THE METHODS AND MATERIALS USED TO GENERATE TWO KEY ELEMENTS OF THE HOUSING UNIT METHOD OF POPULATION ESTIMATION: VACANCY RATES (VR) AND PERSONS PER HOUSEHOLD (PPH)

2 March 2009

Working Paper 09-03
Blakely Center For Sustainable Suburban Development
University of California Riverside
http://cssd.ucr.edu/Papers/Index.html

David A. Swanson
Department of Sociology
University of California Riverside
Riverside, CA 92151 USA
David.swanson@ucr.edu

Table of Contents

This report was prepared under contract to the U.S. Census Bureau, Order No. YA132307SE0374.

I. Introduction

Purpose of the Report

This report deals with the Housing Unit Method (HUM) of population estimation (Bryan, 2004b: 550). Specifically, it provides an examination of methods that can be used to estimate the average number of persons per household (PPH) and vacancy rates (VR), particularly at the county level. The research underlying the report compares the results of using the Census Bureau's current method for estimating these variables with alternative methods. The alternative methods proposed include using trended PPH and VR values as well as PPH and VR values informed by the American Community Survey. Methods selected for examination include the following model based methods: (1) simple and complex trend extrapolation from past decennial census counts (Swanson, Baker, and Van Patten, 1983; Lowe, Pittenger, and Walker, 1979; Smith, Tayman, and Swanson, 2001; Swanson, 1981); (2) regression and related model-based methods (National Research Council, 1980; Rives, 1982; Smith and Lewis, 1980; Smith, Nogle, and Cody, 2002; Swanson, 1980; Swanson and Beck, 1994; Voss, Palit, Kale, and Krebs, 1995); (3) survey-based methods (Griffin and Waite, 2006; Roe, Carlson, and Swanson, 1992; Swanson, 1989; Swanson, 2006); and (4) direct and indirect estimators from other sources (Swanson, Baker, and Van Patten, 1983; National Research Council, 1980; Lowe, Pittenger, and Walker, 1979; Rives, 1982; Smith and Lewis, 1980; Swanson, 1981; Swanson and Lowe, 1979).

The report includes an evaluation of the methods offers recommendations for usage (Chapter IV). It includes a description of each method, its underlying theory and logic (Chapter II) and an example, where feasible, in the form of an illustrative calculation (Chapter III). Each method is assessed relative to a set of evaluation criteria (e.g., face validity, cost, ease of application and explanation, and accuracy) developed using ideas from, among others, Habermann (2006), Swanson, Baker, and Van Patten (1983), Hough and Swanson (2006), National Research Council (1980), Roe, Carlson, and Swanson (1992), Smith and Mandell (1984), Smith, Nogle, and Cody (2002), Smith, Tayman, and Swanson (2001: 279-299), Swanson (2006), Swanson (1989), Swanson (1981), Swanson and Lowe (1979), Swanson and Pol (2005), Swanson and Tedrow (1984), Swanson and Van Patten (1987), and Swanson, Tayman, and Barr (2000).

Context

An important part of the context for this report is the reason why estimates are done in the United States. The census is the most complete and reliable source of information on the number of people in the United States a census, however, is a time-consuming and costly endeavor. The 1990 U.S. census had a cost of about $2.5 billion (U.S. GAO, 2001). The cost of the 2010 U.S. census is estimated at $11.3 billion U.S. GAO, 2006). In the United States, a census of the population is done only once every ten years; in Australia, Canada, England and New Zealand, for example, it is once every five years. Because there is the potential for constant and sometimes quite rapid population change, especially at the sub-national level, census statistics for every tenth and even every fifth year are often inadequate for many purposes (Waldrop, 1995). To fill this gap, population estimates are used by government officials, market research analysts, public and private planners and others for determining national and sub-national fund allocations (Murdock and Ellis, 1991; Serow and Rives, 1995; Siegel, 2002), calculating denominators for vital rates and per capita time series, establishing survey controls, guiding administrative planning, marketing, and for descriptive and analytical studies (Long, 1993a; Pol and Thomas, 2001: 93-95; Swanson and Pol, 2005). In the United States, the Census Bureau is not the only provider of population estimates (Bryan, 2004b: 524-526), but it is the ultimate source of estimates and the data needed to develop them.

The development of methods of population estimation roughly corresponds to the development of censuses and vital statistics registries. For example, in the late 17th century, John Graunt estimated the population of London and then of the whole of England and Wales using what today is known as a censal-ratio method (Devlin 2008: 93-94). Not long afterward, in the 18th century, the French mathematician, Laplace, also used a censal-ratio method in combination with recorded births and a population sample to estimate the population of France (Stigler, 1986: 163-164). However, methodological development really only took off in the late 1930s and early 1940s, fueled in large part by the need for low-cost and timely information generated by the great depression of the 1930s and World War II (Bryan, 2004b; Eldridge, 1947; Hauser and Tepping, 1944; Shryock, 1938; Shryock and Lawrence, 1949). In the United States, the Census Bureau played a major role in this effort, but it was not alone. During the early 1940s, the Washington State Census Board, for example, developed a comprehensive program of annual population determinations based on estimation methods that are still used today (Swanson and Pol, 2005; Swanson and Pol, 2008). Around this same time, demographers also began developing estimation methods for what were then called "underdeveloped countries," (Brass, 1968, Chandrasekaran and Deming, 1949; Popoff and Judson, 2004; United Nations, 1969) and the use of sample surveys as a substitute for complete census counts took hold (Bryan, 2004b).

Today, population estimates are ubiquitous. They are done around the world by a host of governmental and non-governmental entities, as well as individual consultants (Bryan, 2004b;

Siegel, 2002; Swanson and Pol, 2008). The widespread availability of data, methods, and technology has made is possible for many people not only to develop estimates, but to do so more quickly and less expensively than has ever been done before. This trend is not likely to abate, but it carries with it a cost in that estimates may both be made and used with little or no understanding of the issues involved, what constitutes good estimates, and how to identify them.

What is a Population Estimate?

A population estimate is the determination of the size or the characteristics of a population at a current or past date in the absence of census data for the same date. In the United Sates, they usually are made on a De jure basis, which means that people are estimated where they usually reside. This makes sense because the U.S. census is conducted on a De Jure basis. However, there also is a need for estimating the De facto population of a given place at a given time and researchers have developed these estimates (Swanson and Pol, 2005; Swanson and Pol, 2008). These estimates include vacationers (of interest, for example, to the casino industry in Las Vegas and the Hawaii Visitors Bureau), migratory workers (of interest, for example, to health care, school, and other social service providers), and the people who work in the central business district of a large city each day, but leave it largely vacant in the evenings (of interest to the San Francisco City Planning Office, for example). While estimates of de facto populations are of great interest, they are very difficult to make in the United States because of the lack of census benchmarks (Cook, 1996, Smith, 1994). This is an important estimation topic, but it is beyond the scope of my mandate to cover research needs for de facto populations in depth. I only suggest here that the U. S. Census Bureau is the logical agency to develop systematic and comprehensive estimates of De Facto populations in the United States – as are its sister agencies in other countries currently operating similar population estimation programs.

In this context, it is useful here to note that the American Community Survey uses what amounts to a De Facto population as its sample frame and therefore is essentially aimed at estimating the De Facto population and its characteristics earlier (National Research Council, 2006a). However, while the ACS is essentially based upon a De facto population, it is important to note that its raw numbers are then controlled to a De jure population (National Research Council, 2007; Swanson and Hough, 2007). I return to this topic later in the report.

An estimate can be prepared for a nation or a sub-national area such as a state, county, city, town, or census tract. An estimate also can be prepared for groups of sub-national areas, groups of nations, or even the world as a whole. The principal demographic characteristics for which an estimate is made include age and sex. However, in multiracial and multi-ethnic countries such as the United States and Canada, an estimate might be done not only by age and sex, but also by race and ethnicity. An estimate also can be made of social and economic sub-groups of the population, households, and families.

The term "population estimate" is frequently used in the public domain to refer to the determination of the size or the characteristics of a population at a future date. However, most demographers prefer to use the term projection when talking about the possible size and characteristics of a population in the future. In developing a portrait of a given population in the future, it is not uncommon for a series of projections to be made that incorporate a range of plausible assumptions (e.g., expected trends in fertility, mortality, and migration). However, when one of these projections is selected as representing the most likely future, it then becomes the forecast for the population in question. As opposed to a projection or a forecast, a population estimate is concerned with either the present or the past, but not the future (Smith, Tayman, and Swanson, 2001: 3-4). Thus, in this report, I make the following distinctions among the terms "estimate," "projection," and "forecast."

> ***Estimate*** – A calculation of a current or past population, typically based on symptomatic indicators of population change.
>
> ***Projection--*** The numerical outcome of a particular set of assumptions regarding future population trends.
>
> ***Forecast*** – The projection deemed most accurate for the purpose of predicting future population.

Virtually all methods of population estimation can be categorized into one or the other of two traditions: (1) demographic (Bryan, 2004b); and (2) statistical (Kordos, 2000; Platek, Rao, Sarndal, and Singh, 1987; and Rao, 2003). Demographic methods are used to develop estimates of a total population as well as its ascribed characteristics, age, race, and sex. Statistical methods are largely used to estimate the achieved characteristics of a population, and include, for example, educational attainment, employment status, income, and martial status. As is the case in the national statistical agencies of other countries, the US Census Bureau produces estimates using both of these traditions, demographic and statistical.

Demographers and statisticians have developed estimation a wide range of methods designed to meet different information needs at varying levels of accuracy and cost. As noted earlier, for the most part they are based on the concept of a De jure population although there are exceptions (Swanson and Pol, 2005). The methods can be roughly placed into three categories: (1) analytical and statistical models that use data symptomatic of population and its changes; (2) mathematical models that use historical census data; and (3) sample surveys. Methods falling into the first category have generally been developed by and for applied demographers, many of whom work for national, state, and local governments. Methods falling into the second category have generally developed by and for academic demographers, most of whom work at universities and research institutes. The methods falling into the third category have generally been developed by and for statisticians and survey research scientists, but they also are widely used by demographers. Not surprisingly, there also are techniques that combine methods from two or even all three categories.

Population estimation methods also can be identified along a temporal dimension: (1) inter-censal estimates, which refer to a date between two census counts and usually take the results of both counts into consideration; (2) post-censal estimates, which refer to a date subsequent to the latest census count and usually into account one or more previous census counts; and (3) pre-censal estimates, which refer to a date prior to a census count, but usually take into account one or more subsequent census counts. This temporal classification is useful because different methods are typically employed in the development of inter-censal, post-censal, and pre-censal estimates (Bryan 2004b).[1] These definitions and distinctions fall into the demographic tradition.

Among survey statisticians, the demographer's definition of an estimate is generally termed an "indirect estimate" because unlike a sample survey, the data used to construct a demographic estimate do not directly represent the phenomenon of interest (Swanson and Stephan, 2004: 758 and 763). In this paper, I use the demographic tradition's definitions and distinctions unless specifically noted.[2]

There are other ways to classify estimation methods. John Long (1993a), for example, categorizes them generally into two types: (1) "flow" methods; and (2) "stock" methods. Flow methods are also known as component methods, because they require estimation of each component of population change (births, deaths, and migrants) since the last census. Stock methods relate changes in population size since the last census to changes in other measured variables: the number of housing units, automobile registrations, total number of deaths (and births), and tax returns. Long (1993a) also notes that stock and flow methods may be used in combination. Popoff and Judson (2004: 603), make the following useful distinctions between stocks and flows: "…stock data are the numbers of persons at a given date, classified by various characteristics…(and) are recorded from censuses….flow data are the collection of or summation of events. At the most basic level this includes births, deaths, and migration flows…." This distinction is useful for purposes of this paper because, as is discussed later in this section, there are population estimations methods that solely rely on "stock" data while others rely on a combination of "stocks" and "flows."

Finally, it is useful here to consider micro data and aggregated data in the context of population estimation methods. I take micro data to mean records for individual persons. These records are often linked by relationships to form family and household records and I use the term "micro data" to refer to these linked records as well. The "Public Use Microdata Sample" (PUMS) is such a file (Swanson and Stephan, 2004: 772). Aggregated data are summations of records of individuals (families and households) such as one would find in a table. The aggregations are often done to specific geographic areas, but they can also be done for types of people across different geographies. The life table constructed by Kintner and Swanson (1994) for retirees of General Motors is an example of such an aggregation.

The Housing Unit Method

The HUM is designed to generate estimates of the total population by focusing on the population residing in households. As such, it inherently fits within the demographic tradition. However, while the HUM is inherently demographic in nature, the two key HUM elements I cover in this report (PPH and VR) are generated using methods that fit within the statistical tradition. Thus, I cover both traditions in discussing the HUM. Given that The HUM is aimed at the population residing in households, it is easy to see that is used to generate estimates of the total "De Jure" population. This, of course, is the definition of population used by the US Census Bureau, which is based on place of "usual residence" (Cook, 1996; Wilmoth, 2004).

One of the first times that the HUM is mentioned in the literature is found in an article by Starsinic and Zitter (1968) who found that it made a "…surprisingly strong showing…" and that "…it may be worthwhile to devote considerably more effort to refining the input data for estimating the number of households in addition to dealing with the problem of deriving current estimates on the size of households" (Starsinic and Zitter, 1968: 484). The article mentions work by Carl Frisén (1958) on the HUM in the 1950s for the California Department" The work by Frisén involved testing methods of population estimation against special censuses done by the state of California during the 1950s. Earlier work along these lines by was reported by Frisén (1951) when he was at San Jose State University.

However, the HUM was used even before 1950. It was used as early as 1942 under the auspices of the Washington State Census Board, which utilized the sociology graduate program at the University of Washington to carry out a program of annual estimates for cities and towns under the overall direction of Professor Calvin Schmid (Lowe, 2009). In 1967, the operation was transferred to the Washington State government, where today it exists in the state's Office of Financial Management (Lowe, 2009). Washington's use of the HUM is done in conjunction with census counts that allows cities and towns to conduct a special 'headcount' census when disagreements over estimates arise (Washington Office of Financial Management, 1978). These census counts are conducted in accordance with residency and housing definitions used by the Census Bureau with training assistance and supervision (including auditing) from the Washington Office of Financial Management. In 1981, the Washington system of municipal population estimation was adapted by the state of Alaska (Alaska Department of Community and Regional Affairs, 1981a, 1981b; Alaska Department of Labor, 1981, 1982; Swanson, Baker, and Van Patten, 1983). Today, the HUM is arguably the most commonly used method of population estimation in the United States (Bryan, 2004b).

The Housing Unit Method (HUM) is a "stock" method that describes a basic identity in the same way that the balancing equation does (Bryan, 2004b). In the case of the HUM, this identity is usually given as

$$P = H*(1-VR)*PPH + GQ \qquad [1]$$

where P = Population,

H = Housing units,

VR= Vacancy Rate (Proportion Vacant),

PPH = Average number of persons per household, and

GQ = Population residing in "group quarters" and the homeless.

Like the balancing equation, the HUM equation can be expressed in less detail (i.e., P= HH*PPH + GQ, where HH=H*(1-VR), Smith and Cody, 2004: 2) or more detail - by structure type, for example (Swanson, Baker, and Van Patten, 1983). It also can be used in combination with sample data, which opens the door to developing measures of statistical uncertainty for the estimates so produced (Roe, Carlson, and Swanson, 1992).

The HUM is based on the assumption that virtually everyone lives in some type of housing structure. It is generally accepted that the HUM is the most commonly used method for making small area population estimates in the United States (Byerly, 1990; Smith, 1986; Smith, Nogle, and Cody, 2002). Because of how data are collected, the HUM had not been a method that could be used for all sub-national areas and the nation as a whole until recently. However, with the continuous "Master Address File" (MAF), it has now emerged as a method that can be used by the US Census Bureau for all sub-national areas and the nation as a whole (Wang, 1999). This is a new resource for the Census Bureau's estimates program because in the previous "mail-out/mail-back censuses, the MAF was constructed from scratch before each census. As observed nearly 25 years ago by Pittenger (1982) and more recently by Wang (1999), this housing unit inventory is serves as a key resource in the Bureau's ability to construct population estimates (Swanson and McKibben, 2009) .[3] Other resources in regard to the suitability of the MAF, include Perrone (2008), Reese (2006), Swanson and McKibben (2007), U.S. Census Bureau (2004a), U.S. Department of Commerce (2002), and U.S. GAO (2006). Related work on the development of housing unit information includes McDonald and McMillen (2000), Pittenger (2004), and the U.S. Census Bureau (2006, 2007).

In testimony before Congress, Swanson (2006) advocated the use of the HUM by the Census Bureau. In so doing, he: (1) described what he believed was the major challenge faced by the Census Bureau in providing timely, accurate, and cost-effective estimates; (2) provided a suggestion for dealing with this challenge; and (3) identified the major issues presented by his suggestion that need to be resolved.

In regard to the major challenge faced by the Census Bureau in providing timely, accurate, and cost-effective estimates, Swanson observed that this stemmed from the proliferation of federal programs distributing benefits using decennial census data and the knowledge that federal courts were now willing to consider apportionment cases. He noted in this regard that several lawsuits were filed against the Census Bureau following the 1970 census, a practice that has proliferated over the past thirty years and threatens to move into other areas of the Census Bureau's work such as the annual estimates program. Swanson argued that the reason for much of this conflict is clear: Billions of federal dollars are allocated each decade to states and local governments using census counts and inter-censal estimates and these funds are allocated in a "zero-sum" fashion. This situation will lead to even more litigation and other forms of conflict as the states, cities, and counties struggle to get their "populations" counted in the decennial censuses and estimated during the inter-censal periods.

It is this "atmosphere of conflict" that Swanson(2006) believes is the major challenge facing the Census Bureau's decennial census and inter-censal estimates programs. Within the Census Bureau it not only serves to foster a "defensive" working environment, but also takes important resources away from production and research activities. As the defensive climate within the Bureau hardens, states and local governments feel even more frustration in their attempts to work cooperatively with the Bureau and turn to more confrontational forms of communication. This is particularly attractive for the local governments in states lacking strong demographic centers.

Swanson's suggestion for dealing with the challenges facing the Census Bureau revolves around the MAF. He noted in his testimony that, breaking with the past, the Census Bureau decided to retain and update its Master Address File – the MAF - for the 2000 Census. The MAF is a critical resource for the American Community Survey and its retention facilitates the planning and conduct of an accurate and cost-effective 2010 census. The continuously updated MAF and the related TIGER improvements are a fundamental element of success for an accurate 2010 census. Importantly, Swanson argues, the continuously updated MAF also represents an untapped resource for inter-censal estimates. It leads directly to the potential to have timely, accurate, and cost-effective estimates done using a method that is not only simple to apply and explain, but one that offers the potential for a meaningful role for states and local governments to play in the development of these estimates. He identified this method as the HUM, but to be successful this approach he argued that it needs a nationwide system of state demographic centers that participates in a meaningful partnership with the Census Bureau. He also notes that the state demographic centers, in turn, would need an active and meaningful partnership with the local governments within their respective states. Swanson (2006) argued that MAF-based population estimates would contribute toward having more timely, comprehensive, and internally consistent demographic and housing data for the U. S. as a whole and its sub-areas. In regard to geography, he notes that MAF- based data are extremely flexible in that they can be geo-coded to a specific location (as opposed to being assigned to an area defined by administrative or statistical boundaries). This also means that the MAF-based system can be overlaid with other features using GIS capabilities. The TIGER street address file comes to mind first in this regard.

Swanson (2006) testified that this approach to inter-censal population estimation would lead to an entirely new way of looking at the concept of a "small area," in that boundaries could be drawn that are much finer than those allowed by the census defined block. However, he noted that this would allow much higher precision in defining areas for purposes of marketing, site location and micro-simulation analysis, and modeling. Once up and running, this would also allow for greater ease in producing a consistent time series for areas in which administrative boundaries changed over time. The estimates would also provide population controls for the American Community Survey.

Swanson (2006) identified three major issues that needed to be resolved if his suggestion were to be successfully implemented. The first was confidentiality (e.g., Title 13 requirements), not an insignificant problem. However, he argued that this problem is not insurmountable in regard to his proposal for a MAF-based population estimation system. The National Research Council (2006b) has issued recommendations to reconcile access and confidentiality and the Census Bureau itself has appointed a Chief Privacy Officer and worked to put effective procedures in place regarding this reconciliation. Thus, he believes that the Census Bureau is capable of creating a national MAF-based population estimation system that meets confidentiality concerns.

Another important obstacle identified by Swanson (2006) is the financial cost of developing a national system of state demographic centers such that each state center functions according to accepted standards. States need to shoulder a share of these costs. After all, it is to their benefit to have high quality state demographic centers. As such, he proposed that a funding mechanism involving federal-state matching funds be considered.

Swanson (2006) also rhetorically asked if the proposed MAF-based population estimation system could provide accurate data and noted that the GAO identified MAF/TIGER problems that needed to be solved in order to have a good census in 2010. These problems include: (1)

resolving address related issues such as duplication, omission, deletion, and incorrect locations in the MAF; and (2) implementing GPS-based geo-coding of housing units. Swanson noted that these same two problems represent sources of error in the proposed MAF-based system. Consequently, if the Census Bureau solves these problems in regard to the 2010 census, Swanson testified that it would do much in regard to the accuracy of the proposed MAF-based population estimation system.

