. *Climatic Change* 11: 141-177.

Lutz, W., J.W. Vaupel, and D.A. Ahlburg, eds. 1999. *Frontiers of Population Forecasting*. New York: The Population Council (a supplement to *Population and Development Review* 24, 1998).

Mahmoud, E. 1984. Accuracy in forecasting: A survey. *Journal of Forecasting* 3: 139-159.

Manton, K., C. Patrick, and E. Stallard. 1980. Mortality model based on delays in progression of chronic diseases: Alternative to cause elimination model. *Public Health Reports* 95: 580-588.

Manton, K., E. Stallard, and H. Tolley. 1991. Limits to human life expectancy: evidence prospects , and implications. *Population and Development Review* 17: 603-637.

Mason, A. 1996. Population and Housing" *Population Research and Policy Review* 15 (5-6): 419-435.

Massey, D., Alarcon, R., Durand, J., and Gonzalez, H. 1987. *Return to Aztlan: The social process of international migration from Western Mexico.* Berkeley, CA.: University of California Press.

McCleary, R. and R. Hay. 1980. *Applied time series analysis for the social sciences*. Beverly Hills, CA: Sage Publications.

McDonald, J. 1979. A time series approach to forecasting Australian total live-births. *Demography* 16: 575-601.

Mentzer, J. and K. Kahn. 1995. Forecasting technique familiarity, satisfaction, wage, and application. *Journal of Forecasting* 14: 465-476.

Michaud, J., M. V. George, and S. Loh. 1996. *Projections of Persons with Disabilities (Limited at Work/Perception); Canada, Provinces, and Territories, 1993-2016*. Ottawa, Canada: Statistics Canada..

Miller, T. 2001. Increasing longevity and Medicare expenditures. *Demography* 38: 215-226.

Mills, E. and L. Lubuele.1995. Projecting growth in metropolitan areas. *Journal of Urban Economics* 37: 344-360.

Moen, E. 1984. Voodoo forecasting: Technical, political and ethical issues regarding the projection of local population growth. *Population Research and Policy Review* 3: 1-25.

Mosert, W. and J. van Tonder. 1987. *Projections of the South African Population*. Pretoria, South Africa: Institute for Sociological and Demographic Research, Human Sciences Research Council.

Murdock, S., and D. Ellis. 1991. *Applied demography: An introduction of basic,* *concepts, methods, and data*. Boulder, CO: Westview Press.

Murdock, S., F. Leistritz, R. Hamm, S. Hwang, and B. Parpia. 1984. An assessment of the accuracy of a regional economic-demographic projection model. *Demography* 21: 383-404.

Murdock, S., R. Hamm, P. Voss, D. Fannin, and B. Pecotte. 1991. Evaluating small area population projections. *Journal of the American Planning Association* 57: 432-443.

National Center for Health Statistics. 1997. *U.S. decennial life tables for 1989-91. Volume 1, No. 1.* Washington, DC.

NCSS, Inc. 1995, *NCSS 6.0 Users' Guide*. Kaysville, UT: NCSS, Inc.

Olshansky, S. 1987. Simultaneous/multiple cause-delay (SIMCAD): An epidemiological approach to projecting mortality. *Journal of Gerontology* 42: 358-365.

Olshansky, S. 1988. On forecasting mortality. *Milbank Quarterly* 66: 482-530.

O'Neill, B. D. Balk, M. Brickman, and M. Ezra. 2001. A guide to global population projections. *Demographic Research* 4(8): 204-288. (http://www.demographic-research.org, retrieved July 2001).

Opitz, W. and H. Nelson. 1996. Short-term population-based forecasting in the public sector: A dynamic caseload simulation model. *Population Research and Policy Review* 15: 549-563.

Oregon Office of Economic Analysis. 2000. *Oregon Corrections Population Forecast.* Salem, OR: Oregon Department of Administrative Services, Office of Economic Analysis.