Swanson (2006) also identified some lesser problems, ones largely already known to Census Bureau staff and others in regard to using the HUM effectively, to include: (1) tracking new housing units, converted housing units, and deleted housing units; (2) dealing effectively with seasonal populations and seasonal housing. He also notes that with the implementation of the ACS, the seasonality problem is compounded because of differences between the ACS and the decennial census in regard to what constitutes the de jure population. As such, he observed that an accurate MAF-based population estimation system will need to deal with the seasonal housing issue and the differences in the definition of the de jure population found in the ACS and the decennial census. However, Swanson(2006) testified that given the experience being gained by Census Bureau in regard to the MAF/TIGER system, the widespread knowledge use of the Housing Unit Method, and the capabilities of the best of the State Demographic Centers – Alaska, California, Florida, Texas, and Washington, for example, the timeliness and accuracy of MAF-based population estimates based on a comprehensive system of state demographic centers functioning at the level of the best state demographic centers would be sufficient for purposes of resource allocation, research, decision-making, and planning for the national, state and local levels. He also testified that he believed that it would also prove to be cost-effective and equitable and noted that the conflict-free system used in Finland to produce annual population data has the type of state-national participation and cooperation that he proposed, even though Finland uses population data to distribute funds and other resources in a zero-sum fashion to regional and local governments.

Swanson (2006) concluded his testimony with the observation that with the exception of the issue of confidentiality, all of the challenges facing the development of a national MAF-based population estimation system are in the form of costs, technical problems, or a combination of both. The major technical tasks in building and maintaining a MAF-based population estimation system come down to two areas - address data collection and MAF/TIGER update. The feasible way to effect a solution to these problems is to enhance the federal-state-local cooperative programs already part of Census Bureau activities such that local entities are compensated for helping to maintain the system. There are data collection activities in the United States that already follow this model, such as the vital registration system. He also noted in conclusion that by exploiting a functional MAF with the HUM that the Census Bureau and its state and local government partners would have universal means of population estimation for all areas of geography, administrative and statistical, and that state demographic centers should be developed to a uniform level of capability. Thus, he suggested that this proposal be supported by state-federal matching funds. He summarized his testimony bay stating that he believed that this system would lead not only to timely, accurate and cost-effective inter-censal population estimates, but also to greater equity in that there would be a uniformly higher level of demographic human capital in the country.

Data Foundations for Population Estimates.

As can be surmised from the preceding discussion of the MAF and other sources of housing unit data, all estimates, including the HUM, rely on one or more censuses and use administrative record systems on which different estimation methods for census-defined populations rely – vital events, tax returns, housing permits, assessor parcel files, utility hookups, licensed drivers, covered employment, K-12 enrollment, Medicare, and child support payments, among others (Bryan, 2004a; Bryan, 2004b; Bryan and Heuser, 2004). It is important to note that there is some variation in availability and quality of administrative records systems by state and by local jurisdictions in the US as well as variation among countries. For example in many areas of the United States, Kindergarten through 8th grade enrollments are used in the calculations of population estimates to avoid mistaking students who drop out of high school as out migrants from the area (McKibben, 2006).

It also is important to note that the U. S. Census Bureau maintains as much consistency in data sources and methods as it can because among other desirable features, it wants to have a consistent set of estimates for a given "vintage" year (See, Habermann, 2006; Appendix A of this report, U.S. Census Bureau, n. d.).

General Concepts and Methods

Although it is not used directly in the HUM, the fundamental demographic identity known as the balancing equation forms the conceptual framework for most other methods of population estimation (and forecasting). This identity is defined as $P_t = P_0 + I – O$, where P_t is the given population at time 0 + t, P0 is the given population at time 0, I is the number of persons entering the population through birth and in-migration during the period 0 – t, and O is the number of persons exiting the population through death and out-migration during the period 0 – t (Swanson and Stephan, 2004: 753). This identity can be phrased in more detail to separate recognize births, deaths, in-migration, and out-migration and is used as a point of departure to discuss in detail the concept of "stocks and flows" and the measurement thereof encompassed in the following five general types of methods, not including the HUM.

Simple Interpolation and Extrapolation Methods. Although no longer widely used in their own right, interpolation methods (see, .g., Judson and Popoff, 2004) and extrapolation methods (see, e.g., Smith, Tayman, and Swanson, 2001) represent ways to construct, respectively, inter-censal estimates and post-censal estimates. These methods range from being relatively simple (e.g., linear trending) to very complex (ARIMA models). Both interpolation and extrapolation are based on mathematical formulas that are applied to "stock" data to produce "flows" that, in turn, generate estimates. As such, the principles underlying these methods, particularly extrapolation, are often found in other estimation methods (e.g., regression methods).

Regression Methods. This approach to population estimation represents a "stock" method in which measures of change in the ratios of indicators to population are used as "flow" estimates that are extrapolated to generated population estimates (Bryan, 2004b). The flow estimates serve as independent variables in these forms, while the dependent variable is a measure of population change. Measures of change can be in the form of ratios, lagged ratios, and differences (Bryan 2004b). These regression methods require a nested set of geographies (e. g., the counties within a given state) and they are inherently embedded in statistical inference (Swanson, 2004). As observed by Prevost and Swanson (1985), the "ratio-correlation" form can be viewed as a regression-based version of the so-called "synthetic" method of estimation.[4]

Component Methods. These are directly based on the fundamental demographic identify known as the balancing equation. As such, they are stock and flow methods. Included in this set are "Component Method II," "Cohort-Component Method," and the "Tax Return Method," each of which is described by Bryan (2004b). The stock data are comprised of census counts in each of these methods, which use administrative records (e. g, vital events) to develop flow estimates.

Administrative Records. So-called direct estimates can be acquired from selected types of administrative records systems, namely the national population registration systems found in the Nordic countries (Bryan, 2004a: 31-33; Statistics Finland, 2004). Although the United States lacks a national population registration system, it has several national administrative record systems that serve as partial population registers, including those relating to social insurance and welfare and the payment of income taxes (Bryan, 2004a; Judson, 2000).[5] It is worthwhile at this point to consider the MAF, which represents a national housing registration system that can be used to generate estimates using the Housing Unit Method (Swanson, 2006).

Other Methods (Not including the HUM). Here, I include the economic-demographic models and urban systems models described by Smith, Tayman, and Swanson (2001: 185-237) as well as the iterative proportional fitting, log-linear, and multiregional methods described by

Judson and Popoff (2004). To this list can be added the methods developed for statistically underdeveloped countries and those for estimating wildlife populations (briefly discussed in Endnote # 2) as well as the imputation methods used by the US Census Bureau to compensate for missing data (see, e.g., Swanson and Stephan, 2004: 762).

In concluding this brief overview of methods of population estimation other than the HUM, I note that it is often the case that various data adjustments must be made to effectively operate the preceding methods and that these adjustments serve as "other methods" in themselves (Wang, 1999). For example, the presence of non-household populations, such as found in prisons, school dormitories, and long-term care facilities, can affect the accuracy of virtually all of the methods just described, as can the presence of seasonal populations, undocumented aliens, and the occurrence of disasters, natural and otherwise.[6]

Organization of this Report

The remainder of the report consists of four chapters, endnotes, references, and two appendices (one of which has its own endnotes). The following Chapter (II) provides descriptions of the theory, logic, data, and assumptions found in the methods used to generate PPH and VR. It includes, as appropriate, critical commentary. Chapter III consists of illustrations of the calculations underlying the methods and the data to produce PPH and VR estimates, along with the theory and logic underlying them, and the steps they use. For some methods, such illustrations are not feasible (e.g., a windshield survey). Chapter IV provides an evaluation of the methods identified in Chapter III, along with recommendations and suggestions. Endnotes and references follow Chapter IV.

Appendix A is a reproduction of the principles underlying the US Census Bureau's estimates and projections programs (Habermann, 2006; U.S. Census Bureau, n.d.). Appendix B consists of two parts: (1) an annotated bibliography of population estimation methods, with a focus on the HUM and methods for estimating PPH and VR; and (2) a glossary of demographic and statistical terms relating to populations estimation methods, with a focus on the HUM. Appendix B contains its own endnotes.

While the report is focused on the United States at least some of what it covers applies to certain other countries that, like the United States, have strong administrative record systems, but lack a population registry system and rely upon regular census counts for population information. Like the United States, these countries are largely English-speaking and include, among others, Australia, Canada, England, and New Zealand. It also is worth noting that the report goes beyond the HUM at times and discusses other methods for estimating population.

II. Methods used to Generate VR and PPH Data

In the preceding chapter, I mentioned the HUM study reported by Starsinic and Zitter (1968) who suggested that it was likely to be worthwhile to devote some effort to improving estimates of the number of households and their average size. There are two variables required to estimate the number of households: (1) housing units; and (2) a vacancy rate. By multiplying the number of housing units by the vacancy rate and subtracting this product from the number of housing units one arrives at the number of occupied housing units, which is synonymous with the number of households. So, I start with vacancy rate estimation.

Vacancy Rates

Following up on the suggestion by Starsinic and Zitter (1968) to work on ways to generate household numbers, Lowe, Pittenger, and Walker (1977) describe a method for generating vacancy rates using "windshield surveys." This method is labor-intensive and time-consuming (Alaska Department of Labor, 1981; Lowe, Pittenger, and Walker, 1977); Swanson, Baker, and Van Patten, 1982; Washington Office of Financial Management, 1978), but given the sample survey operations already being conducted by regional offices, it is potentially feasible for the Census Bureau. I return to this in more detail in the next chapter and in the final chapter

There are three other potential methods for obtaining vacancy rates that also are viable for the Census Bureau: (1) holding vacancy rates from the most recent decennial census constant until the subsequent census; (2) the use of US Postal Service (USPS) delivery data; and (3) modeling, which could also use the vacancy rates from the most recent decennial census and the

USPS delivery data as inputs. In terms of holding vacancy rates constant since the last census, there is not much to describe, other than being aware of boundary changes and the use of structure type classifications.

In regard to USPS delivery data, Theresa Lowe (1988) examined the accuracy of postal survey data in reporting residential housing unit occupancy estimates against vacancy rates found in the 1970 and 1980 U.S. decennial census counts for 26 Washington State cities. The postal surveys were conducted by the Federal Department of Housing and Urban Development in the 1970s within 2 months of collection of census data. She found that postal surveys almost always show lower vacancy rates than census data because they do not include unfinished or new units, or concealed unoccupied conversions in single family homes. Suburban single family housing generally had the highest occupancy rates. However, she also found that postal data were much more accurate than census data in areas where occupancy rates were subject to high variation, as is found, for example, in cities near military bases, and in multi-unit structures. Because of this variation, Lowe (1988) argued that it was difficult to model vacancy rates in such areas.

Lowe (2000b) subsequently examined real estate vacancy surveys, which are aimed at the market for apartment rentals (multi-unit structures) and found that because they do not use random sample procedures, they did not match up well with the vacancy rates found in a decennial census, the later typically showing higher rates of vacancy. She suggested that the tendency of real estate vacancy surveys to be lower was primarily due to two factors: (1) many 'rented' units are not 'occupied' in the same manner that the census defines occupancy; and (2) the surveys only cover apartment units that are currently on the rental market (excluding, among other things for examples, newly constructed units that according to census definitions are unoccupied housing units). In her examination, she found that real estate vacancy rates tended to be around five percentage points lower than equivalent census vacancy rates and she provides adjustment factors for real estate vacancy surveys of multi-unit structures so that they more closely match the equivalent census vacancy rates.

Lowe, Mohrman, and Brunink (2003) examined postal delivery data within a context of factors affecting vacancy rates in general using the 2000 census of Washington as a benchmark. Acknowledging that United States Postal Service (USPS) delivery data recognize postal deliveries rather than housing units, they found that for the state of Washington as a whole, residential postal delivery data exceeded the 2000 census count of 2,451,075 housing units by 7.6 percent. However, when post office box deliveries were excluded, they found that residential deliveries fell about 7.1 percent units short of the 2000 census count of housing units for the state as a whole. When looking at Washington's 39 counties, they found, however, that metropolitan counties had lower differences than did non-metropolitan counties. Considering the 2,001,325 housing units counted in the 2000 census for metropolitan counties, the postal delivery data were 7.3 percent higher when all deliveries were included and -3.3 percent lower when post office box deliveries were excluded. Considering the 449,750 housing units counted in the 2000 census for non-metropolitan counties, they found that the postal delivery data were 9.2 percent higher when all deliveries were included and 24.1 percent lower when the post office box deliveries were excluded.

Moving on to vacancy rates themselves, Lowe, Mohrman and Brunink (2003: 5) note that postal delivery data recognize deliveries as "possible" and "active." "Active" deliveries are reported within "possible" deliveries so by subtracting "active" from "possible," to obtain a residual set, "possible, but not active" that corresponds roughly to vacant units (Lowe, Mohrman, and Brunink, 2003: 5). Carrying out these operations, they compare the "possible, but not active" set to vacant housing units by county in Washington using 2000 data. They find that for the state as a whole, the 2000 census found a housing unit vacancy rate of 7.33 percent while the comparable rate from the USPS data was 1.78 percent when all deliveries are used and 1.33 percent when post office box deliveries are excluded. Following the state as a whole, Lowe, Mohrman, and Brunink (2003: 6-8) the 2000 USPS data produce, on average, estimates of vacancy rates that are 11 to 12 percentage points lower than the vacancy rates from the 2000 census for Washington's 39 counties. They find that the USPS data are only about 2 percentage points lower than the census vacancy rates in metropolitan counties, however. The largest arithmetic differences are found in counties that have substantial seasonal housing, which are non-metropolitan and that USPS data are, on average, 12-13 percentage points lower than the census vacancy rates across all non-metropolitan counties. Lowe, Mohrman, and Brunink (2003) also examined changes in the USPS data subsequent to the 2000 test and found that they were in accordance with expected vacancy rate changes due to population and housing changes. They conclude that USPS data may be a useful tool for 'adjusting" (modeling) decennial census vacancy rates at the county level. However, they advise that counties be examined individually in accordance with metropolitan/non-metropolitan classifications and the presence of substantial seasonal housing stock, among other variables.

Lowe and Mohrman (2003) extend the research reported by Lowe, Mohrman, and Brunink (2003) by examining the consistency of 2002 HUM-based county population estimates using USPS adjustments with 2002 population estimates made by the US Census Bureau. They used all possible residential deliveries, including post office boxes and vacancy rates from the 2000 census were adjusted on a county-by-county basis. They fond that the mean algebraic percent difference (MALPE, which includes the sign of the percent difference) across all 39 counties in 2002 was only 0.14 percent and that in 17 counties the HUM-based estimates exceeded the Census Bureau estimate and that in 22 counties they were lower. The highest positive difference (2.80%) was for the non-metropolitan county of Garfield, which is not adjacent to a metropolitan county; the highest negative difference (-5.39%) was for Island County, which is adjacent to the metropolitan county of Snohomish. Lowe and Mohrman (2003) also ‘backed-into” the USPS adjustments that would be required to match the “most likely” populations of these counties in 2002 (which were a combination of state and Census bureau estimates, accounted for the population in group quarters, and maintained 2000 PPH values). They found that at the state level, a 25.2 percent change was required of the USPS data and that most counties required between a 20 and 40 percent change. They concluded that the process they used needed to be extended to more years subsequent to 2000 to assess the stability of the relationship between the (assumed) underlying actual vacancy rates and the rates derived from the USPS data.

Moving away from USPS delivery data, Swanson, Baker, and Van Patten (1983) discuss vacancy rates within the context of an overall assessment of the HUM by state demographic centers. They point out that the HUM is optimal when it is done in conjunction with an active “headcount” census program, which can be used to update the elements of the HUM, including VR and PPH, if there is a dispute between the state demographic center and local agency preparing a population estimate using the HUM. As was described earlier, they also point out that reasonably estimates of vacancy rates can be obtained from “windshield surveys, for which detailed procedures are given in the Housing Unit Method Manual produced by the Alaska Department of Labor (1981), which was adapted from a similar manual developed by the Washington Office of Financial Management (1978).

Smith and House (2007) suggest the use of the ACS, but not for directly estimating vacancy rates. Instead, they argue for using it to develop estimates of 'temporary' migrants, a topic that affects areas with seasonal populations and, hence, vacancy rates of seasonal housing units. An important issue in considering the ACS as a source of VR data is the fact that it does not use the same residency definition as the decennial census (Swanson and Hough, 2007; Swanson and McKibben, 2009). The decennial census (along with the CPS, and SIPP, among other products), use the De jure rule of residency while the ACS uses what amounts to a De facto residency rule. However, it is useful to note here that when the micro-level ACS data are aggregated to geographic areas, they are controlled to number produced for these areas by the Census Bureau's annual population estimates program, which are produced on a De Jure basis. As Swanson and McKibben (2009) observe, this may not be a huge issue at the national level, but at sub-state levels, the effects of these different residency rules could be substantial.

Most of the "modeling" techniques for developing estimates of occupied and vacant residential units (and the corresponding rates) come from the field of housing economics. Edelstein and Tang (2007) develop and test a theoretical model for residential housing market cyclical dynamics. The model employs an interactive supply and demand framework to engender housing price dynamics. Under our set of assumptions, the two equation system is econometrically identified: the first equation, housing demand, relates rent, property values, and capitalization rates with demand fundamentals. The second equation, housing supply, relates housing investment and property values with supply fundamentals. Edelstein and Tang (2007) use their model to analyze empirically the cyclical dynamics for residential properties in Los Angeles, San Francisco, San Diego and Sacramento for the 1988–2003 time period. The authors argue that their empirical analyses suggest that fundamentals, such as employment growth and interest rates, are key determinants of residential real estate cycles, but that local fundamentals tend to have greater cyclical impacts than those of national or regional fundamentals.

Gabriel and Nothaft (2001) make use of inter-metropolitan and time-series data from the BLS to model the incidence and duration of rental vacancies and to assess their importance to the price adjustment mechanism for rental housing. They find that duration varies with measures of MSA housing costs and housing stock heterogeneity, while incidence varies with measures of population mobility, public housing availability, and population growth. Results support a more general specification of rental price adjustment in which the rate of real rent change reflects deviations in observed vacancy incidence and duration from their equilibrium levels.

Hendershott, MacGregor and Tse (2002) note that rental adjustment equations have been estimated for a quarter century and that in the U.S., models have used the deviation of the actual vacancy rate from the natural rate as the main explanatory variable, while in the UK, drivers of the demand for space have dominated the estimation. They derive a more general model incorporating both supply and demand factors and find that it is greatly superior to the vacancy rate model. They also construct a variant of the general model with a separate vacancy rate equation and find that it also performs better than those produced earlier. .

Hsueh, Tseng, and Hsieh (2007) develop a model using 1990 and 2000 data for Taiwan that simultaneously looks at housing price, vacancy rate, and moving rate. The results for 1990 show that in a booming market situation, both expected housing price and current housing price had a strong, positive impact on the vacancy rate; however, the housing vacancy rate did not display a negative impact on housing price as expected. The results for 2000 show that housing price did not significantly affect the vacancy rate; however, the vacancy rate had a negative impact on housing price that was highly statistically significant. This result reflected the fact that housing market operation had swung to another extreme after the real estate bubble that started in the late 1980s and burst in the mid-1990s.

Khor, Ming, and Yuan (2000) look at the concept of a “natural” vacancy rate, which they define as an equilibrium level of inventory of space, in the sense that both the matching process between landlord and tenant is facilitated, and that building owners hold an optimal buffer stock of inventory to meet future leasing contingencies. They find that when vacancy rates are above the natural vacancy rate, rents will fall and vacancies will drift upward toward equilibrium. The determination of the natural vacancy rate is therefore significant in that it can facilitate the monitoring of the market conditions since a vacancy rate below the natural vacancy rate signifies a tight market. The authors find that the converse is true if the vacancy rate is above the equilibrium level.

Extrapolation models are discussed in the following section on PPH. However, they are not discussed in regard to VR because they are not useful. As is discussed in regard to PPH, extrapolation models match up well with the demographic determinants underlying changes in PPH and because of the inertia in these determinants, the models have a good track record in regard to post-censal PPH estimates. Extrapolations do not match up well with the determinants underlying changes in VR, however, so I neither describe nor evaluate them for purposes of generating VR values.

<u>Persons Per Household</u>

The development of PPH data has been examined by demographers more than VR data development has.   Similar to the examination of VR modeling by economists, more demographic theory comes to bear on PPH values than on VR values. I describe some of the theory here and return to this topic in Chapter IV.

Akkerman (1980) finds that a household composition matrix by age of head and age of other household members operates as a linear transformation from the vector of household distribution by age of head to the vector of population age distribution. He continues this observation by showing that the first row of the matrix may be interpreted as representing a vector of average household fertility rates. Akkerman (1980) observes that if a linear relationship between household and population distributions is fully implemented, then a relationship between household fertility and the size of the youngest age group can be derived. He concludes by noting that an extension of this result enables the simultaneous projection of population and households. Following up on these observations, Akkerman (2000) shows that the formal relationship between age of head to age of other household members is equivalent to the input-output relationship in the Leontief model of the open economy. Thus, he argues that the notions of household composition and household accommodation, which have emerged independently over the past two decades, are formally linked within this relationship.

Akkerman (2004) applies his earlier work to a case study of the Czech Republic during its post-communist transition to market economy. He uses the household composition matrix as a demographic gauge to the behavioral response of households to Czech housing markets and policy. He observes that the age-specific household size shown for the various regions of the Czech Republic follow a Gamma function, with anomalies detected in the trajectory of age-specific household size for Prague confirming unique housing market conditions and a commensurate demographic response. Acs and Nelson (2004) also examine the effect of policies on PPH values, but do so in the US and without the household composition matrix. They suggest that welfare policies may have contributed to the decline in single parenting and the rise in cohabitation between 1997 and 1999, a consequence of which is an increase in PPH. Brown (1999) also found household compositional matrices to be useful in examining PPH variation across small areas.

Burch (1967) examined census data from a range of countries and found that PPH mainly reflected past fertility rates rather than changes in family structure. Specifically, he finds that smaller PPH values had little to with the then-common view that urbanization was linked “breakdowns" of the extended residential family that had been more common in less urbanized societies. Burch (1970) followed with a broader examination that looked at the influence of demographic variables (viz., mortality, fertility, and age at marriage) on average household size under different family systems--nuclear, extended and stem. Using the model he developed, Burch found that under all three family systems, average household size is positively correlated with fertility, life expectancy, and average age at marriage. Other studies identifying the demographic determinants of PPH include Bumpass (1990), Burch et al. (1987) and Coale (1965). Ellen and O’Flaherty (2007) extend the work on the demographic determinants of PPH by examining the effect of government policies. Using data from a survey of households in New York City, they find that these policies appear to have an impact.

Using census data from 1790 to 1970, Korbin (1976) examined the long-term fall in household size in the United States and found that household changes were due to demographic changes. Sweet (1984) decomposed the growth of households during the 1970s and found that about one-third of the increase in the number of households was due to increased age by marital status propensity to form households while the remaining two-thirds was due to shifts in the age by marital status distribution and population growth. Specifically, he found that the increased propensity to form households had its major impact at ages under 35, and primarily among never-married persons while the composition component had its primary impact at ages 25-44 as a result of the baby boom, and also because of the increased fractions never married and separated and divorced during this period.

Myers and Doyle (1990) develop a demographic framework for determining PPH and propose models for estimating these values with the age structure of the household population. Mason and Racelis (1992) consider four models that explicitly incorporate the impact of changes in the number of men and women on the number and joint age distribution of husband-wife households. The models are applied to the Philippines using data from the 1988 National Demographic Survey to project households to 2010. The models are also evaluated by 'backcasting' and comparing the results with special tabulations from the 1970 and 1980 censuses and the 1975 National Demographic Survey. Van Imhoff et al. (1995) provide a wide range of demographic models that can be used to forecast and estimate households. Zeng, Land, Wang, and Gu (2006) propose a household model that uses demographic rates as input and projects more detailed household types, sizes, and living arrangements for all members of the population.

Although, it contains less demographic theory than the studies just reviewed, work by Smith, Nogle, and Cody (2002) found that regression models using symptomatic indicators of PPH change performed very well in estimating county-level PPH values in four states (Florida, Illinois, Texas, and Washington). Kimpel and Lowe (2007) and Washington Office of Financial management (2007) found less encouraging results in Washington for PPH estimates developed from regression models using symptomatic indicators. Swanson and Lowe (1979) examined the use of IRS data on average exemptions per return as a symptomatic indicator of PPH and found that while it had deficiencies, it held promise.

Even less demographic theory (i.e., virtually none) is found in the trend extrapolation models used in Washington and Alaska, among other places to develop post-censal PPH estimates (Alaska Department of Labor, 1981; Lowe, Pittenger, and Walker, 1977; Swanson, Baker and Van Patten, 1982; Washington Office of Financial Management, 1978). Work by Findley and Reinhardt (1980) found that PPH change during the 1970s was non-linear. This work serves to confirm that simple geometric trend extrapolation has the potential to develop accurate PPH estimates.

Sample surveys offer a means of obtaining current PPH data (Alaska Department of Labor, 1981; Hogan, 2007; Washington Office of Financial Management, 1978; Swanson and Hough, 2007). The obvious survey of choice for the Census Bureau is the ACS (Hogan, 2007).

III. Illustrations of Methods used to Generate VR and PPH Data

Vacancy Rates

There is not much in the way of illustrations in this section. The models described earlier, the ones developed largely by economists, are also not suitable for illustration. Holding vacancy rates constant since the last census does not require an illustration. For USPS delivery data, one needs to obtain the data (see, e.g. http://www.usps.com/ncsc/addressinfo/deliverystatistics.htm), learn about what they represent, and then apply the data to given areas with the appropriate adjustments. Adjustments are needed because of what the delivery data represent relative to vacancy rates (as explained earlier). An illustration of such an adjustment is given by Lowe and Mohrman (2003), who apply the rate of change found in the USPS delivery data from the date of the previous census to the current (estimate) date to the vacancy rate found in the previous census. This effectively brings forward the census VR to the current estimate date. This can be done for all housing units combined, with subsequent adjustments by structure type, as needed.

The Windshield surveys mentioned in the preceding chapter can be done on a full count basis or a sample basis (Alaska Department of Labor, 1981; Washington Office of Financial Management, 1978). This type of field work generally works best in towns of less than 100,000 that are not characterized by massive numbers of multi-unit housing units. It can be done in urban areas with concentrations of multi-unit housing, but typically this requires a fair amount of work on foot, which adds to costs. The detailed procedures are not presented here; rather, an overview is given.

A windshield represents field work, which is expensive. However, a windshield survey is far less expensive than door-to-door canvassing. The first step is to map out the area for which VR is desired in terms of blocks (Alaska Department of Labor, 1982). This map then serves as the basis for a full count and as the frame for a sample. If a full count is to be done, these blocks can then be organized into enumeration areas. If a sample is to be done, the blocks serve as PSUs, which can be randomly selected (or more preferably, randomly selected with variations, such as the more cost-effective random selection of clusters of blocks and then full counting of the blocks within these clusters, where the clusters are equivalent to "enumeration areas." Once the areas are mapped, two-person teams systematically canvass the blocks in automobiles using the same in a set manner (e.g., driving such that only turns are made to the right, starting in the northeast corner of a block and ending up where the canvassing began) with the driver focused on driving and the second person focused on counting, classifying, and recording.

Basic training is required, but generally this can be done in a single day. The training includes census definitions of housing units and vacancies, as well as practical field procedures Evidence from Washington, suggests that in areas with less than 20,000 people, the full count was preferable to a sample survey. When areas had between 20,000 and 100,000 then a sample was preferable because of the cost savings. Many more of the nuances and details (e.g., identifying large clusters of certain types of housing structure during the mapping phase such as mobile home parks and where possible, obtaining vacancy information from resident managers, once they understood the nature of a census-defined vacancy) are found in the HUM Manuals produced by the Alaska Department of Labor (1981) and the Washington Office of Financial Management (1978), with supplemental information found in the State of Alaska Census

Administrators Manual (Alaska Department of Labor, 1982) and the State of Alaska Census Enumerator's Manual (Alaska Department of Labor, 1981).

Persons Per Household

Current PPH values have often been obtained by using the value from the most recent census or extrapolating trends found from the two most recent decennial censuses, with the geometric model being typically used (Alaska Department of Labor, 1981; Washington Office of Financial management, 1978; Bryan, 2004b; Lowe, Pittenger, and Walker, 1977; Swanson, Baker, and Van Patten, 1983). In this approach, the rate of change is benchmarked to two most recent successive census counts and then applied to the PPH value found in the most recent census count, which is then extrapolated beyond the most recent census by applying the rate of change to it.

The process takes place in two steps. The first is the calculation of the rate of change in PPH:

$$r = (PPH_l/PPH_b)^{(1/y)} - 1 \qquad [2]$$

where

r = rate of change

PPH = Persons Per Household

l = Launch Year (most recent census)

b = Base Year (Census preceding launch Year

y = Number of years between l and b (10 years)

The second step is applying the rate to the launch year to find PPH values:

$$PPH_t = (PPH_l)\ [(\ 1 + r)^{(y)}]$$

where

r = rate of change (from step 1)

PPH = Persons Per Household

t = Target Year

l = Launch Year (most recent census)

y = number of years between t and l

Table 1 shows the results of this process for counties in Washington. The base year is 1980 and the launch year is 1990. As an example of the two steps described above, consider Kitsap County for which the first step yields a rate of -0.0013, which is equal to $[(2.6469/2.6820)^{(1/10)} – 1]$, and for which the second step yields a 2000 PPH estimate of 2.6123, which is equal to $\{(2.6469) [(1 – 0.0013)^{(10)}]\}$

Table 1 also provides a comparison of the PPH values estimated for 2000 with the reported PPH values from the 2000 Census. As the summary statistics at the bottom of Table 1 show, the estimated PPH values are accurate, appoint to which I return in the following chapter.

Other forms of extrapolation can be used, with the caution that a linear extrapolation model is not as desirable as non-linear ones such as the geometric model (Findley and Reinhart, 1980). For example, a double-ratio geometric model could be used, although it is a bit more complicated than the simple geometric model. It requires three time points of data. The first step is to calculate separate rates of change for the two time periods (e.g., 1980 to 1990 and 1990 to 2000) as is done in Equation [2] for the simple geometric model. If the rate of change for second (more current) period is less than the rate of change for the first period then one calculates the ratio of the absolute rate of change from the first period

Table 1. Accuracy Test of the Geometric Method of Estimating PPH Values for Counties: Washington State 2000

Washington State PPH Values By County, 1980, 1990, and 2000								
	1980 Persons Per Household	1990 Persons Per Household	2000 Persons per Household	1980-1990 Geometric Rate of Change	Estimated 2000 Persons Per Household	Absolute Error	Percent Error	MAPE
STATE	2.6086	2.5348	2.5349	-0.0029	2.4631	-0.0718	-2.83%	2.83%
Adams	2.9113	2.9405	3.0949	0.0010	2.9700	-0.1249	-4.03%	4.03%
Asotin	2.5662	2.4727	2.4162	-0.0037	2.3826	-0.0336	-1.39%	1.39%
Benton	2.7971	2.6516	2.6795	-0.0053	2.5137	-0.1658	-6.19%	6.19%
Chelan	2.4827	2.4863	2.6192	0.0001	2.4899	-0.1293	-4.93%	4.93%
Clallam	2.5374	2.4007	2.3066	-0.0055	2.2714	-0.0353	-1.53%	1.53%
Clark	2.7625	2.6625	2.6900	-0.0037	2.5661	-0.1239	-4.61%	4.61%
Columbia	2.5254	2.4368	2.3628	-0.0036	2.3513	-0.0115	-0.49%	0.49%
Cowlitz	2.6619	2.5588	2.5531	-0.0039	2.4597	-0.0934	-3.66%	3.66%
Douglas	2.7591	2.6769	2.7554	-0.0030	2.5971	-0.1583	-5.74%	5.74%
Ferry	2.8567	2.6978	2.4938	-0.0057	2.5477	0.0539	2.16%	2.16%
Franklin	2.8817	3.034	3.2637	0.0052	3.1943	-0.0693	-2.12%	2.12%
Garfield	2.5955	2.3948	2.3911	-0.0080	2.2096	-0.1815	-7.59%	7.59%
Grant	2.7986	2.7407	2.9204	-0.0021	2.6840	-0.2364	-8.09%	8.09%
Grays Harbor	2.5966	2.4813	2.4826	-0.0045	2.3711	-0.1115	-4.49%	4.49%
Island	2.6706	2.6149	2.5223	-0.0021	2.5604	0.0381	1.51%	1.51%
Jefferson	2.4537	2.3089	2.2122	-0.0061	2.1726	-0.0395	-1.79%	1.79%
King	2.4868	2.3982	2.3905	-0.0036	2.3128	-0.0777	-3.25%	3.25%
Kitsap	2.682	2.6469	2.6007	-0.0013	2.6123	0.0115	0.44%	0.44%
Kittitas	2.3976	2.3251	2.3314	-0.0031	2.2548	-0.0766	-3.29%	3.29%
Klickitat	2.7211	2.6409	2.5361	-0.0030	2.5631	0.0270	1.06%	1.06%
Lewis	2.6732	2.5997	2.5690	-0.0028	2.5282	-0.0408	-1.59%	1.59%
Lincoln	2.5726	2.4308	2.4233	-0.0057	2.2968	-0.1265	-5.22%	5.22%
Mason	2.5458	2.5162	2.4891	-0.0012	2.4869	-0.0022	-0.09%	0.09%
Okanogan	2.6674	2.5877	2.5762	-0.0030	2.5104	-0.0658	-2.56%	2.56%
Pacific	2.4465	2.3499	2.2711	-0.0040	2.2571	-0.0140	-0.62%	0.62%
Pend Oreille	2.8088	2.6029	2.5074	-0.0076	2.4121	-0.0953	-3.80%	3.80%
Pierce	2.6586	2.6231	2.6047	-0.0013	2.5881	-0.0166	-0.64%	0.64%
San Juan	2.2946	2.2489	2.1587	-0.0020	2.2041	0.0454	2.10%	2.10%
Skagit	2.5656	2.5495	2.6032	-0.0006	2.5335	-0.0697	-2.68%	2.68%
Skamania	2.7896	2.6921	2.6120	-0.0036	2.5980	-0.0140	-0.54%	0.54%
Snohomish	2.7606	2.67935	2.6547	-0.0030	2.6005	-0.0542	-2.04%	2.04%
Spokane	2.5789	2.4747	2.4646	-0.0041	2.3747	-0.0899	-3.65%	3.65%
Stevens	2.907	2.7318	2.6439	-0.0062	2.5672	-0.0768	-2.90%	2.90%
Thurston	2.6441	2.553	2.4987	-0.0035	2.4650	-0.0337	-1.35%	1.35%
Wahkiakum	2.7724	2.4762	2.4243	-0.0112	2.2116	-0.2127	-8.77%	8.77%
Walla Walla	2.5411	2.4955	2.5388	-0.0018	2.4507	-0.0880	-3.47%	3.47%
Whatcom	2.5902	2.5324	2.5113	-0.0023	2.4759	-0.0354	-1.41%	1.41%
Whitman	2.4668	2.3868	2.3115	-0.0033	2.3094	-0.0021	-0.09%	0.09%
Yakima	2.7711	2.8039	2.9576	0.0012	2.8371	-0.1205	-4.08%	4.08%

COUNTY LEVEL SUMMARY STATISTICS	
Mean Error	-0.0680
MAPE	2.97%
MALPE	-2.60%
N ABS % ERROR >5	6

to the absolute rate of change for the second time period and, which is then used to modify the launch year PPH; If the rate of change for second period is greater than the rate of change for the first period then one uses PPH as given in the launch year (most recent census).

Other forms of extrapolation can be used, with the caution that a linear extrapolation model is not as desirable as non-linear ones such as the geometric model (Findley and Reinhart, 1980). For example, a double-ratio geometric model could be used, although it is a bit more complicated than the simple geometric model. It requires three time points of data. The first step is to calculate separate rates of change for the two time periods (e.g., 1980 to 1990 and 1990 to 2000) as is done in Equation [2] for the simple geometric model. If the rate of change for second (more current) period is less than the rate of change for the first period then one calculates the ratio of the absolute rate of change from the first period to the absolute rate of change for the second time period and, which is then used to modify the modify the launch year PPH; If rate of change for second period is greater than the rate of change for the first period then one uses PPH as given in the launch year (most recent census)from the most recent census

IRS data on the number of exemptions per return have been examined with an eye toward using them to model current PPH values (Swanson and Lowe, 1979; Voss and Krebs, 1979). One problem noted by Swanson and Lowe (1979) is that the use of the IRS data effectively collapses all structure types into one category. This is a problem because of the different PPH levels associated with different structure types (Swanson and Lowe, 1979), which lead to their suggestion to use IRS data in conjunction with PPH data from the Current Population Survey. Their report preceded the ACS by 20 years, so an appropriate revision of their suggestion would be to use the IRS data in conjunction with PPH values taken from the ACS.

Stan Smith, June Nogle, and Scott Cody (2002) considered four models when they examined the accuracy of a regression approach for estimating PPH using county data from four states, Florida, Illinois, Texas, and Washington (N = 462).

Model 1 (Basic model) is defined as:

$$PPH_{it} = a + b_1*Births_{it} + b_2*School_{it} + b_3*Medicare_{it}$$

where
a = the intercept term and b_1, b_2, and b_3 = regression coefficients
i = county (i = 1 to 462)
t = time (a given year, 1970, 1980, and 1980)
PPH = Persons Per Household
Births = Births Per Household
School = K-12 School Enrollment Per Household
Medicare = Medicare Enrollment (age 65+) Per Household

Smith and his colleagues found that the basic model had the following characteristics for 1970, 1980, and 1990, respectively:

$$\hat{PPH}_{i70} = 2.477 + 10.199*Births_{i70} + 0.359*School_{i70} -0.487*Medicare_{i70}$$

Adj. $R^2 = 0.756$ and all coefficients are statistically significant ($p \leq .01$)

$$\hat{PPH}_{i80} = 2.072 + 11.192*Births_{i80} + 0.512*School_{i80} -0.164*Medicare_{i80}$$

Adj. $R^2 = 0.744$ and all coefficients are statistically significant ($p \leq .01$)

$$\hat{PPH}_{i90} = 2.077 + 12.213*Births_{i90} + 0.473*School_{i90} -0.288*Medicare_{i90}$$

Adj. $R^2 = 0.763$ and all coefficients are statistically significant ($p \leq .01$)

Model 2 (Ratio model) is defined as:

$$PPH_{ikt} = a + b_1*(Births_{ik}/Births_k)_t + b_2*(School_{ik}/School_k)_t + b_3*(Medicare_{ik}/Medicare_k)_t$$

where

a = the intercept term and b_1, b_2, and b_3 = regression coefficients

ik = county i in state k (i = 1 to 67, k=1; i=1 to 102, k=2; i=1 to 254, k=3; i=1 to 39, k=4)

k = state (1 = Florida, 2 = Illinois, 3 = Texas, and 4 = Washington)

t = time (a given year, 1970, 1980, and 1980)

PPH = Persons Per Household

Births = Births Per Household

School = K-12 School Enrollment Per Household

Medicare = Medicare Enrollment (age 65+) Per Household

Smith and his colleagues found that the ratio model had the following characteristics for 1970, 1980, and 1990, respectively:

$$\hat{PPH}_{ik70} = 0.727 + 0.124*(Births_{ik}/Births_k)_{70} + 0.180*(School_{ik}/School_k)_{70} - 0.035* (Medicare_{ik}/Medicare_k)_{70}$$

Adj. R^2 = 0.843 and all coefficients are statistically significant ($p \leq .01$)

$$\hat{PPH}_{ik80} = 0.751 + 0.086*(Births_{ik}/Births_k)_{80} + 0.180*(School_{ik}/School_k)_{80} - 0.021* (Medicare_{ik}/Medicare_k)_{80}$$

Adj. R^2 = 0.821 and all coefficients are statistically significant ($p \leq .01$)

$$\hat{PPH}_{ik90} = 0.738 + 0.125*(Births_{ik}/Births_k)_{90} + 0.154*(School_{ik}/School_k)_{90} -0.021* (Medicare_{ik}/Medicare_k)_{90}$$

Adj. R^2 = 0.836 and all coefficients are statistically significant ($p \leq .01$)

Model 3 (Change model) is defined as:

$$(PPH_{it} - PPH_{ic}) = a + b_1*(Births_{it} - Births_{ic}) + b_2*(School_{it} - School_{ic}) + b_3*(Medicare_{it} - Medicare_{ic}) + b_4*(PPH_{ic})$$

where

a = the intercept term and b_1, b_2, and b_3 = regression coefficients
i = county (i = 1 to 462)
t = time (a given year, 1970, 1980, and 1990)
c = preceding census (1970, 1980)
PPH = Persons Per Household
Births = Births Per Household
School = K-12 School Enrollment Per Household
Medicare = Medicare Enrollment (age 65+) Per Household

Smith and his colleagues found that this model had the following characteristics for 1970-80, and 1980-90, respectively:

$$\widehat{(PPH_{i80} - PPH_{i70})} = 0.241 + 1.851*(Births_{i80} - Births_{i70}) + 0.225*(School_{i80} - School_{i70}) - 0.621*(Medicare_{i80} - Medicare_{i70}) - 0.162*(PPH_{70})$$

Adj. $R^2 = 0.559$ and all coefficients are statistically significant ($p \leq .01$)

$$\widehat{(PPH_{i90} - PPH_{i80})} = 0.014 + 1.544*(Births_{i90} - Births_{i80}) + 0.934*(School_{i90} - School_{i80}) - 0.219*(Medicare_{i90} - Medicare_{i80}) - 0.033*(PPH_{80})$$