Pant, P. and W. Starbuck. 1990. Innocents in the forest: Forecasting and research methods. *Journal of Management* 16: 433-460.

Pflaumer, P. 1992. Forecasting U.S. population totals with the Box-Jenkins approach. *International Journal of Forecasting* 8: 329-338.

Pielou, E. C. 1969. *An introduction to mathematical ecology*. New York, NY: Wiley-Interscience.

Pignataro, L. and Epling, J. 2000. The TELUS story: Information tool for transportation planning makes it debut. *Transportation Research News* 210: 9-12, 25.

Pittenger, D. 1976. *Projecting state and local populations*. Cambridge, MA: Ballinger Publishing Company.

Pol, L. and R. Thomas. 2001. *The Demography of Health and Health Care. 2nd Edition*. New York, NY: Kluwer Academic/Plenum Press.

Putman, S. 1991. *Integrated urban models: II*. London, UK: Pion Limited.

Ravenstein, E. 1889. The laws of migration. *The Journal of the Royal Statistical Society,* LII, 241-301.

Rogers, A. 1985. *Regional population projection methods*. Beverly Hills, CA: Sage Publications.

Rogers, A. 1995. *Multiregional demography: Principles, methods and extensions*. Chichester, UK: John Wiley and Sons.

Rogers, A. and L. Castro. 1978. Model Migration Schedules and Their Application. *Environment and Planning A* 10: 457-502.

Rogers, A., and P. Williams. 1986. Multistate demoeconomic modeling and projection. pp. 177-202 in A. Isserman (Ed.), *Population change in the economy: Social science theory and methods.* Boston, MA: Kluwer-Nijhoff.

Rosenthal, N. 1999. The quality of BLS projections: a historical account. *Monthly Labor Review* 122 (May): 27-35.

Ryder, N. 1990. What is going to happen to American fertility? *Population and Development Review* 16: 433-453.

Saboia, J. 1974. Modeling and forecasting populations by time series: The Swedish case.

Demography 11: 483-492.

Sanderson, W. 1999. Knowledge can improve forecasts: a review of selected socioeconomic population projection models. *Population and Development Review* 24: 88-117.

San Diego Association of Governments. 1999. Demographic and economic forecasting model: Volume 2: Technical description. San Diego, CA.

San Diego Association of Governments. 1998. Urban development model, Volume 2: Technical description. San Diego, CA.

Schmidt, R., C. Barr, and D. Swanson. 1997. Socioeconomic impacts of the proposed federal gaming tax. *International Journal of Public Administration* 20: 1675-1698.

Schmitt, R, and A. Crosetti. 1951. Accuracy of the ratio method for forecasting city population. *Land Economics* 27: 346-348.

Schnaars, S. 1986. A comparison of extrapolation models on yearly sales forecasts. *International Journal of Forecasting* 2: 71-85.

Shryock, H. and J. Siegel. 1973. *The methods and materials of demography*. Washington, DC: U.S. Government Printing Office. *Demography* 9: 51-68.

Siegel, J. 2002. *Applied Demography: Applications to business, government, law, and public policy*. New York, NY: Academic Press.

Smith, S. 1986. Accounting for migration in cohort-component projections of state and local populations. *Demography* 23: 127-135.

Smith, S. 1987. Tests of forecast accuracy and bias for county population projections. *Journal of the American Statistical Association* 82: 991-1003.

Smith, S. and B. Ahmed. 1990. A demographic analysis of the population growth of states, 1950-1980. *Journal of Regional Science* 30: 209-227.

Smith, S. and M. Shahidullah. 1995. An evaluation of population projection errors for census tracts. *Journal of the American Statistical Association* 90: 64-71.

Smith, S. and T. Sincich. 1988. Stability over time in the distribution of population forecast errors. *Demography* 25: 461-474.