Adj. $R^2 = 0.490$ and all coefficients are statistically significant ($p \leq .01$) except the intercept

Model 4 (Ratio Change model) is defined as:

$$[(PPH_{ik} /PPH_k)_t - (PPH_{ik} /PPH_k)_c] = a + b_1*[(Births_{ik}/Births_k)_t - (Births_{ik}/Births_k)_c]$$

$$+ b_2*[(School_{ik}/School_k)_t - (Births_{ik}/Births_k)_c]$$

$$+ b_3*[(Medicare_{ik}/Medicare_k)_t - (Medicare_{ik}/Medicare_k)_c]$$

$$+ b_4*(PPH_i)_c$$

where

a = the intercept term and b_1, b_2, and b_3 = regression coefficients

ik = county i in state k (i = 1 to 67, k=1; i=1 to 102, k=2; i=1 to 254, k=3; i=1 to 39, k=4)

k = state (1 = Florida, 2 = Illinois, 3 = Texas, and 4 = Washington)

t = time (a given year, 1980, and 1990)

c = preceding census (1970, 1980)

PPH = Persons Per Household

Births = Births Per Household

School = K-12 School Enrollment Per Household

Medicare = Medicare Enrollment (age 65+) Per Household

Smith and his colleagues found that this model had the following characteristics for 1970-80 and 1980-90, respectively:

$$\widehat{[(PPH_{ik} /PPH_k)_{80} - (PPH_{ik} /PPH_k)_{70}]} = 0.138 + 0.029*[(Births_{ik}/Births_k)_{80} - (Births_{ik}/Births_k)_{70}]$$

$$+ 0.070*[(School_{ik}/School_k)_{80} - (Births_{ik}/Births_k)_{70}]$$

$$- 0.047*[(Medicare_{ik}/Medicare_k)_{80} - (Medicare_{ik}/Medicare_k)_{70}]$$

$$-0.041*(PPH_i)_{70}$$

Adj. $R^2 = 0.490$ and all coefficients are statistically significant ($p \leq .01$)

$$[\widehat{(PPH_{ik}/PPH_k)_{90}} - (PPH_{ik}/PPH_k)_{80}] = 0.011 + 0.040*[(Births_{ik}/Births_k)_{90} - (Births_{ik}/Births_k)_{80}]$$

$$+ 0.132*[(School_{ik}/School_k)_{90} - (Births_{ik}/Births_k)_{80}]$$

$$- 0.022*[(Medicare_{ik}/Medicare_k)_{90} - (Medicare_{ik}/Medicare_k)_{80}]$$

$$- 0.006*(PPH_i)_{80}$$

Adj. $R^2 = 0.470$ and all coefficients are statistically significant ($p \leq .01$) except the intercept and b4 (-0.006, for lagged persons per household, $(PPH_i)_{80}$)

While the preceding models are a bit dated and cover only four states (Florida, Illinois, Texas, and Washington), they offer guidance on the construction of regression models for estimating PPH using more current data for all counties in the U.S. It appears that the state-specific models offer more promise than those that cross state boundaries

IV. Evaluation of Methods used to Generate VR and PPH Data

Evaluation Criteria

Without question, an estimate should be accurate, but accuracy is not the only criterion by which an estimate should be judged. Following the argument presented by Swanson and Tayman (1995), I suggest that attention be focused on the broader concept of utility. As described in Section I, there are many methods that in principle can be used to estimate a population, and improvements are a regular feature of these methods. Further, there is a wide range of decision-making situations in which population estimates are used. It follows, therefore, that no method should be universally judged to be superior to others and, by the same token, neither should any method be judged universally inferior to all others. I suggest instead, that relative to a given use, utility is gained by selecting a method that provides a sufficient amount of information for the purpose(s) at hand, while keeping cost and time to a minimum. In the case of an estimate, the sufficiency of the information provided is judged on the ability of using it to make good decisions. So, if an estimate is produced at minimal cost but provides timely information sufficient to make good decisions, then it has high utility. If an estimate does not meet these conditions then it has low utility. An important underlying component of sufficiency is "transparency." That is, the ability of a decision-maker to understand how an estimate was done so that he or she can determine if the assumptions, methods, and data are reasonable. Another important component of utility that affects agencies responsible for (hierarchically-structured) sets of estimates across geographic areas is consistency, especially when different methods are used at different levels of geography. This is sometimes referred to as a "one-number roll-up" and it is obvious that this is an important component for the Census Bureau.

Like most other methods for generating population estimates (See, e.g., Bryan, 2004b), it has been possible to evaluate methods for estimating total population estimates by using decennial census counts (recognizing that unlike the total population counts and the ascribed characteristics such as age, race, and sex, virtually all of the ascribed population characteristics in the decennial census counts were derived from the sample-based "long form"). Fortunately, this will not change in terms of VR and PPH in 2010, when the Census Bureau abandons the long form in favor of the ACS. Thus, unlike many population and housing characteristics, the decennial census rather than the ACS will remain the future "gold standard" against which methods for estimating VR and PPH can be evaluated. This is fortuitous, because traditional frameworks and criteria for evaluating the accuracy of VR and PPH estimates will not require modification However, it is worthwhile to note in this regard that even though decennial census data provide the most convenient and accurate standard against which to evaluate population estimation methods, there always have been several important considerations that were taken into account before estimates are compared with census values. First, there was often a tendency to assume that earlier and later censuses are completely consistent, but such consistency cannot be taken for granted. Second, subnational areas often differ in geography and populations covered, and census definitions may have changed as well. Third, where a method was based on a past census that differs from a more recent census against which estimates resulting from the method in question were compared, in any significant way, an accurate evaluation was compromised.

All of this leads up to the point that comparisons of estimates resulting from different methods against the census have to be considered "measures of difference" rather than "measures of error" because it is virtually impossible to precisely determine the degree to which error in the census and error in the estimate contribute to the overall difference. This affects the evaluation of methods used to estimate VR and PPH.

Similarly, it is important to note that a direct "method-to-method" comparison is rarely possible when attempting to make a population estimate. Often, what might be the most accurate method may not be practicable due to excessive time, cost and resources. Other hindrances may include unavailability or inconsistency of necessary data. Furthermore, it will be seen, certain methods are better suited to particularly large or small areas of geography. While a certain method may generate "good results" at a national level, they may be wholly inadequate for other levels of geography. Thus the amount of resources available, the level of geography as well as historical accuracy of each method must always be considered. General criteria that apply at all levels for evaluating methods for estimating the total population (and its ascribed characteristics , include continuity; timeliness of information; refinement; production; cost, and replication. Generally speaking, these criteria can be applied to data, methods, and administrative structures, topics to which I now turn in regard to VR and PPH.

Evaluation of VR and PPH Data and the Methods for Obtaining Them

This section starts with modeling, moves to administrative data and ends with an evaluation of the uses of surveys as sources of VR and PPH. The ACS is treated separately from other sources and I evaluate the ACS in terms of both VR and PPH data toward the end of this section, just before the summary. One caution before starting to look at the individual elements making up an HUM estimate is the interaction effect observed by Lowe, Weisser, and Myers (1988). The authors observe that of some of the terms in the equation for the Housing Unit Method are themselves correlated and point out that an accuracy improvement in the data used in a given term may end up having an adverse impact on other terms such that the resulting estimate is less accurate. This is an important point because it underscores the need to examine the HUM and its data elements as a system rather than in terms of its individual terms and their data elements. Lowe, Weisser, and Myers (1988) provide suggestions for dealing with this issue.

Vacancy Rates

Holding rates constant over the last census is feasible, but this approach to determining current VR levels should be done in a system such that doing so appears warranted, rather than simply using this approach to the exclusion of others. If other data suggest little change since the last census, then holding PPH values constant may be appropriate. One way to judge where constant rates would work is to examine the IRS migration data and USPS data. If little change is indicated then constant rates may be appropriate. The process of determining if past census VR levels should be changed can be largely automated, with thresholds and decisions points empirically established by examining IRS and USPS data changes relative to past census counts.

Similar to the suggestion for setting up an automated system to determine if it appears feasible to hold VR constant since the last census and assuming that USPS delivery data would be part of it, then the same system could be used to apply USPS trends to previous census VR values as suggested by the work of Lowe (1988), Lowe and Mohrman (2003), and Lowe, Mohrman, and Brunink (2003). Using the USPS data for this purpose would appear to fit the Census Bureau's needs in regard to national coverage, consistency of definitions, and having single source of data. There would be much work to do in regard to the findings of Lowe and her colleagues concerning metropolitan and metropolitan counties, the presence of special populates (e.g., military bases) and the like. However, the Census Bureau is well equipped to undertake the needed analyses, given it has basically all of the data on hand already. Refinements to this general approach – a research agenda – are found in work by Lowe, Weisser, and Myers (1984) as well as Bousfield (1977), Clogg, Schockey, and Elisason (1990), and Weidman et al. (2008). An important potential refinement can be found in Fonseca and Tayman (1989) who develop and evaluate a method for deriving postcensal estimates of household income distributions for counties. The modified lognormal probability curve they use as a model of income distribution cold be adopted for use with VR, with the USPS delivery data and census VR data serving the role that IRS income data and decennial income data do for Fonseca and Tayman (1989), respectively.

The windshield surveys mentioned in the preceding chapter could become a tool in the "automated review" system I have just proposed. If indications are such that neither holding rates constant nor adjusting previous census data using trends shown from USPS data looks viable for a given area, then a windshield survey might be considered, given the size and type of area (as described in the previous chapter. If any are used, it would be useful to have areas 'clustered" such that findings from one windshield survey (e.g., changes since the last census) could be applied to other areas within the same cluster. The variables used to identify such clusters could include size, metropolitan status, presence of a downtown core, percent of multiple unit structure, and so forth. Given the size training, and experience of the Census Bureau's field operations and staff in regard to collecting data, conducting windshield surveys as indicated by the automated review system may be a tractable solution.

Persons Per Household

As was suggested in regard to VR values, holding PPH values constant over the last census is feasible, but it should be done in a system such that doing so appears warranted, rather than simply using this approach to the exclusion of others. Thus, it appears preferable to use administrative data (e.g., IRS migration, USPS deliveries, vital statistics, K-12 school enrollment, Medicare, and covered employment) to determine if PPH change is likely. If little change is indicated then holding PPH values constant may be appropriate. Like the process of determining if VR values could be held constant, the process of determining if past census PPH levels should be held constant can be largely automated, with thresholds and decisions points empirically established by examining administrative data changes relative to past census counts.

Trend extrapolation is inexpensive and easy to implement. It also has been found to provide accurate projections when used judiciously (Smith , Tayman, and Swanson, 2001: 167-183). That is, by not extending the past too far into the future and by understanding the demographic dynamics being modeled by an extrapolation technique such as the geometric model in conjunction with PPH data from successive and recent census counts. It is worthwhile here to point out that while trend extrapolation may not be well suited to capture certain aspects of demographic change (e. g, cohort effects on the total population), the general approach used to generate the estimated 2000 PPH values found in Table 1 is well suited to capture PPH changes. It is so, because two successive decennial census points provide a good time frame for the effects of demographic determinants on PPH to be observed and a geometric model used to extrapolate this change effectively exploits the inertia underlying it. Recall that Burch (1967) and others (Bumpass, 1990; Burch et al., 1987; and Coale, 1965 have identified these determinants and found that their effects on PPH are not played out over a short period of time. These aspects generally mean that extrapolation methods such as found in the geometric model can work well, as is shown in Table 1.

IRS data on the number of exemptions per return have been examined with an eye toward using them to model current PPH values (Swanson and Lowe, 1979; Voss and Krebs, 1979). One problem noted by Swanson and Lowe (1979) is that the use of the IRS data effectively collapses all structure types into one category. This is a problem because of the different PPH levels associated with different structure types (Swanson and Lowe, 1979) , which leads to their suggestion to use IRS data in conjunction with PPH data from the Current Population Survey. Their report preceded the ACS by 20 years, so an appropriate revision of their suggestion would be to use the IRS data in conjunction with PPH values taken from the ACS, which could be massaged using, for example, techniques described by Bousfield (1977) , Clogg, Schockey, and Eliason (1990), and Lowe (2000a).

The regression models for estimating PPH that were constructed by Stan Smith, June Nogle, and Scott Cody (2002) show that this approach has considerable promise for the Census Bureau, with model 2 (Ratio model) standing out. Recall that this model is state-specific and defined as follows:

$$PPH_{ikt} = a + b_1*(Births_{ik}/Births_k)_t + b_2*(School_{ik}/School_k)_t + b_3*(Medicare_{ik}/Medicare_k)_t$$

where
a = the intercept term and b_1, b_2, and b_3 = regression coefficients
ik = county i in state k (i = 1 to 67, k=1; i=1 to 102, k=2; i=1 to 254, k=3; i=1 to 39, k=4)

k = state (1 = Florida, 2 = Illinois, 3 = Texas, and 4 = Washington)
t = time (a given year, 1970, 1980, and 1980)
PPH = Persons Per Household
Births = Births Per Household
School = K-12 School Enrollment Per Household
Medicare = Medicare Enrollment (age 65+) Per Household

The results of the models constructed by Smith and his colleagues (2002) for 1970, 1980, and 1990 are very good in terms of their statistical properties, as is shown below.

<u>1970 Model</u>

$$\hat{PPH}_{ik70} = 0.727 + 0.124*(Births_{ik}/Births_k)_{70} + 0.180*(School_{ik}/School_k)_{70} - 0.035*(Medicare_{ik}/Medicare_k)_{70}$$

Adj. $R^2 = 0.843$ and all coefficients are statistically significant ($p \leq .01$)

<u>1980 Model</u>

$$\hat{PPH}_{ik80} = 0.751 + 0.086*(Births_{ik}/Births_k)_{80} + 0.180*(School_{ik}/School_k)_{80} - 0.021*(Medicare_{ik}/Medicare_k)_{80}$$

Adj. $R^2 = 0.821$ and all coefficients are statistically significant ($p \leq .01$)

<u>1990 Model</u>

$$\hat{PPH}_{ik90} = 0.738 + 0.125*(Births_{ik}/Births_k)_{90} + 0.154*(School_{ik}/School_k)_{90} -0.021*(Medicare_{ik}/Medicare_k)_{90}$$

Adj. $R^2 = 0.836$ and all coefficients are statistically significant ($p \leq .01$)

However, even though the characteristics do not seem as favorable as this is the case with Model 2 (Ratio Model), I also would not discount Model 4 (Ratio Change Model). I make this suggestion based on findings about ratios combined with change in regression models used to estimate population by Swanson (2004) and the fact that Model 4 could be modified using ideas from Swanson (1980), Swanson and Beck (1994), Swanson, Tayman, and Beck (1995), and Swanson and Tedrow (1984). While I cannot say for certain if Model 4 is sufficient, I believe it would be worth examining it further, with the idea that it may be sufficient as is, or that revisions may make it so.

In regard to the work of Smith and his colleagues (2002) in developing and examining their models, it should be noted that the "traditional" methods against they were evaluated did not include geometric and other non-linear models. Instead, the models were assessed relative to the following three 'traditional methods: (1) holding PPH constant for each county from the last census; taking the percent change in PPH for each during the previous decade, applying it to the most recent census PPH value and bringing it forward accordingly; and (3) using the percent change in PPH at the state level since the most recent census and applying this uniformly to the most recent census PPH value in each county and bringing them each forward accordingly. Thus, I suggest that if the Census Bureau examines regression methods for estimating PPH relative to other methods that the geometric model and its non-linear relatives be included.

One important finding that emerges from the work of Smith, Nogle and Cody (2002) is that it is consistent with other work that stresses that 'no-size fits all' when it comes to the HUM and other methods of population methods. The use of specific local area (in this case, county) data is important as is using state-specific models. Thus, I suggest that the Census Bureau take these two related issues in to account as it considers how to implement the HUM nationwide. In addition, in regard to the HUM itself, it is wise to consider the findings of Lowe, Weisser, and Myers (1988) in regard to the interactions of terms in the HUM. It is important to realize that knowledge that improvement in the accuracy of a given data element such as housing units may adversely impact the accuracy of the HUM overall, which suggests that improvements to data used in the HUM need to be assessed in a comprehensive manner rather than simply assuming that accuracy improvements in one element will automatically lead to improvement in the accuracy of the overall population estimates.

An important factor in regard to regression models, is that they can be used with sample data (e. g, Ericksen , 1973, 1974), which brings in the ACS as a data source, a point to which I now turn.

The ACS

Background

The American Community Survey (ACS) is a U. S. Census Bureau product designed to provide accurate and timely demographic and economic indicators on an annual basis for both large and small geographic areas within the United States (Citro and Kalton, 2007; U. S. Census Bureau, 2004b). Operational plans call for ACS to serve not only as a substitute for the decennial census long-form, but as a means of providing annual data at the national, state, county, and sub-county levels (Cork, Cohen, and King, 2004; Smith, 1998; U. S. Census Bureau, 2001a, 2001b, 2003, 2004b) . In addition to being highly ambitious, this approach represents a major change in how data are collected and interpreted (Citro and Kalton, 2007; Hough and Swanson, 1998, 2006). Two of the major questions facing the ACS are its functionality and usability (Citro and Kalton, 2007).

As has been noted earlier, the Decennial Census, (along with the CPS, and SIPP, among other products) uses the De jure rule of residency while the ACS uses what amounts to a De facto residency rule. However, when the micro-level ACS data are aggregated to geographic areas, they are controlled to number produced for these areas by the Census Bureau's annual population estimates program. This may not be a huge issue at the national level, but at sub-state levels, the effects of these different residency rules could be substantial. Where ACS micro-level data are used in conjunction with micro-level data from the CPS, SIPP, or the (2000) census, it is highly likely that mismatches occur. Add to this difference, there is evidence that ACS values at the sub-state level are subject to high levels of variation in general (Fay, 2007; Van Auken, Hammer, Voss, and Veroff, 2006). Moreover, there is now evidence that PPH values taken from the ACS are affected by both of these issues (Swanson and Hough, 2007) .

Putting these two concerns together leads to a high level of uncertainty in regard to the applicability of ACS-derived PPH values top given areas. As an illustration of the potential magnitude of this uncertainty, consider Table 2 (in four parts, a, b, c, and d), which compares "official population estimates" produced by the Census Bureau for 115 cities in California with "ACS population estimates" for these cities (San Francisco is shown in the table, but excluded from the analysis because the city boundaries are the same as the county boundaries). Table 2 also shows the margins of error around the ACS population estimates and shows which cities have official estimates that are not within the respective margins or error. Even with relatively wide margins of error the official estimates are not contained within them in 20 of the 115 cities. That is, 17.4% of the cities have official estimates that are beyond the margins of error of the corresponding ACS estimates. It likely to be discouraging for all users that an agency would be producing two sets of estimates for cities nationwide. Particularly disturbing is the fact that given there are two sets of city numbers, over 17 percent of the 2007 official estimates for California cities are beyond the ACS margins of errors. Hopefully, on-going research within the Census Bureau (e.g., Robinson and Dixon, 2009; Weidman et al., 2008) along with administrative cognizance of the adverse effects of having two sets of estimates will lead to resolutions of these two related issues.

Table 2.a. Comparison of 2007 Official Population Estimates for California Cities with ACS estimates, Alemeda through Fullerton.

City	ACS TOTAL POP 2007	REGULAR ESTIMATE TOTAL POP 2007	Numeric Difference	ACS MARGIN OF ERROR, TOTAL POP	DIFFERENCE WITHIN ACS MARGIN OF ERROR?
Alameda city, California	75,642	70,272	5,370	+/-6,407	YES
Alhambra city, California	88,393	86,352	2,041	+/-7,561	YES
Anaheim city, California	342,856	333,249	9,607	+/-16,199	YES
Antioch city, California	104,426	99,619	4,807	+/-7,963	YES
Apple Valley town, California	69,835	70,322	-487	+/-7,377	YES
Bakersfield city, California	324,540	315,837	8,703	+/-11,127	YES
Baldwin Park city, California	76,945	77,800	-855	+/-8,390	YES
Bellflower city, California	69,477	73,434	-3,957	+/-7,658	YES
Berkeley city, California	111,680	101,377	10,303	+/-5,974	NO
Buena Park city, California	85,992	79,281	6,711	+/-8,490	YES
Burbank city, California	96,972	103,286	-6,314	+/-7,251	YES
Carlsbad city, California	95,796	95,439	357	+/-7,106	YES
Chico city, California	83,460	83,128	332	+/-4,963	YES
Chino city, California	83,914	82,830	1,084	+/-8,228	YES
Chula Vista city, California	227,336	217,478	9,858	+/-11,597	YES
Citrus Heights city, California	88,576	84,469	4,107	+/-7,880	YES
Clovis city, California	92,987	90,808	2,179	+/-7,419	YES
Compton city, California	100,037	94,425	5,612	+/-9,928	YES
Concord city, California	124,300	120,844	3,456	+/-8,089	YES
Corona city, California	156,394	150,308	6,086	+/-11,341	YES
Costa Mesa city, California	114,057	108,978	5,079	+/-8,421	YES
Daly City city, California	104,752	100,882	3,870	+/-7,752	YES
Downey city, California	109,920	108,109	1,811	+/-11,536	YES
El Cajon city, California	97,964	92,533	5,431	+/-7,993	YES
Elk Grove city, California	138,072	131,212	6,860	+/-9,718	YES
El Monte city, California	113,308	122,272	-8,964	+/-9,809	YES
Escondido city, California	128,819	136,246	-7,427	+/-8,744	YES
Fairfield city, California	111,007	103,992	7,015	+/-7,979	YES
Folsom city, California	74,795	67,401	7,394	+/-5,299	NO
Fontana city, California	193,716	183,502	10,214	+/-11,369	YES
Fremont city, California	214,957	201,334	13,623	+/-10,482	NO
Fresno city, California	476,460	470,508	5,952	+/-11,446	YES
Fullerton city, California	126,955	132,066	-5,111	+/-8,303	YES

Table 2.b. Comparison of 2007 Official Population Estimates for California Cities with ACS estimates, Garden Grove through Norwalk.