Smith, S. and T. Sincich. 1991. An empirical analysis of the effect of the length of forecast horizon on population forecast errors. *Demography* 28: 261-274.

Smith, S. and T. Sincich. 1992. Evaluating the forecast accuracy and bias of alternative population projections for states. *International Journal of Forecasting* 8: 495-508.

Smith, S. and D. A. Swanson. 1998. In defense of the net migrant. *Journal of Economic and Social Measurement* 24: 249-264.

Smith, S., J. Tayman, and D. A. Swanson. 2001. *State and Local Population Projections: Methodology and Analysis*. New York, NY: Kluwer Academic/Plenum Publishers.

Spencer, G. 1989. Projections of the population of the United States by age, sex, and race: 1988 to 2008. *Current Population Reports* P-25, No. 1018. Washington DC: U.S. Bureau of the Census.

Statistics Canada, 2001. *Population projections for Canada, Provinces, and Territories, 2000-2026.* Catalogue No. 91-520. Demography Division. Ottawa, Ontario: Statistics Canada.

Statistics Canada, 2002. *Annual Demographic Statistics, 2001*. Catalogue No. 91-213. Demography Division. Ottawa, Ontario: Statistics Canada.

Statistics New Zealand. 2000. *Subnational Population Projections: 1996 (Base) to 2021 (May 2000 Release*). Last accessed May 2002 online at (http://www.stats.govt.nz)

Stoto, M. 1983. The accuracy of population projections. *Journal of the American Statistical Association* 78:13-20.

Swanson, D. and D. Beck. 1994. A new short-term county population projection method.

Journal of Economic and Social Measurement 20: 25-50.

Swanson, D. and B. Klopfenstein. 1987. How to forecast VCR penetration. *American Demographics* (December): 44-45.

Swanson, D., G. Hough, J. Rodriguez, and C. Clemans. 1998. K-12 enrollment forecasting: Merging methods and judgment. *ERS Spectrum* 16: 24-31.

Tabuchi, T. 1985. Time series modeling of gross migration and dynamic equilibrium. *Journal of Regional Science* 25: 65-83.

Tayman, J. 1996a. The accuracy of small-area population forecasts based on a spatial interaction land-use modeling system. *Journal of the American Planning Association* 62: 85-98.

Tayman, J. 1996b. Forecasting, growth management, and public policy decision making. *Population Research and Policy Review* 15: 491-508.

Tayman, J. and D. Swanson. 1996. On the utility of population forecasts. *Demography* 33: 523-528.

Tayman, J., B. Parrott, and S. Carnevale. 1994. Locating fire station sites: The response time component. pp. 203-217 in H. Kintner, P. Voss, P. Morrison, and T. Merrick (Eds.), *Applied demographics: A casebook for business and government*. Boulder, CO: Westview Press.

Texas Water Development Board. 1997. Water for Texas: A consensus-based update to the state water plan Vol. II, technical planning appendix. Austin, TX: Texas Water Development Board.

Thomas, R. 1994. Using demographic analysis in health services planning: A case study in obstetrical services. pp. 159-179 in H. Kintner, T. Merrick, P. Morrison, and P. Voss (Eds.), *Demographics: A casebook for business and government*. Boulder, CO: Westview Press.

Thompson, A. 1999. Industry output and employment projections to 2008. *Monthly*

Labor Review 122 (November): 33-50.

Thompson, W. and P. Whelpton. 1933. *Population trends in the United States*. New York, NY: McGraw-Hill Book Company.

Treadway, R. 1997. Population projections for the state and counties of Illinois. Springfield, IL: State of Illinois.

Treyz, G. 1995. Policy analysis: Application of REMI economic forecasting and simulation models. *International Journal of Public Administration,* 18: 13-42.

Treyz, G. 1993. *Regional economic modeling: A systematic approach to economic forecasting and policy analysis*. Boston, MA: Kluwer.

Treyz, G., D. Rickman, G. Hunt, and M. Greenwood. 1993. The dynamics of U.S. internal migration. *Review of Economics and Statistics* 75: 209-214.

United Nations. 1998. *World Population Projections to 2150*. New York, NY: United Nations Population Division.

United Nations. 1999. *World Population Prospects: The 1988 Revision*. New York, NY: United Nations Population Division.

United Nations. 1973. *Manual VII. Methods of Projecting Households and Families*. New York, NY: United Nations.

U.S. Bureau of Economic Analysis. 1995. *BEA regional projections to 2045: Volume I: States.* Washington, DC.

U.S. Bureau of the Census. 1957. Illustrative projections of the population, by state, 1960, 1965, and 1970. Current Population Reports, P-25, No. 160. Washington, DC.

U.S. Bureau of the Census. 1966. Illustrative projections of the population of states: 1970 to 1985. Current Population Reports, P-25, No. 326. Washington, DC.

U.S. Bureau of the Census. 1972. Preliminary projections of the population of states:

1975-1990. Current Population Reports, P-25, No. 477. Washington DC.

U.S. Bureau of the Census. 1979. Illustrative projections of state populations by age, race, and sex: 1975 to 2000. Current Population Reports, P-25, No. 796. Washington, DC.

U.S. Census Bureau . 2001. *International Data Base*. Last Accessed March 2002 online at (http://www.census.gov/ipc/www/idbnew.html7/9/01)

Verma, R. B. P. and M.V. George. 1993. Projecting the Age Distribution of Immigrants in Canada: An alternative approach by linking Immigrants flow and stock data..” Paper prepared for the Conference of European Statisticians, Joint ECE/Eurostat Work Session on Migration Statistics, Geneva, Switzerland.

Verma, R. B. P., S. Loh., S.Y. Dai, and D. Ford. 1994. Fertility Projections for Canada, Provinces and Territories, 1993-2016. Background Paper, Population Projections Section, Demography Division, Statistics Canada. Ottawa, Ontario.

Voss, P. and B. Kale. 1985. Refinements to small-area population projection models: Results of a test based on 128 Wisconsin communities. Paper presented at the meeting of the Population Association of America. Boston, MA.

Waddell, P. 2000. A behavioral simulation model for metropolitan policy analysis and planning: Residential location and housing market components of UrbanSim. *Environment and Planning B* 27: 247-263.

Washington State Office of Financial Management. 2000. *Historical/Current Data Set: Total Resident Population by Year for County, Washington, 1960 to 2000*. Olympia, WA: Washington State Office of Financial Management. Last accessed March, 2002 online at: (http://www.ofm.wa.gov/pop6098/pop6098toc.htm12/28/01

Washington State Office of Financial Management. 2001. *Forecast of the State Population by Age and Sex, 1990 to 2030, November 2001 Forecast.* Olympia, WA:

Washington State Office of Financial Management . Last accessed May, 2002 online at: (http://www.ofm.wa.gov/popagesex19702020/popagesex19702020toc.htm)

Wetrogan, S. 1990. Projections of the population of states by age, sex, and race: 1989-2010. Current Population Reports, P-25, No. 1053. Washington, DC: U.S. Bureau of the Census.

Whelpton, P. 1928. Population of the United States, 1925 to 1975. *American Journal of Sociology* 34: 253-270.

White, H. 1954. Empirical study of the accuracy of selected methods of projecting state populations. *Journal of the American Statistical Association* 49: 480-498.

Yaffee, R. 2000. Introduction to Time Series Analysis and Forecasting, with Applications of SAS and SPSS. New York, NY: Academic Press.

Yokum, J. and J. Armstrong. 1995. Beyond accuracy: Comparison of criteria used to select forecasting methods. *International Journal of Forecasting* 11: 591-597.

SUGGESTED READINGS

General Works on Projection and Forecasting

Armstrong, J. 1985. *Long range forecasting: From crystal ball to computer*. New York, NY: John Wiley and Sons.