City	ACS TOTAL POP 2007	REGULAR ESTIMATE TOTAL POP 2007	Numeric Difference	ACS MARGIN OF ERROR, TOTAL POP	DIFFERENCE WITHIN ACS MARGIN OF ERROR?
Garden Grove city, California	145,923	165,610	-19,687	+/-13,426	NO
Glendale city, California	200,859	196,979	3,880	+/-10,101	YES
Hawthorne city, California	92,321	84,422	7,899	+/-9,239	YES
Hayward city, California	129,885	140,943	-11,058	+/-8,203	NO
Hemet city, California	77,001	70,288	6,713	+/-7,235	YES
Hesperia city, California	90,312	85,515	4,797	+/-8,245	YES
Huntington Beach city, California	188,056	192,885	-4,829	+/-10,437	YES
Indio city, California	70,791	83,937	-13,146	+/-7,026	NO
Inglewood city, California	106,581	113,376	-6,795	+/-9,709	YES
Irvine city, California	205,813	201,160	4,653	+/-8,408	YES
Lake Forest city, California	78,130	75,688	2,442	+/-8,483	YES
Lakewood city, California	89,289	78,956	10,333	+/-7,840	NO
Lancaster city, California	155,902	143,616	12,286	+/-11,940	NO
Livermore city, California	79,213	79,532	-319	+/-6,366	YES
Long Beach city, California	458,302	466,520	-8,218	+/-18,630	YES
Los Angeles city, California	3,806,003	3,834,340	-28,337	+/-43,027	YES
Lynwood city, California	69,537	70,336	-799	+/-6,745	YES
Merced city, California	73,224	76,879	-3,655	+/-6,075	YES
Milpitas city, California	66,494	66,770	-276	+/-5,593	YES
Mission Viejo city, California	92,673	94,586	-1,913	+/-5,823	YES
Modesto city, California	198,456	203,955	-5,499	+/-10,352	YES
Moreno Valley city, California	190,990	188,936	2,054	+/-11,306	YES
Mountain View city, California	70,000	70,436	-436	+/-5,467	YES
Murrieta city, California	89,885	90,555	-670	+/-7,722	YES
Napa city, California	71,664	74,247	-2,583	+/-4,183	YES
Newport Beach city, California	89,125	79,554	9,571	+/-5,726	NO
Norwalk city, California	112,001	103,720	8,281	+/-10,988	YES

Table 2.c. Comparison of 2007 Official Population Estimates for California Cities with ACS estimates, Oakland through San Leando.

City	ACS TOTAL POP 2007	REGULAR ESTIMATE TOTAL POP 2007	Numeric Difference	ACS MARGIN OF ERROR, TOTAL POP	DIFFERENCE WITHIN ACS MARGIN OF ERROR?
Oakland city, California	358,829	401,489	-42,660	+/-13,801	NO
Oceanside city, California	168,814	168,602	212	+/-9,661	YES
Ontario city, California	156,027	170,936	-14,909	+/-11,593	NO
Orange city, California	142,097	134,299	7,798	+/-11,764	YES
Oxnard city, California	167,412	184,725	-17,313	+/-8,354	NO
Palmdale city, California	132,266	140,882	-8,616	+/-10,047	YES
Pasadena city, California	136,936	143,400	-6,464	+/-9,751	YES
Pleasanton city, California	69,348	66,544	2,804	+/-5,983	YES
Pomona city, California	142,111	152,631	-10,520	+/-11,043	YES
Rancho Cucamonga city, California	157,777	170,266	-12,489	+/-12,011	NO
Redding City	87,130	89,780	-2,650	+/-5,302	YES
Redlands city, California	73,539	69,941	3,598	+/-8,059	YES
Redondo Beach city, California	70,948	67,019	3,929	+/-6,838	YES
Redwood City city, California	69,559	73,603	-4,044	+/-5,891	YES
Rialto city, California	108,969	98,713	10,256	+/-9,628	NO
Richmond city, California	97,279	101,454	-4,175	+/-9,020	YES
Riverside city, California	316,154	294,437	21,717	+/-14,637	NO
Roseville city, California	114,958	108,759	6,199	+/-6,578	YES
Sacramento city, California	451,404	460,242	-8,838	+/-15,995	YES
Salinas city, California	140,499	143,517	-3,018	+/-8,046	YES
San Bernardino city, California	203,691	199,285	4,406	+/-11,585	YES
San Buenaventura (Ventura) city, California	105,673	103,219	2,454	+/-6,800	YES
San Diego city, California	1,276,740	1,266,731	10,009	+/-22,810	YES
San Francisco city, California	764,976	764,976	0	*****	N/A
San Jose city, California	922,389	939,899	-17,510	+/-16,294	NO
San Leandro city, California	96,186	77,725	18,461	+/-9,192	NO

Table 2.d. Comparison of 2007 official Population Estimates for California Cities with ACS estimates, San Marcos through Yorba Linda.

City	ACS TOTAL POP 2007	REGULAR ESTIMATE TOTAL POP 2007	Numeric Difference	ACS MARGIN OF ERROR, TOTAL POP	DIFFERENCE WITHIN ACS MARGIN OF ERROR?
San Marcos city, California	75,217	78,286	-3,069	+/-7,344	YES
San Mateo city, California	91,461	91,768	-307	+/-5,967	YES
Santa Ana city, California	327,780	339,555	-11,775	+/-14,085	YES
Santa Barbara city, California	89,959	86,204	3,755	+/-6,453	YES
Santa Clara city, California	105,591	109,756	-4,165	+/-6,451	YES
Santa Clarita city, California	177,740	169,951	7,789	+/-13,076	YES
Santa Maria city, California	86,160	85,685	475	+/-6,581	YES
Santa Monica city, California	86,857	87,212	-355	+/-6,317	YES
Santa Rosa city, California	147,516	154,241	-6,725	+/-8,276	YES
Simi Valley city, California	127,053	120,464	6,589	+/-8,047	YES
South Gate city, California	104,031	97,110	6,921	+/-8,719	YES
Stockton city, California	295,070	287,245	7,825	+/-10,999	YES
Sunnyvale city, California	135,548	131,140	4,408	+/-8,627	YES
Temecula city, California	93,743	94,767	-1,024	+/-8,318	YES
Thousand Oaks city, California	128,519	123,349	5,170	+/-7,821	YES
Torrance city, California	143,628	141,420	2,208	+/-8,316	YES
Tracy city, California	82,383	79,705	2,678	+/-6,540	YES
Turlock city, California	69,330	68,133	1,197	+/-6,246	YES
Tustin city, California	63,524	70,869	-7,345	+/-6,624	YES
Union City city, California	73,212	70,075	3,137	+/-6,373	YES
Upland city, California	78,260	72,464	5,796	+/-9,002	YES
Vacaville city, California	93,795	92,084	1,711	+/-6,076	YES
Vallejo city, California	106,608	115,552	-8,944	+/-6,297	NO
Victorville city, California	97,534	107,221	-9,687	+/-9,214	NO
Visalia city, California	115,899	118,603	-2,704	+/-8,732	YES
Vista city, California	97,977	90,839	7,138	+/-10,065	YES
West Covina city, California	103,154	106,388	-3,234	+/-8,704	YES
Westminster city, California	91,994	88,678	3,316	+/-6,606	YES
Whittier city, California	82,755	82,850	-95	+/-8,684	YES
Yorba Linda city, California	57,550	65,434	-7,884	+/-4,415	NO

This section serves as an introduction to an evaluation of the use of ACS data in regard to VR and PPH. In the following section, I focus specifically on VR and then on PPH.

ACS Vacancy Rates

In regard to ACS vacancy rates, it is highly likely that they have too much variance to use. However, this serves as a research question rather than a conclusion because VR itself is subject to higher levels of variation than PPH values over a given interval of time. In addition to the issue of statistical variance, the interaction of seasonality and residency differences between the ACS and the decennial census also must be considered. It is more difficult to assess the effects of statistical uncertainty and the interaction of seasonality and residency differences on VR than it is on PPH because changes in the determinants of VR lack the inertia associated with changes in the demographic determinants of PPH. With these point in mind, suggestions for analysis of the suitability of ACS VR values for use in an HUM estimation system include: (1) conducting a broad scale comparison, taking note of county-level 'market conditions' that are likely to have impacts on VR levels and their changes; and (2) making adjustments to ACS VR values (deriving model-based PPH values from the ACS) that may provide more statistical stability.

ACS Persons Per Household[7]

The data used in this exploration of the usability and functionality of ACS PPH data are taken from 18 counties that were in the 1999 ACS test sites (See Exhibit 1). The examination proceeds by comparing ACS PPH values for these 18 counties to PPH values generated using a geometric model based on PPH change from Census 1990 to Census 2000. The ACS PPH values represent what could be called the "statistical perspective" because variations in the values of specific variables over time and space are viewed largely by statisticians with an eye toward sample (and non-sample) error (Citro and Kalton, 2007; Fay, 2005; Kish, 1998; Purcell and Kish, 1979, 1980; U. S. Census Bureau, 2001a, 2001b, 2003, 2004b). The model-based PPH values represent a "demographic perspective" because PPH values are largely viewed by demographers as varying systematically, an orientation stemming from theory and empirical evidence that PPH values respond to demographic and related determinants (Burch, 1967, 1970; Burch et al., 1987; Coale, 1965; Goldsmith, Jackson, and Shambaugh, 1982; Kimpel and Lowe, 2007; Korbin, 1976; Myers and Doyle, 1990; Smith, Nogle, and Cody, 2002).

EXHIBIT 1. The 18 COUNTIES USED IN THE ANALYSIS

Pima County, AZ	Madison County, MS
Jefferson County, AR	Douglas County, NE
San Francisco County, CA	Bronx County, NY
Tulare County, CA	Rockland County, NY
Broward County, FL	Franklin County, OH
Lake County, IL	Multnomah County, OR
Black Hawk County, IA	Schuylkill County, PA
Calvert County, MD	Sevier County, TN
Hampden County, MA	Yakima County, WA

As noted earlier, the HUM is based on the assumption that virtually everyone lives in some type of housing structure. It is generally accepted that the HUM is the most commonly used method for making small area population estimates in the United States (Byerly, 1990; Smith, Nogle, and Cody, 2002). One of the reasons for this is that current data for two of its elements are generally available, the number of households and the group quarters population (Smith, Nogle, and Cody, 2002). The other remaining element needed to get the household population is PPH. Until the full implementation of the ACS, current PPH values were obtained by using the value from the most recent census or extrapolating trends found from the two most recent decennial censuses (Bryan, 2004b; Smith, Nogle, and Cody, 2004; Swanson, Baker, and Van Patten, 1983). With the expansion of the ACS to its full design in 2005 (Griffin and Waite, 2006), it is not surprising that among the large number of HUM users, more than a few are interested in seeing if the ACS can provide usable PPH values. Thus, this evaluation.

Evaluation Data

The U. S. Census Bureau established the operational structure for the ACS in 1994 when it put in place the "Continuous Measurement Office," which implemented the first operational test of the ACS in four test sites in 1995 (Griffin and Waite, 2006). These test sites were subsequently expanded, and by 1999, operational tests took place in 36 counties spread across 26 states (Griffin and Waite, 2006). Three year ACS averages centered on 2000 were set up for these counties to support comparisons with Census 2000. Relevant among the many findings of these tests was that the arithmetic mean (2.63) of the PPH values found in the ACS for these 36 counties was the same as that found in Census 2000 and that there were no statistically significant differences for PPH (U. S. Census Bureau, 2004b: 17). It was also noted that this result was not unexpected because the total household population and the total number of housing units found in Census 2000 are used as control variables in ACS weighting (U. S. Census Bureau 2004b: 17).

Among the 36 ACS test counties, annual PPH values estimated from single-year ACS collections are available online for 21 of them for the period 2001 to 2006; annual PPH values estimated from three-year ACS collections are available online for 18 of these same 21 counties for the period 1999-2001 to 2003-2005. (See Exhibit 1). It is for these 18 counties that both single-year and three-year ACS PPH values are used in our comparison with model-based PPH values.

The analytical method for generating the model-based PPH values is one method commonly used by applied demographers for this purpose, namely, the geometric rate of change (Lowe, Pittenger, and Walker, 1977; Smith, Nogle, and Cody, 2002; Smith, Tayman, and Swanson, 2001; Swanson, Baker, and Van Patten, 1983). In this approach, the rate of change is benchmarked to two most recent successive census counts and then applied to the PPH value found in the most recent census count, which is then extrapolated beyond the most recent census by applying the rate of change to it.

The process takes place in two steps. The first is the calculation of the rate of change in PPH, which is described in the preceding chapter as equation, along with the second step (applying the rate to the launch year to obtain current PPH values), so I will not repeat them here.

Results

The data for the 18 counties are shown in exhibits 2 through 19. Each of these exhibits is divided into two parts. The first part shows the single-year ACS PPH values for each year from 2001 to 2006 while the second part shows the three-year ACS PPH values for each year from 2001 to 2005, the latter corresponding to the ACS collections from 1999-2001 to 2003-2005.

Both parts of each of the exhibits also show the annual ACS values generated using the geometric model. The ACS PPH values are labeled as “ACSPPH” in each of the two parts and the model-generated PPH values are labeled as “ADPPH” (where “AD” stands for Analytically Derived).

EXHIBIT 2.1*

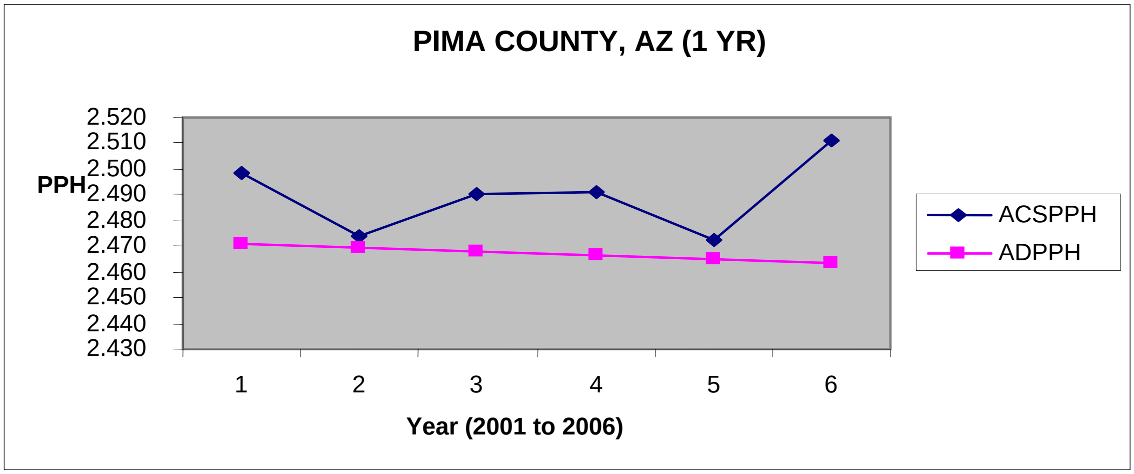
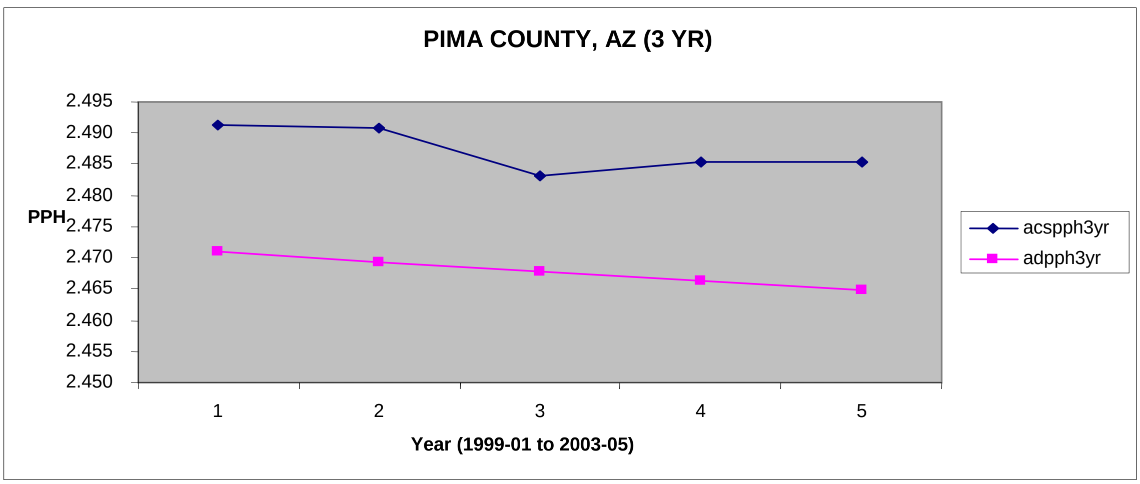


EXHIBIT 2.2*



*The ACS PPH values are labeled as “ACSPPH” and the model-generated PPH values are labeled as “ADPPH” (where “AD” stands for Analytically Derived). “(1 YR)” stands for single year ACS data and “(3 YR)” stands for 3 year ACS data.

EXHIBIT 3.1*

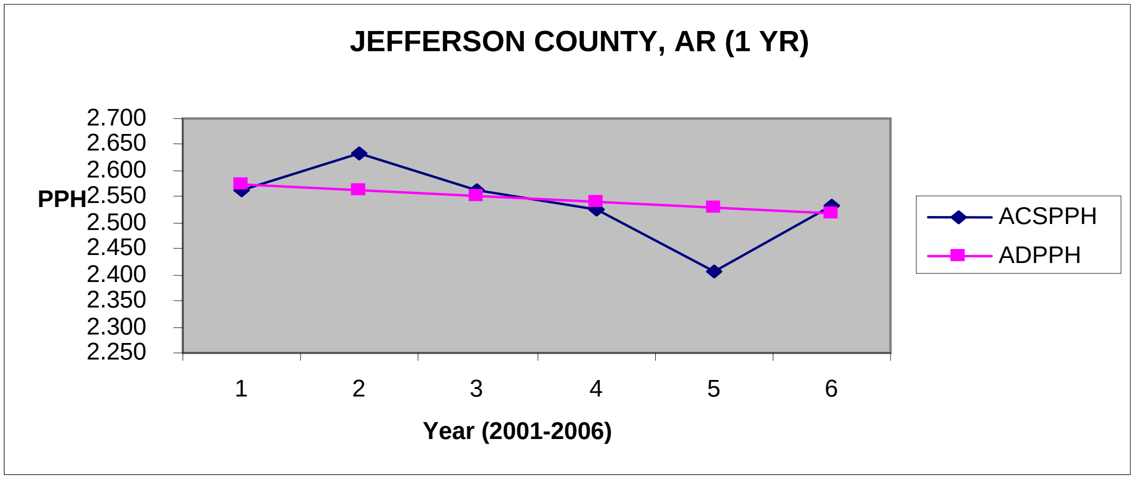
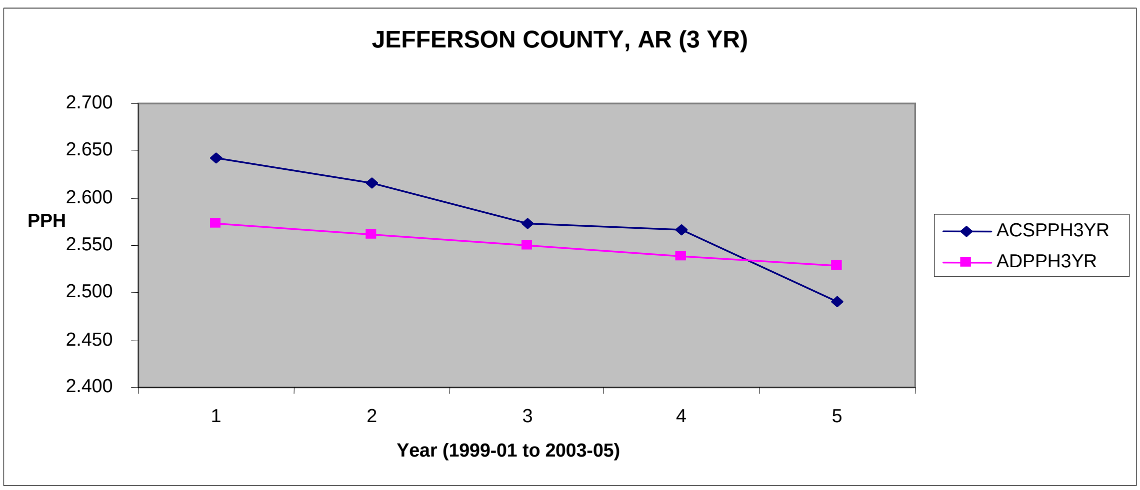


EXHIBIT 3.2*



*The ACS PPH values are labeled as "ACSPPH" and the model-generated PPH values are labeled as "ADPPH" (where "AD" stands for Analytically Derived). "(1 YR)" stands for single year ACS data and "(3 YR)" stands for 3 year ACS data.