Cohen, J. 1995. *How many people can the earth support*? New York, NY: W.W. Norton and Company.

Makridakas, S., S. Wheelwright, and R. Hyndman. 1998. *Forecasting: methods and* *applications.* New York, NY: Wiley.

United Nations. 1993. *Projection Methods for Integrating Population Variables into Development Planning. Volume 1: Methods for Comprehensive Planning. Module 3: Techniques for Preparing Projections of Household and other Incomes, Household Consumption, and Savings and Government Consumption and Investment*. New York, NY: United Nations.

Specialized Methods

Dalkey, N., and O. Helmer, O. 1963. An experimental application of the Delphi method to the use of experts. *Management Science* 9: 458-467.

McNees, S. 1992. The uses and abuses of 'consensus' forecasts. *Journal of Forecasting,* 11, 703-710.

Morrill, R., A. Rogers, and G. Thrall. 1988. *Spatial Diffusion*. Newbury Park, CA: Sage Publications.

Rowe, G. and G. Wright. 1999. “The Delphi Technique as a Forecasting Tool: Issues and Analysis. *International Journal of Forecasting* 15: 353-375.

Trend Extrapolation

Fildes, R. 1992. The evaluation of extrapolative forecasting methods. *International Journal of Forecasting,* 8: 81-98.

Marchetti, C., P. Meyer, and J. Ausubel. 1996. Human population dynamics revisited with the logistic model: How much can be modeled and predicted. *Technological Forecasting and Social Change,* 52, 1-30.

McNown, R., A. Rogers, and J. Little. 1995. Simplicity and complexity in extrapolative population forecasting models. *Mathematical Population Studies* 5: 235-257.

Pearl, R. and L. Reed. 1920. On the rate of growth of the population of the United States since 1790 and its mathematical representation. *Proceedings of the National Academy of Science* 6: 275-287.

Schmitt, R. and A. Crosetti. 1953. Short-cut Methods of Forecasting City Population. *Journal of Marketing* 17: 417-424.

Swanson, D. A., D. Beck., and J. Tayman. 1995. "On the Utility of Lagged Ratio-Correlation as a Short-cut Population Projection Method: A Case Study of Washington State." *Journal of Economic and Social Measurement* 21: 1-16.

Neural Networks

Adya, M., and F. Collopy, F. 1998. How effective are neural networks at forecasting and prediction: A review and evaluation. *Journal of Forecasting,* 17: 481-495.

Hill, T., L. Marquez, M. O'Conner, and W. Remus. 1994. Artificial neural network models for forecasting and decision making. *International Journal of Forecasting,* 10, 5-15.

Lippmann, R. 1987. An introduction to computing with neural nets. *ISEE ASSP Magazine,* 4: 4-22.

Wasserman, P. 1989. *Neural computing: Theory and practice*. New York, NY: Van Nostrand Rheinhold.

Zhang, G., B. Patuwo, and M. Hu. 1998. Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting,* 14: 35-62.

Structural Models

Harris, B. 1994. The real issues concerning Lee's requiem. *Journal of the American Planning Association,* 60: 31-34.

Hunt, D., and D. Simmons. 1993. Theory and application of an integrated land-use and transport modelling framework. *Environment and Planning B,* 20: 221-244.

Kim, T. 1989. *Integrated urban systems modeling: Theory and applications*. Cambridge, MA: Harvard University Press.

Klosterman, R. 1994. Large scale urban models: Retrospect and prospects. *Journal of the American Planning Association* 60: 3-6.

Landis, J. 1994. The California urban futures model: A new generation of metropolitan simulation models. *Environment and Planning B* 31: 399-420.

Landis, J. 1995. Imagining land use futures: Applying the California urban futures model. *Journal of the American Planning Association,* 61: 438-457.