EXHIBIT 4.1*

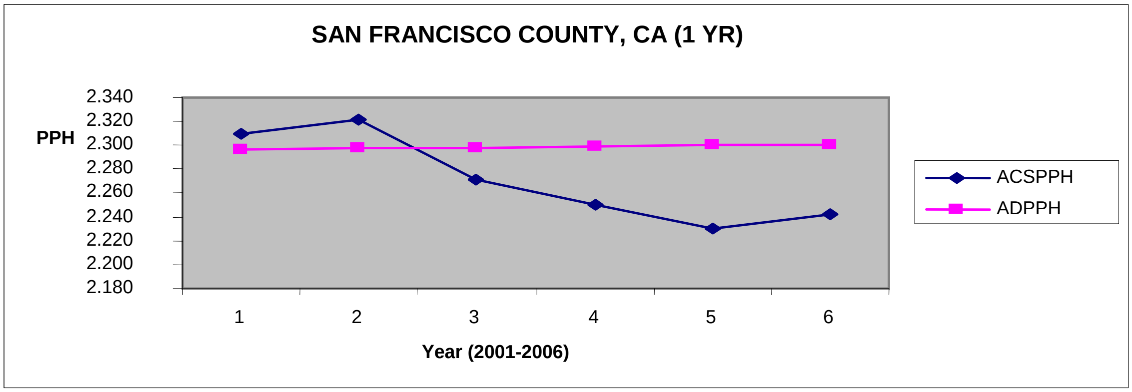
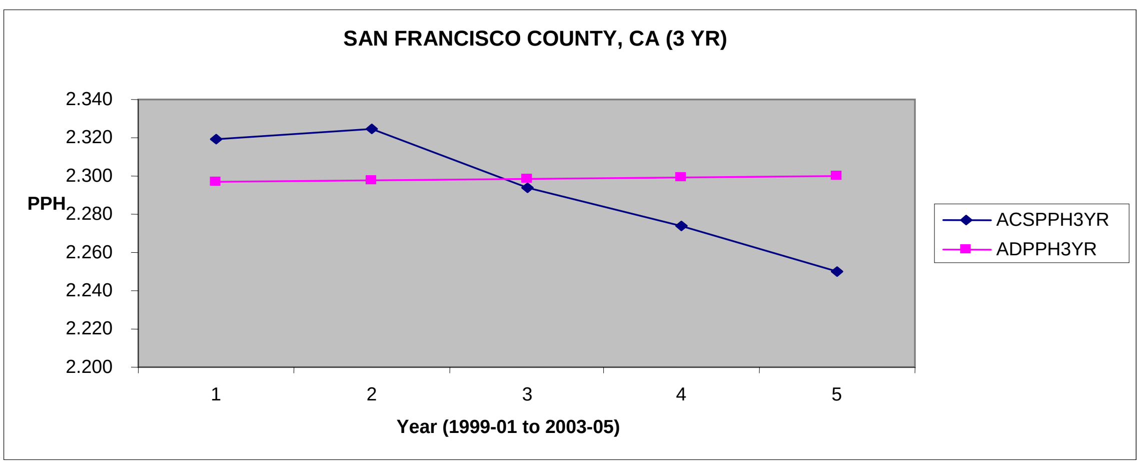


EXHIBIT 4.2*



*The ACS PPH values are labeled as "ACSPPH" and the model-generated PPH values are labeled as "ADPPH" (where "AD" stands for Analytically Derived). "(1 YR)" stands for single year ACS data and "(3 YR)" stands for 3 year ACS data.

EXHIBIT 5.1*

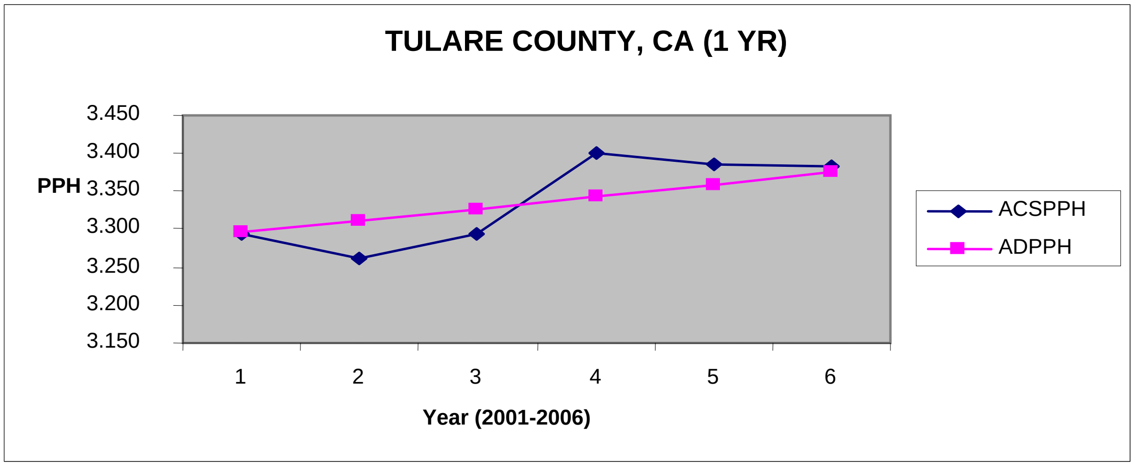
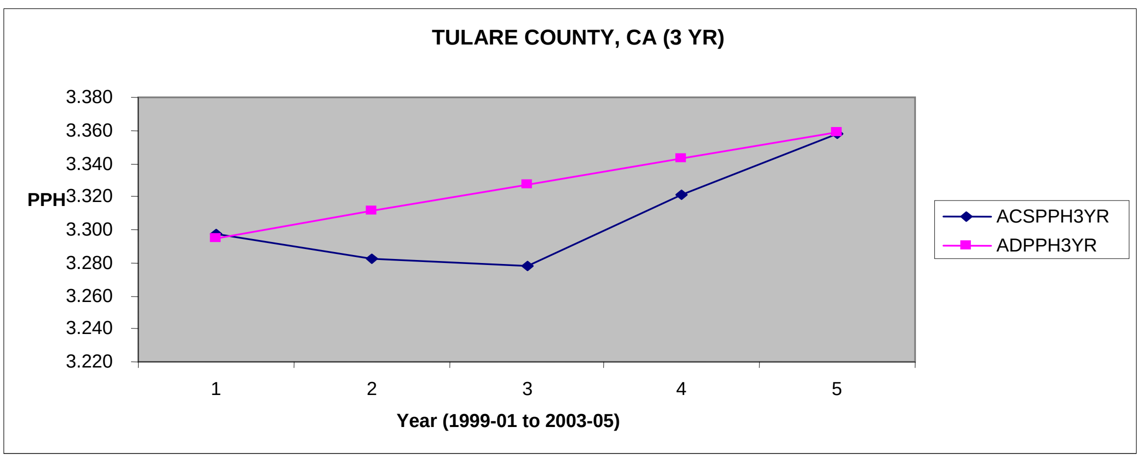


EXHIBIT 5.2*



*The ACS PPH values are labeled as "ACSPPH" and the model-generated PPH values are labeled as "ADPPH" (where "AD" stands for Analytically Derived). "(1 YR)" stands for single year ACS data and "(3 YR)" stands for 3 year ACS data.

EXHIBIT 6.1*

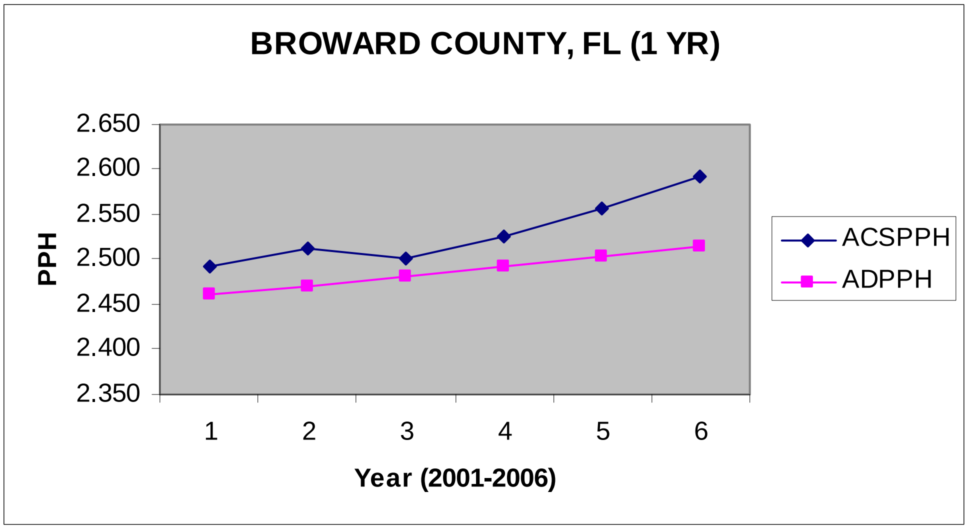
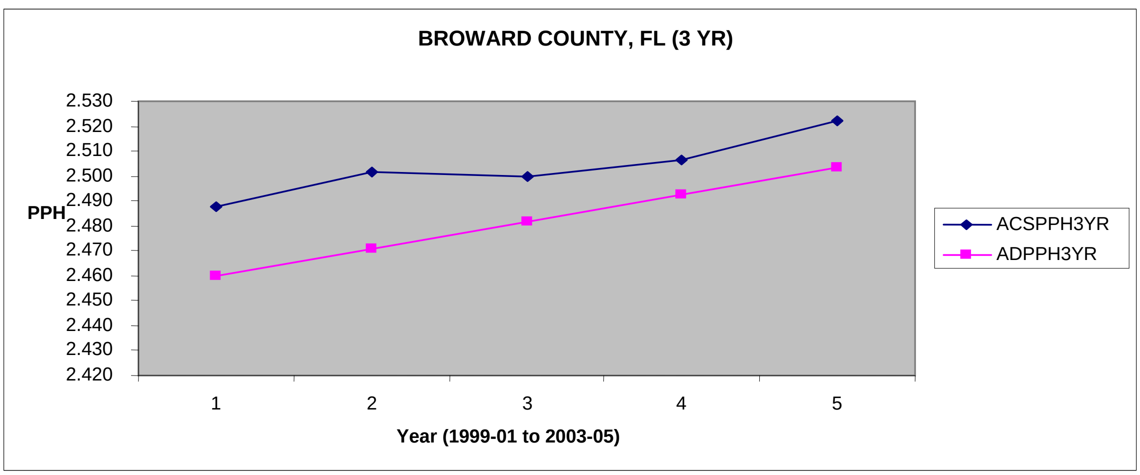


EXHIBIT 6.2*



*The ACS PPH values are labeled as "ACSPPH" and the model-generated PPH values are labeled as "ADPPH" (where "AD" stands for Analytically Derived). "(1 YR)" stands for single year ACS data and "(3 YR)" stands for 3 year ACS data.

EXHIBIT 7.1*

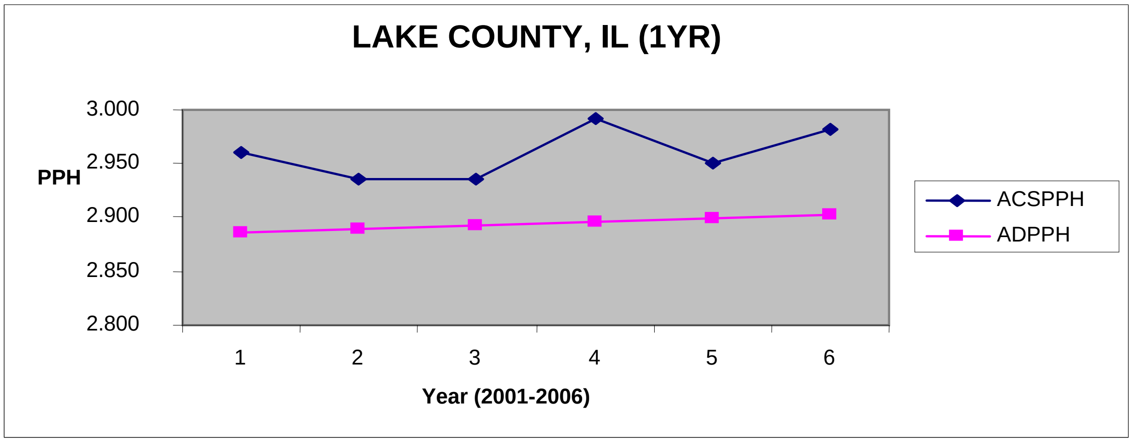
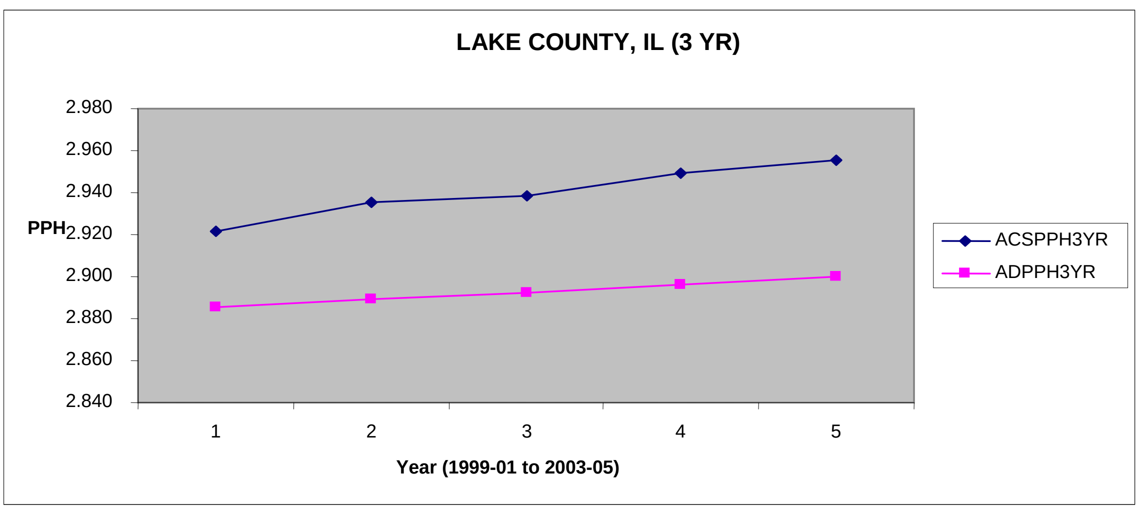


EXHIBIT 7.2*



*The ACS PPH values are labeled as "ACSPPH" and the model-generated PPH values are labeled as "ADPPH" (where "AD" stands for Analytically Derived). "(1 YR)" stands for single year ACS data and "(3 YR)" stands for 3 year ACS data.

EXHIBIT 8.1*

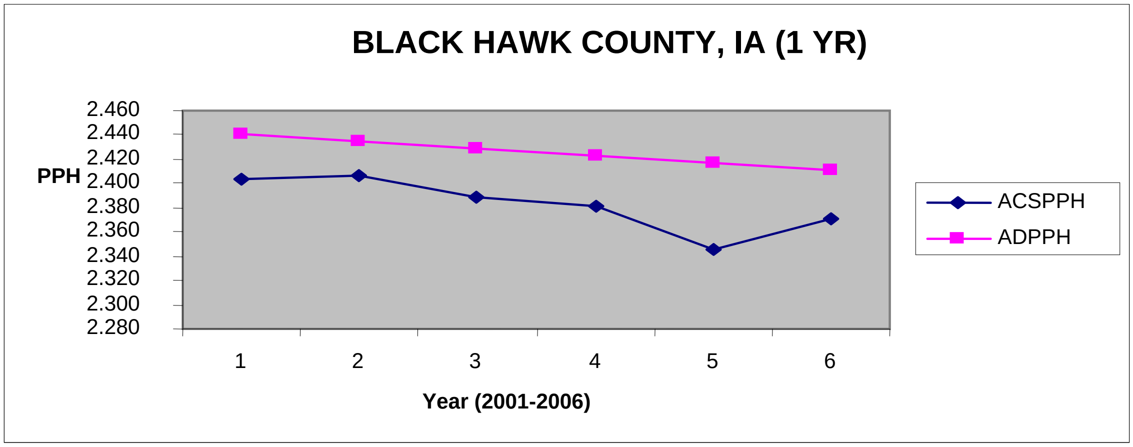
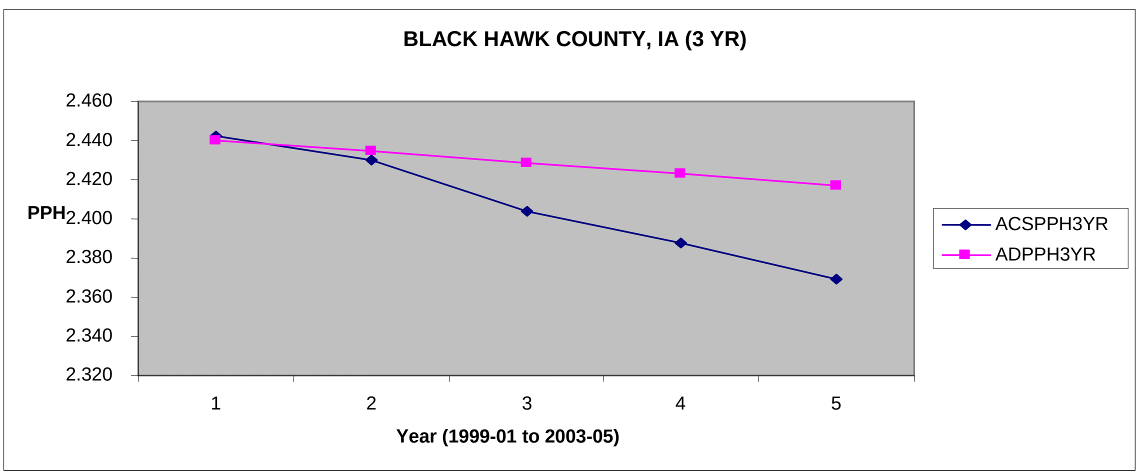


EXHIBIT 8.2*



*The ACS PPH values are labeled as "ACSPPH" and the model-generated PPH values are labeled as "ADPPH" (where "AD" stands for Analytically Derived). "(1 YR)" stands for single year ACS data and "(3 YR)" stands for 3 year ACS data.

EXHIBIT 9.1*

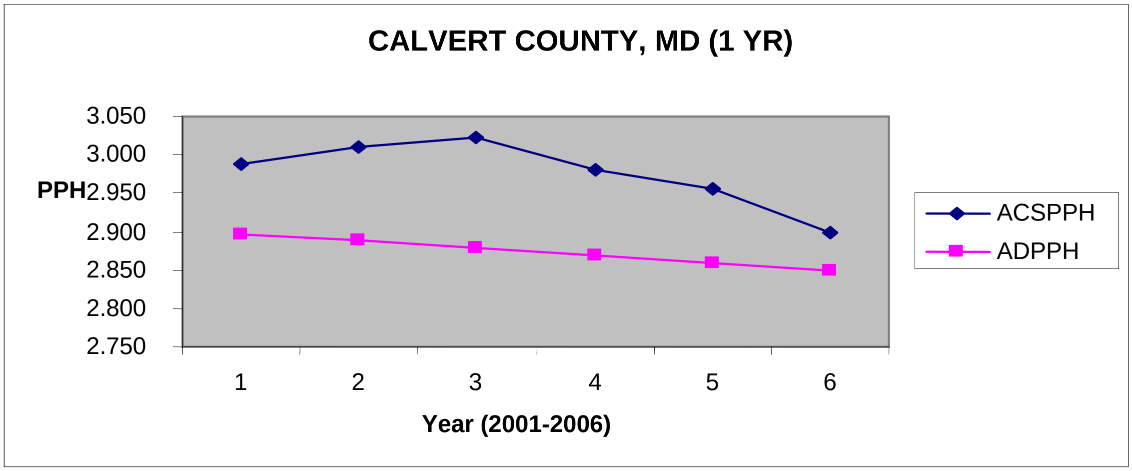
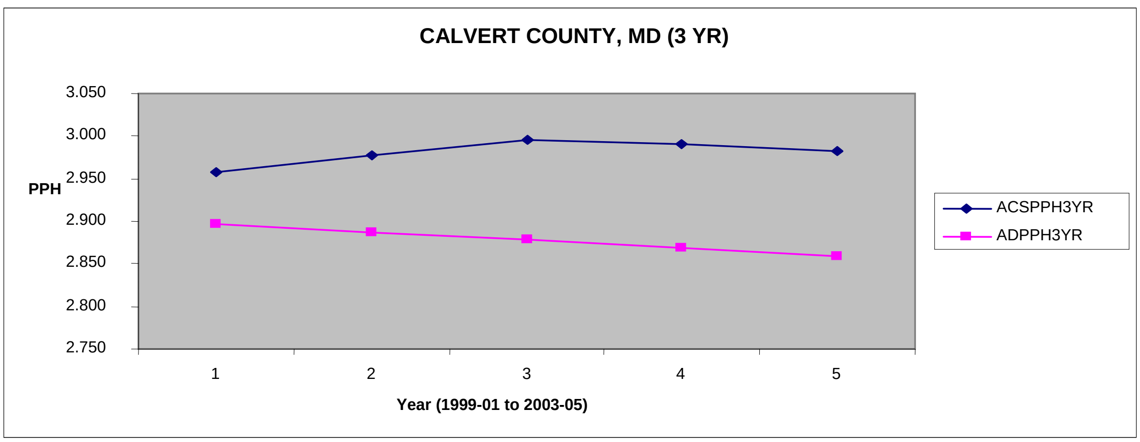


EXHIBIT 9.2*



*The ACS PPH values are labeled as “ACSPPH” and the model-generated PPH values are labeled as “ADPPH” (where “AD” stands for Analytically Derived). “(1 YR)” stands for single year ACS data and “(3 YR)” stands for 3 year ACS data.