Landis, J., Zhang, M., and Zook, M. 1998. CUFII: The second generation of the California urban futures model. Berkeley, CA: UC Transportation Center, University of California.

Lee, D. 1973. Requiem for large-scale models. *Journal of the American Institute for Planners,* 39: 163-178.

Lowry, I. 1964. A model of metropolis. Report RM 4125-RC. Santa Monica, CA: The

Rand Corporation.

Wegener, M. 1994. Operation urban models: State of the art. *Journal of the American Planning Association,* 60: 17-30.

Forecast Uncertainty and Accuracy

Ahlburg, D., W. Lutz and J. Vaupel. 1998. Ways to Improve Population Forecasting: What should be done differently in the future? *Population Development Review* 24: 191-199

Alho, J. 1997. Scenarios, uncertainty and conditional forecasts of the world population. *Journal of the Royal Statistical Society A,* 160, part 1: 71-85.

Armstrong, J., and Collopy, F. 1992. Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting,* 8: 69-80.

Armstrong, J., and Fildes, R. 1995. Correspondence on the selection of error measures for comparisons among forecasting methods. *Journal of Forecasting,* 14: 67-71.

Cohen, J. 1986. Population forecasts and the confidence intervals for Sweden: A comparison of model-based and empirical approaches. *Demography* 23: 105-126.

Keilman, N. 1998. How accurate are the United Nations world population projections? Pp. 15-41 in W. Lutz, J. Vaupel, and D. Ahlburg (Eds.), *Frontiers of Population Forecasting*. New York, NY: The Population Council. (A supplement to *Population and Development Review*, 24).

Lutz, W., Sanderson, W., and Scherbov, S. 1998. Expert-based probabilistic population projections. In W. Lutz, J. Vaupel, and D. Ahlburg (Eds.), *Frontiers of Population Forecasting.* (pp. 139-155). New York, NY: The Population Council. (A supplement to *Population and*

Development Review, 24).

Lutz, W., P. Saariluoma, W. Sanderson, and S. Scherbov. 2000. New developments in the methodology of expert- and argument-based probabilistic population projection. Interim Report 1R-00-020. Laxenburg, Austria: IIASA.

Makridakis, S. 1993. Accuracy measures: Theoretical and practical concerns. *International Journal of Forecasting,* 9: 527-529.

Pflaumer, P. 1988. Confidence intervals for population projections based on Monte Carlo methods. *International Journal of Forecasting,* 4: 135-142.

Shryock, H. S. 1954. "Accuracy of Population Projections for the United States." *Estadistica*, 12: 587-597.

Smith, S., and Sincich, T. 1990. The relationship between the length of the base period and population forecast errors. *Journal of the American Statistical Association,* 85: 367-375.

Swanson, D. and J. Tayman. 1995. Between a rock and a hard place: The evaluation of demographic forecasts. *Population Research and Policy Review* 14: 233-249.

Tayman, J., and Swanson, D. 1999. On the validity of MAPE as a measure of population forecast accuracy. *Population Research and Policy Review,* 18: 299-322.

Tayman, J., Schafer, E., and Carter, L. 1998. The role of population size in the determination and prediction of population forecast errors: An evaluation using confidence intervals for subcounty areas. *Population Research and Policy Review,* 17: 1-20.

Projecting Mortality

Bennett, N., and Olshansky, S. 1996. Forecasting U.S. age structure and the future of social security: The impact of adjustments to official mortality schedules. *Population and Development Review,* 22, 703-727.

Budlender, D. and R. Dorrington. 2002. *ASSA AIDS and demographic models: user guide*. Cape Town, South Africa: Center for Actuarial Research, University of Cape Town.

Horiuchi, S., and J. Wilmoth. 1998. Deceleration in the age pattern of mortality at older ages. *Demography,* 35, 391-412.

Olshansky, S., Carnes, B., and C. Cassel. 1990. In search of Methuselah: Estimating the upper limits to human longevity. *Science,* 250, 634-640.

White, K., and S. Preston. 1996. How many Americans are alive because of improvements in mortality? *Population and Development Review,* 22: 415-429.

Projecting Fertility

Ahlburg, D. 1982. How accurate are the U.S. Bureau of the Census projections of total live births. *Journal of Forecasting* 1: 365-374.

Ahlburg, D. 1983. Good times, bad times: A study of the future path of U.S. fertility. *Social Biology,* 30: 17-23.

Akers, D. S. 1965. “Cohort Fertility Versus Parity Progression as Methods of Projecting Births.” *Demography* 2: 414-428.

Bhrolchain, M. 1992. Period paramount? A critique of the cohort approach to fertility. *Population and Development Review,* 18: 599-629.

Bloom, D., and J. Trussell. 1984. What are the determinants of delayed childbearing and permanent childlessness in the United States? *Demography,* 21: 591-611.

Bongaarts, J., and G. Feeney. 1998. On the quantum and tempo of fertility. *Population and Development Review,* 24: 271-291.

Bumpass, L. and C. Westoff. 1969 "The Prediction of Completed Fertility." *Demography* 6: 445-454

Easterlin, R. 1987. *Birth and fortune: The impact of numbers on personal welfare*. (2nd ed.). Chicago, IL: University of Chicago Press.

Friedman, D., M. Hechter, and S. Kanazawa. 1994. A theory of the value of children. *Demography,* 31: 375-401.

Golini, A. 1998. How low can fertility be? An empirical exploration. *Population and Development Review,* 24: 59-73.

Lee, R. 1974. Forecasting births in post-transition populations: Stochastic renewal with serially correlated fertility. *Journal of the American Statistical Association,* 69: 607-617.

Lee, R. 1976. Demographic forecasting and the Easterlin hypothesis. *Population and Development Review,* 2, 459-468.

Lesthaeghe, R., and P. Wilhems. 1999. Is low fertility a temporary phenomenon in the European Union? *Population and Development Review,* 25: 211-228.

Mason, K. 1997. Explaining fertility transitions. *Demography,* 34: 443-454.

Ryder, N. 1988. Observations on the history of cohort fertility in the United States. *Population and Development Review,* 12: 617-643.

Projecting Migration

Clark, D., and J. Cosgrove. 1991. Amenities versus labor market opportunities: Choosing the optimal distance to move. *Journal of Regional Scienc,* 31: 311-328.

Clark, D., and Murphy, C. 1996. Countywide employment and population growth: An analysis of the 1980s. *Journal of Regional Scienc,* 36: 235-256.

Clark, D., Knapp, T., and White, N. 1996. Personal and location-specific characteristics and elderly interstate migration. *Growth and Change* 27: 327-351.

Cushing, B. 1987. A note on the specification of climate variables in models of population migration. *Journal of Regional Scienc,* 27: 641-649.

DaVanzo, J. 1978. Does unemployment affect migration? Evidence from micro-data. *Review of Economics and Statistic,* 60: 504-514.

Edmonston, B., and J. Passel, J. 1992. Immigration and immigrant generations in population projections. *International Journal of Forecastin,* 8: 459-476.

Evans, A. 1990. The assumption of equilibrium in the analysis of migration and interregional differences: A review of some recent research. *Journal of Regional Scienc,* 30: 515-531.

Graves, P. 1980. Migration and Climate. *Journal of Regional Science* 20: 227-237.

Graves, P., and P. Mueser. 1993. The role of equilibrium and disequilibrium in modeling growth and decline. *Journal of Regional Science* 33: 69-84.

Greenwood, M. 1981. *Migration and economic growth in the United States: National, regional and metropolitan perspectives*. New York, NY: Academic Press.

Greenwood, M. 1997. Internal migration in developed countries. pp. 647-720 in M. Rosenzweig and O. Stark (Eds.), *Handbook of Population and Family Economics*. Amsterdam, Holland: Elsevier Science B.V.