EXHIBIT 10.1*

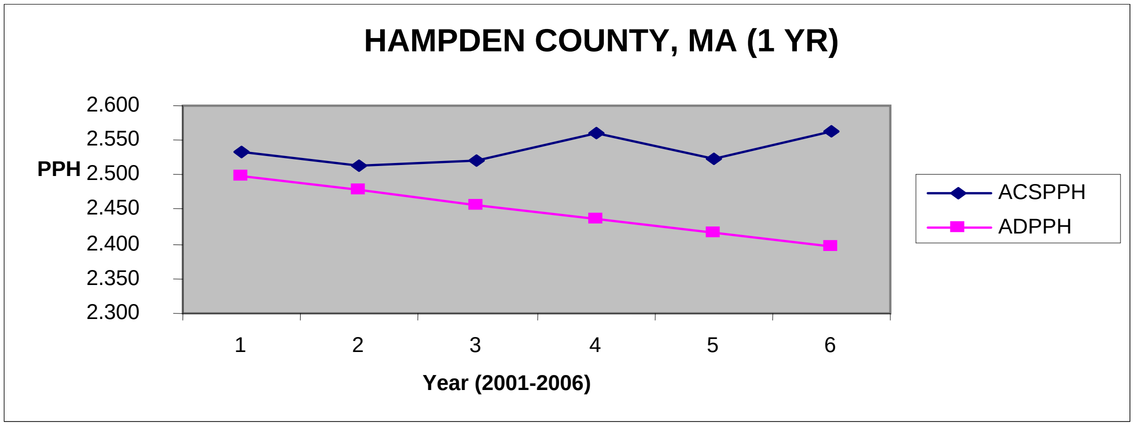
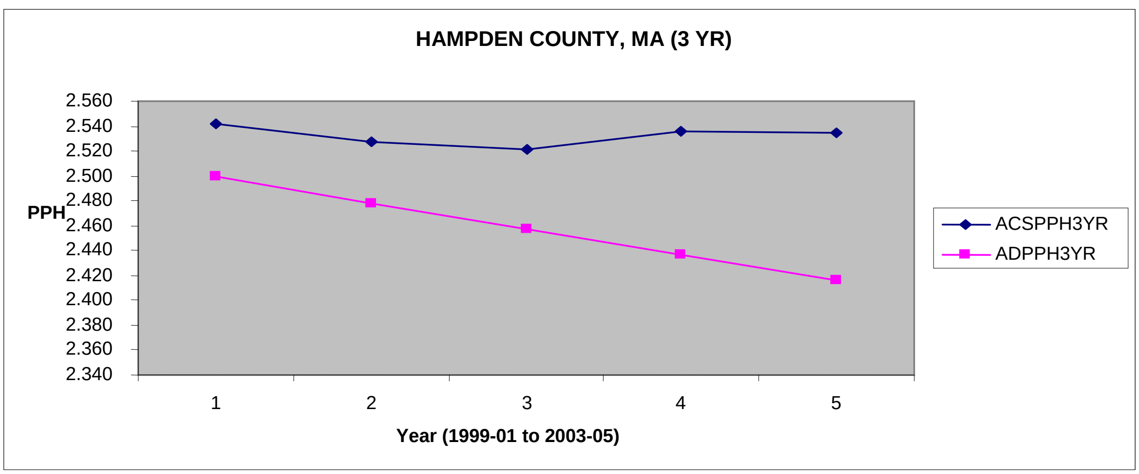


EXHIBIT 10.2*



*The ACS PPH values are labeled as "ACSPPH" and the model-generated PPH values are labeled as "ADPPH" (where "AD" stands for Analytically Derived). "(1 YR)" stands for single year ACS data and "(3 YR)" stands for 3 year ACS data.

EXHIBIT 11.1*

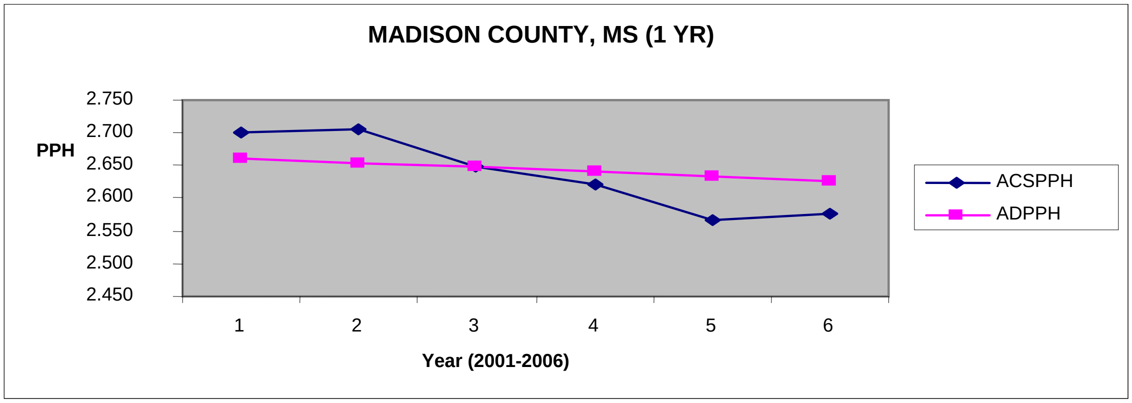
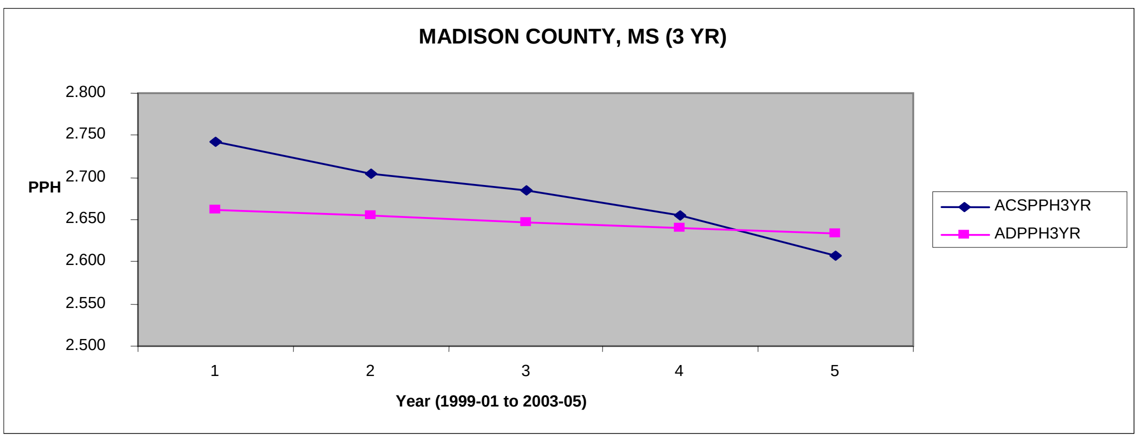


EXHIBIT 11.2*



*The ACS PPH values are labeled as "ACSPPH" and the model-generated PPH values are labeled as "ADPPH" (where "AD" stands for Analytically Derived). "(1 YR)" stands for single year ACS data and "(3 YR)" stands for 3 year ACS data.

EXHIBIT 12.1*

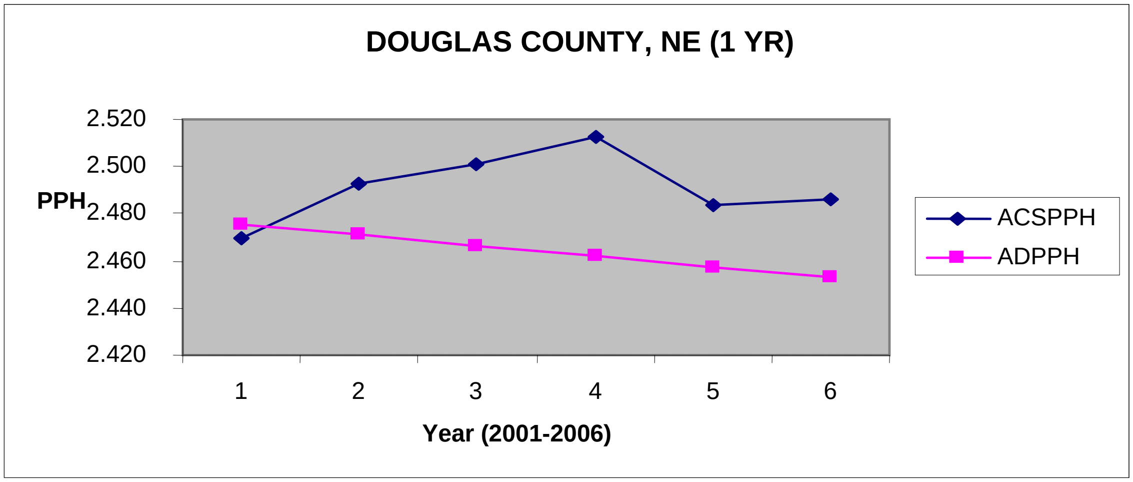
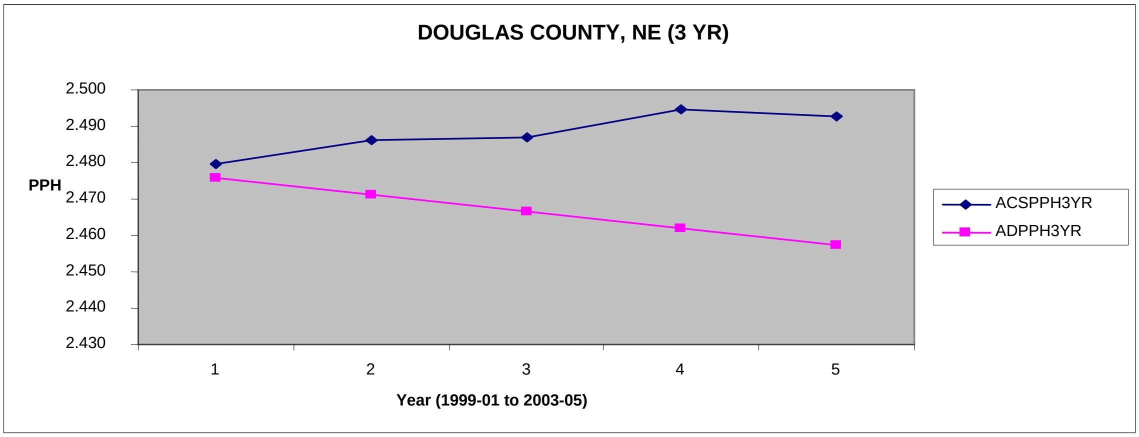


EXHIBIT 12.2*



*The ACS PPH values are labeled as "ACSPPH" and the model-generated PPH values are labeled as "ADPPH" (where "AD" stands for Analytically Derived). "(1 YR)" stands for single year ACS data and "(3 YR)" stands for 3 year ACS data.

EXHIBIT 13.1*

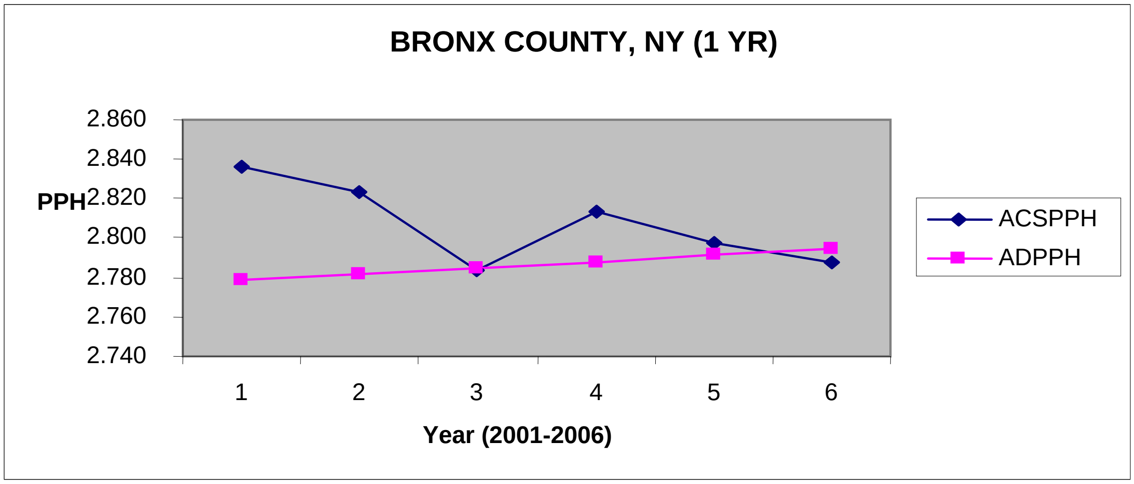
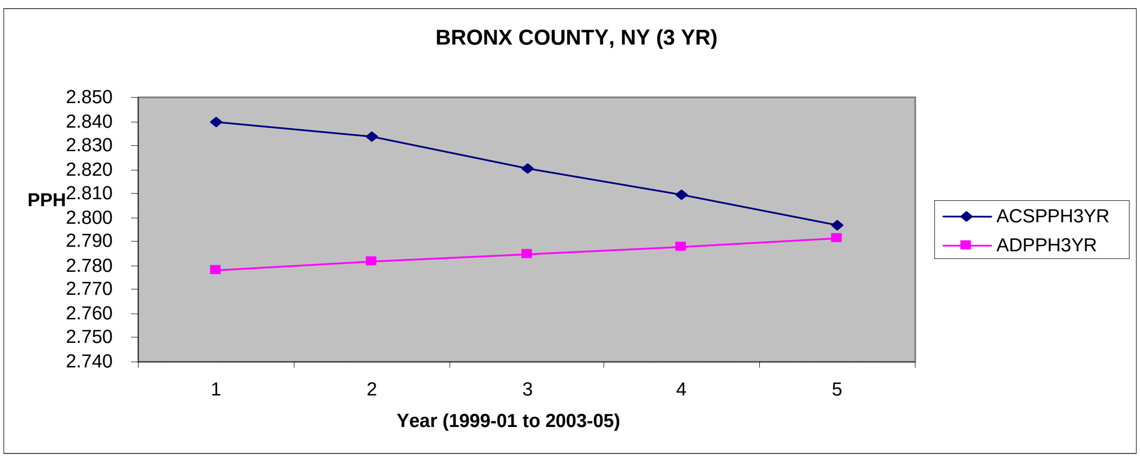


EXHIBIT 13.2*



*The ACS PPH values are labeled as "ACSPPH" and the model-generated PPH values are labeled as "ADPPH" (where "AD" stands for Analytically Derived). "(1 YR)" stands for single year ACS data and "(3 YR)" stands for 3 year ACS data.

EXHIBIT 14.1*

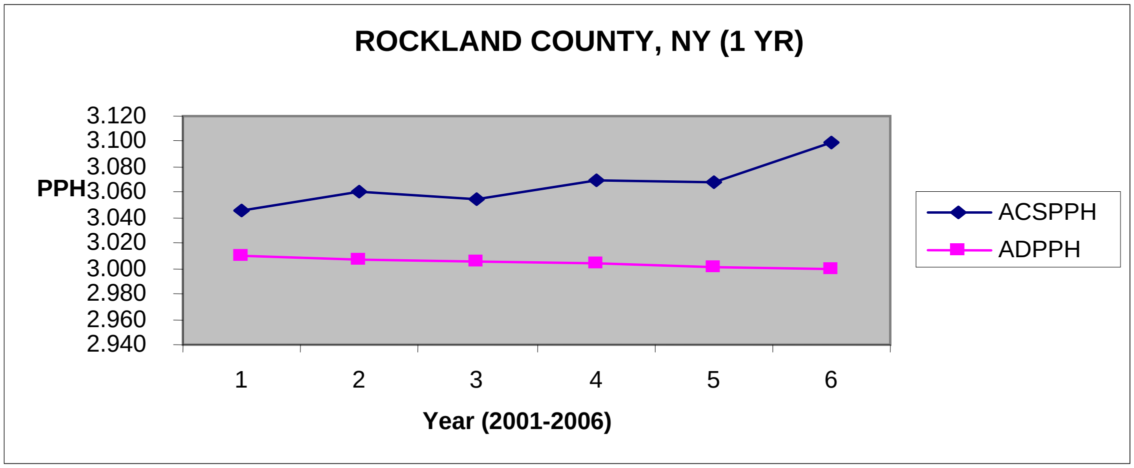
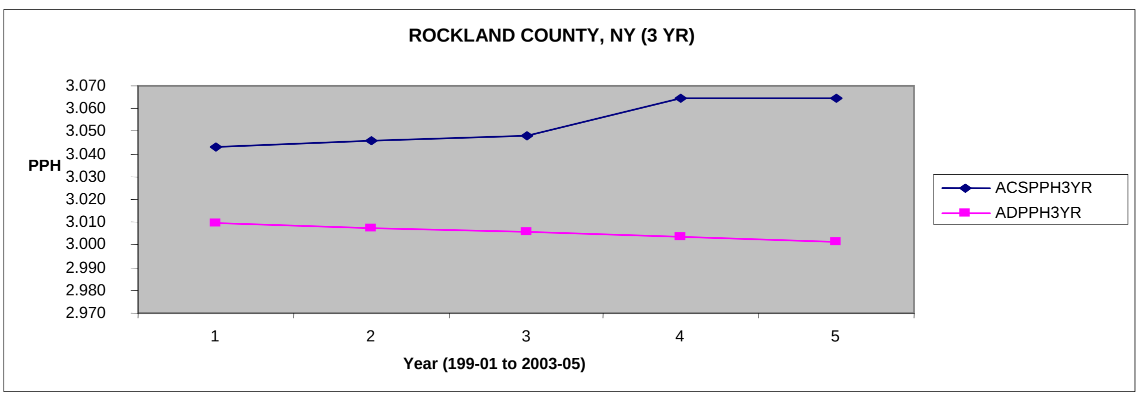


EXHIBIT 14.2*



*The ACS PPH values are labeled as "ACSPPH" and the model-generated PPH values are labeled as "ADPPH" (where "AD" stands for Analytically Derived). "(1 YR)" stands for single year ACS data and "(3 YR)" stands for 3 year ACS data.