Lee, E. 1966. A theory of migration. *Demography* 3: 47-57.

Long, L. 1988. *Migration and residential mobility in the United States*. New York, NY: Russell Sage Foundation.

Long, L. 1991. Residential mobility differences among developed countries. *International Regional Science Review* 14: 133-147.

Massey, D. 1999. International Migration at the dawn of the twenty-first century: The Role of the States. *Population and Policy Review* 25: 303-322.

Massey, D., J. Arango, G. Hugo, A. Kouaouci, A. Pellegrino, and J. Taylor. 1998. *Worlds in Motion: Understanding International Migration at the End of the Millenium*. Oxford, UK: Oxford University Press.

McHugh, K., Hogan, T., and Happel, S. 1995. Multiple residence and cyclical migration: A life course perspective. *Professional Geographer* 47: 251-267.

Meuser, P., and White, M. 1989. Explaining the association between rates of in-migration and out-migration. *Papers of the Regional Science Association* 67: 121-134.

Mohlo, I. 1986. Theories of migration: A review. *Scottish Journal of Political Economy* 33: 396-419.

Morrison, P. 1971. Chronic movers and the future redistribution of population: A longitudinal analysis. *Demography* 8: 171-184.

Nam, C. and D. Sly. (Eds.). 1990. *International Handbook on Internal Migration*. New York, NY: Greenwood Press.

Plane, D. 1989. Population migration and economic restructuring in the United States. *International Regional Science Review* 12: 263-280.

Plane, D. 1993. Demographic influences on migration. *Demography* 27: 375-383.

Rogers, A. 1992. Elderly migration and population redistribution in the United States. pp.

226-248 in A. Rogers (Ed.), *Elderly migration and population redistribution*. London, UK: Belhaven Press.

Rogers, A., and Castro, L. 1984. Model migration schedules. pp. 41-88 in A. Rogers (Ed.), *Migration, urbanization and spatial population dynamics*. Boulder, CO: Westview Press.

Smith, S., and H. Fishkind. 1985. Elderly migration in rapidly growing areas: A time series approach. *Review of Regional Studies* 15: 11-20.

Steinnes, D. 1982. 'Do people follow jobs' or 'do jobs follow people'? A causality issue in urban economics. *Urban Studies* 19: 187-192.

Vijverberg, W. 1993. Labor market performance as a determinant of migration. *Economica* 60: 143-160.

Projecting Socio-economic Characteristics

De Wolff, P. 1967. "Employment Forecasting by Professions." pp. 109-114 in United Nations, *World Population Conference*, 1965 (Belgrade), Vol. 111.

Fullerton, H. 1997. Labor Force 2006: Slowing down and changing composition. *Monthly Labor Review*.

Goldstein, H. 1966. "Projections of Manpower Requirements and Supply." *Industrial Relations* 5(3): 17-27.

Tilak, V. R. K. 1963. The Future Manpower Situation in India, 1961-76. *International Labor Review* 80: 435-446.

Projecting Households and Families

George, M.V. and J. Perreault. 1993. Projecting Households and Families in Canada by the Headship Rate Method. pp.385-402 in K. Mahadevan and P. Krishnan (Eds.) *Methodology for Population Studies and Development.* New Delhi: Sage Publications.

Van Imhoff, E., A. Kuijsten, P. Hooimeijer, and L. van Wissen. 1995. *Household Demography and Household Modeling*. New York, NY: Plenum Press.

Walkden, A. H. 1961. "The Estimation of Future Numbers of Private Households in England and Wales." *Population Studies* 15: 174-186.

Zeng, Y., J. Vaupel, and Z. Wang. 1998. Household projection using conventional demographic data. pp. 59-87 in W. Lutz, J. Vaupel, and D. Ahlbrrg (Eds.). *Frontiers of population forecasting*. New York, NY: The Population Council.