EXHIBIT 15.1*

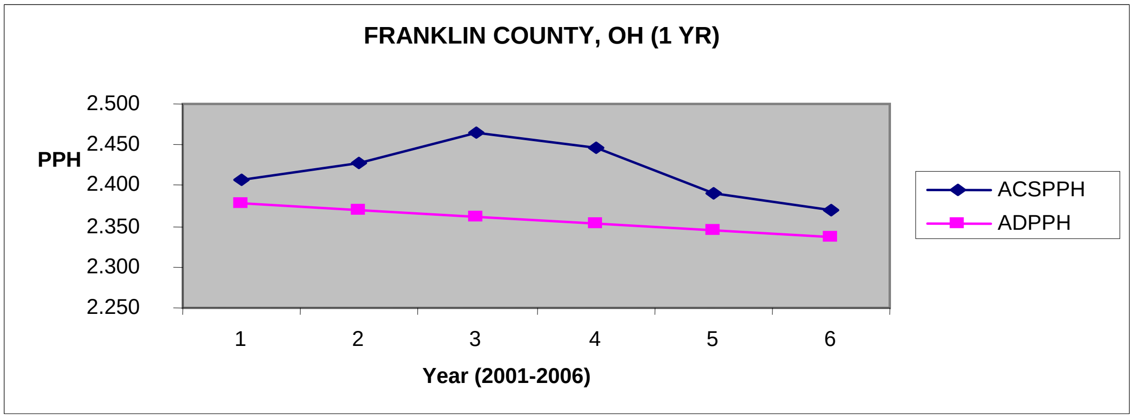
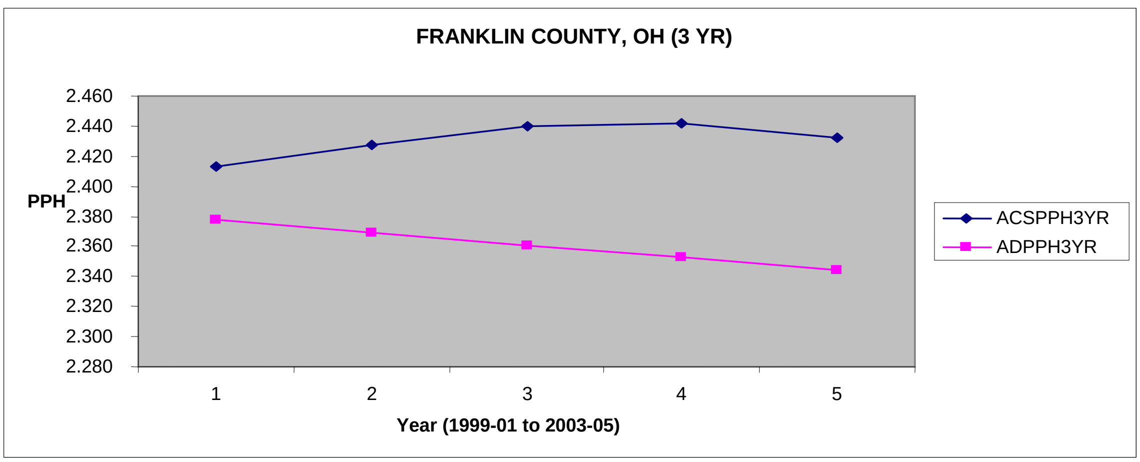


EXHIBIT 15.2*



*The ACS PPH values are labeled as "ACSPPH" and the model-generated PPH values are labeled as "ADPPH" (where "AD" stands for Analytically Derived). "(1 YR)" stands for single year ACS data and "(3 YR)" stands for 3 year ACS data.

EXHIBIT 16.1*

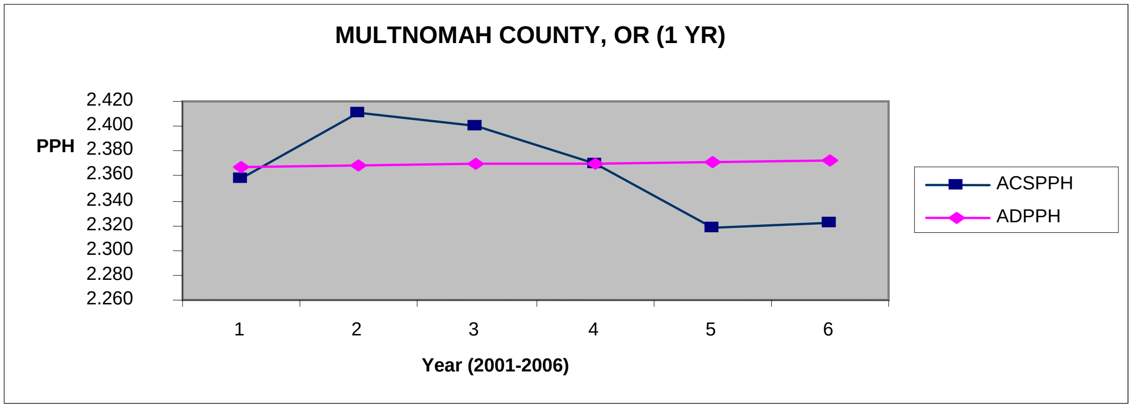
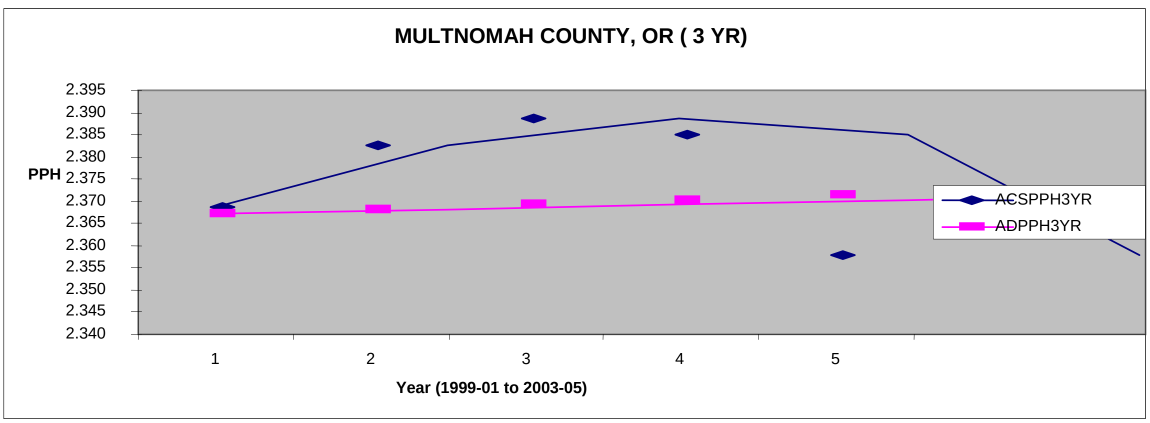


EXHIBIT 16.2*



*The ACS PPH values are labeled as "ACSPPH" and the model-generated PPH values are labeled as "ADPPH" (where "AD" stands for Analytically Derived). "(1 YR)" stands for single year ACS data and "(3 YR)" stands for 3 year ACS data.

EXHIBIT 17.1*

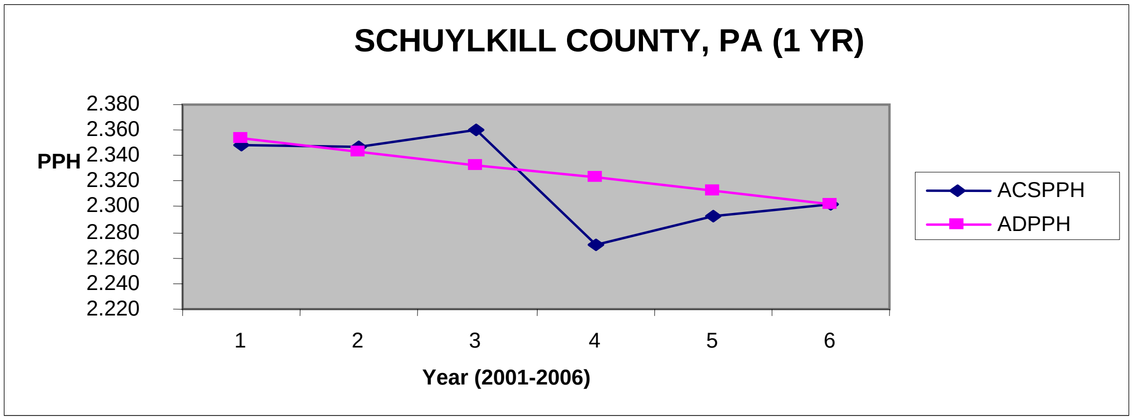
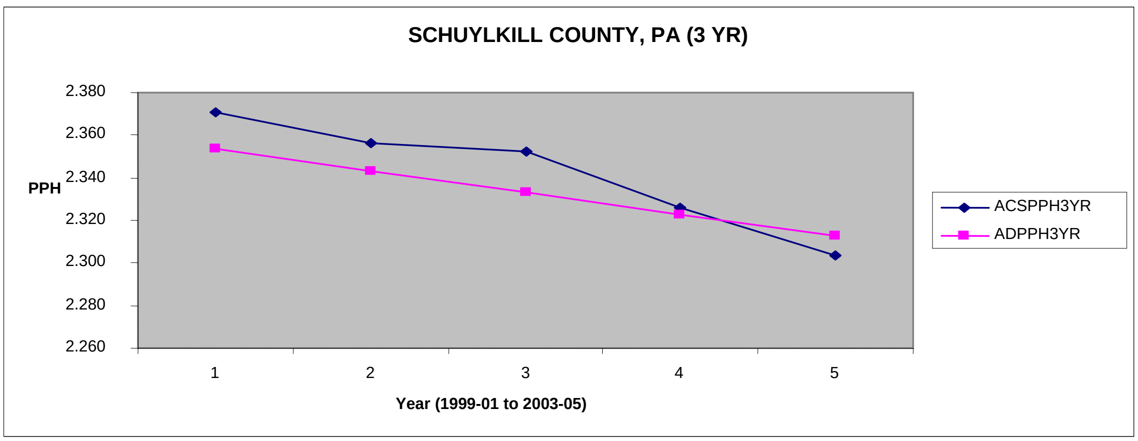


EXHIBIT 17.2*



*The ACS PPH values are labeled as "ACSPPH" and the model-generated PPH values are labeled as "ADPPH" (where "AD" stands for Analytically Derived). "(1 YR)" stands for single year ACS data and "(3 YR)" stands for 3 year ACS data.

EXHIBIT 18.1*

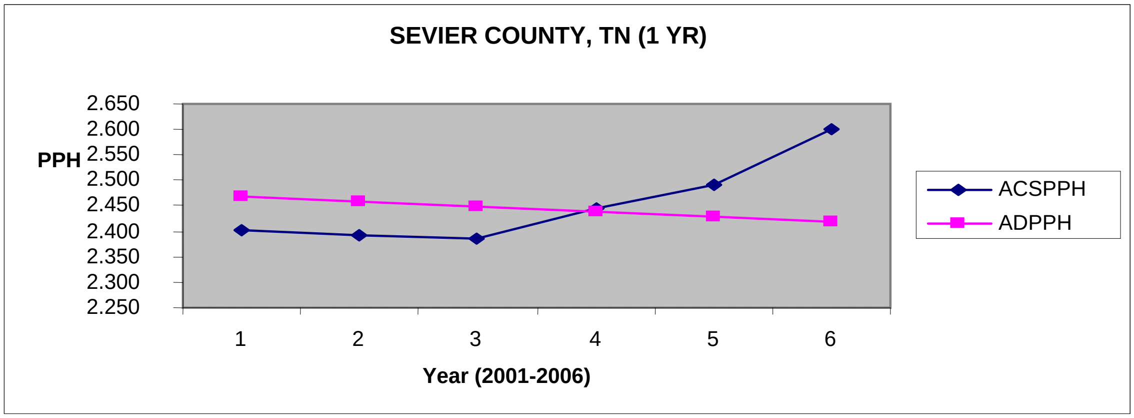
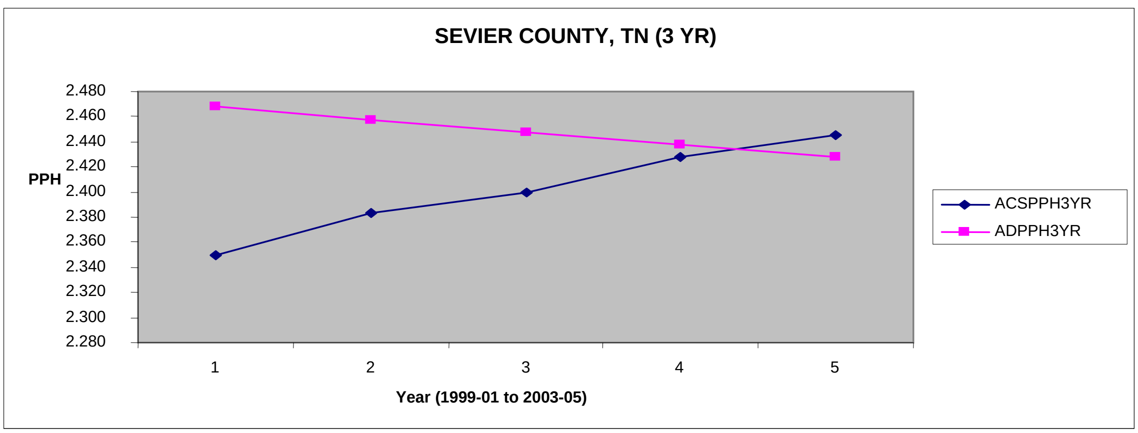


EXHIBIT 18.2*



*The ACS PPH values are labeled as "ACSPPH" and the model-generated PPH values are labeled as "ADPPH" (where "AD" stands for Analytically Derived). "(1 YR)" stands for single year ACS data and "(3 YR)" stands for 3 year ACS data.

EXHIBIT 19.1*

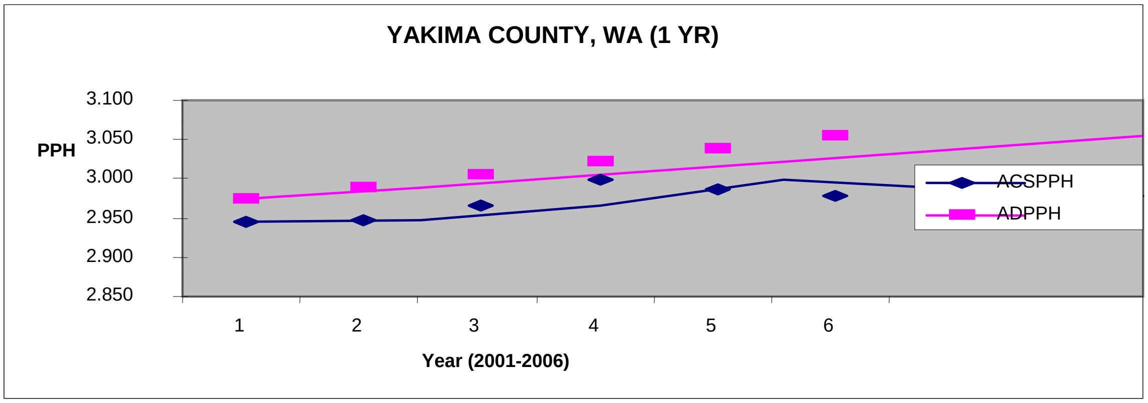
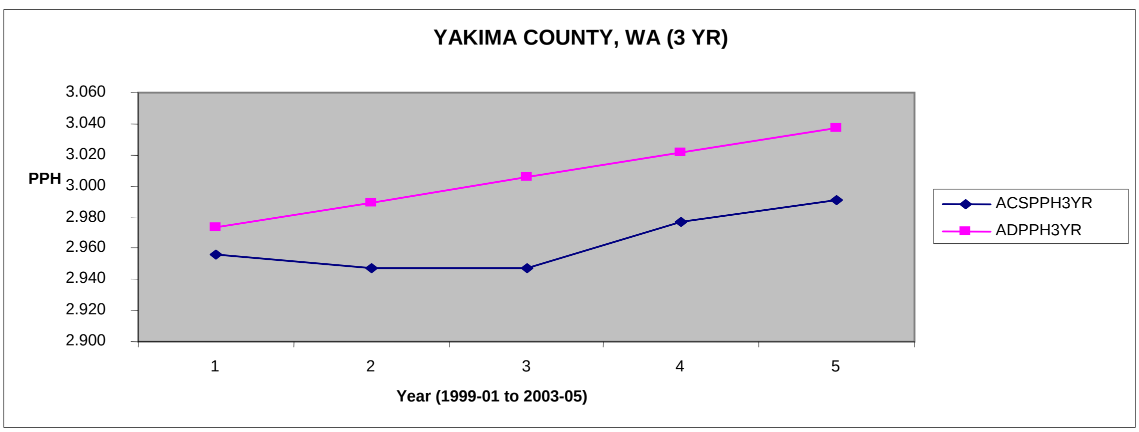


EXHIBIT 19.2*



*The ACS PPH values are labeled as "ACSPPH" and the model-generated PPH values are labeled as "ADPPH" (where "AD" stands for Analytically Derived). "(1 YR)" stands for single year ACS data and "(3 YR)" stands for 3 year ACS data.

Before discussing the results, it is important to note that the PPH values generated by geometric trend extrapolation are used as benchmarks not because they are inherently more accurate than those derived from other models or from samples such as the ACS, but, rather, because they represent the type of systematic change demographers expect to see in PPH values.

However, in order to provide evidence that county level PPH values generated by the geometric trend extrapolation method are reasonably accurate, refer to Table 1 which is found in Chapter 3.

In this test, Census 1980 and 1990 PPH values are used as input to the geometric model, which is applied to the Census 1990 PPH values to generate PPH values for 2000. These estimated PPH values are then compared to Census 2000 PPH values. The results support the argument that the geometric method is capable of generating PPH values sufficiently accurate for use in post-censal HUM estimates: (1) The mean error is 0.068; (2) the mean absolute percent error is 2.97; (3) the mean algebraic percent error is -2.60; and (4) the number of absolute percent errors that are 5.0 or greater is six.

In comparing the single-year ACS PPH values to the model-based PPH values, the ACS PPH values are above the model-based PPH values in seven counties for the entire period, 2001-2006, that they are below the model-based values in two counties for the entire period and cross over the model-based values in nine counties (three of which (Bronx, Multnomah, and Schuylkill) have two crossovers each and one of which (Jefferson) has three crossovers). In terms of directional changes, the single-year ACS PPH values change direction three or more times in three counties, twice in nine counties, and once in six counties.

The three-year ACS PPH values remain above the model-based values for the entire period , 1999-2001 to 20003-2005 in nine counties, while in only one county (Yakima) they remain below the model-based values, and cross over the model based values nine times. The three-year ACS PPH values change direction twice in two counties and once in seven counties. In the remaining nine counties no directional changes are observed, although there are some in which trends become flattened for some of the time. The model-based PPH values show a secular decline in 11 counties and an increase in seven.

In some of the counties with declining model-based PPH values, the trends are very slight (e.g., Pima County, Arizona) and in others, more distinct (Schuylkill County, Pennsylvania). Similarly, some of the counties with increasing model-based PPH values have a very slight upward trend (e.g., San Francisco County, California), in others they are much more pronounced (e.g., Tulare County, California).

Table 2 provides mean PPH values across the 18 counties (and their standard deviations) by year. Not surprisingly, the single-year ACS PPH values exhibit the least systematic change over time and the most variation each year. In two of the six years, these values are less than the model-based PPH values while in the remaining four years they exceed the model-based values The means of the three-year ACS PPH values show a systematic decline over time with annual variations comparable to the model-based PPH values.

Discussion of ACS PPH Data

As noted earlier, the U. S. Census Bureau found encouraging results for the three-year ACS PPH values among the set of 1999 test counties when it compared the 1999-2001 numbers to the PPH values of the 2000 Census (U. S. Census Bureau 2004b). As also was noted earlier, this finding was no surprise because the total household population and the total number of housing units found in Census 2000 are used as control variables in ACS weighting. Given this, the results found here are a bit discouraging, given that these same variables are also used as control variables in ACS weighting – with one major change – once beyond the 2000 census, the total household populations and housing units are not enumerated directly, but, instead, estimated.

Table 2. Mean ACS Values by Year and Their Standard Deviations

Year	Mean 1-Year ACS PPH Values*	Mean Model-Based PPH Values*	Mean 3-Year ACS PPH Values*	Year
2001	2.503 (0.295)	2.627 (0.281)	2.648 (0.290)	1999-2001
2002	2.509 (0.287)	2.625 (0.286)	2.647 (0.286)	2000-2002
2003	2.642 (0.294)	2.622 (0.289)	2.642 (0.289)	2001-2003
2004	2.647 (0.319)	2.620 (0.300)	2.644 (0.300)	2002-2004
2005	2.623 (0.323)	2.618 (0.303)	2.635 (0.312)	2003-2005
2006	2.717 (0.312)	2.625 (0.309)	N/A	N/A
*The value shown in parentheses is the standard deviation (N=18)				

Not surprisingly, it is the single-year ACS PPH estimates that are the most discouraging. They jump around a great deal from year to year in many of the counties, a characteristic that is not desirable for both demographers who use the HUM and the stakeholders for whom HUM estimates are done. This is because there is an expectation on the part of both these demographers and the stakeholders that PPH values should exhibit systematic changes unless there is compelling substantive evidence (e.g., the PPH values jumped because of a surge of in-migrants with high fertility and large family sizes) to the contrary. If such PPH values are used in the absence of compelling substantive evidence justifying their temporal instability then it appears to me that the risk of challenges and related administrative and legal actions increases (See, e.g., Walashek and Swanson, 2006), especially when these estimates are used to allocate resources, which is often the case (National Research Council, 1980, 2003; Scire, 2007).

In considering the three-year ACS PPH values, the results are not as discouraging, as those for the single-year values, but neither are they strongly encouraging. These values change more systematically than do the single-year ACS PPH values, but they still exhibit temporal instability. However, how one uses the three-year data is not very clear. Can one use them for the first year of the three-year interval as well as the second and third years? The Census Bureau stresses that they are interval rather than point estimates, but the fact that they are linked to point estimates (e.g., controlled to annual population estimates) and are needed in return for point estimates (e.g., annual population estimates) leaves many questions about their use.

In addition to the temporal instability issue, which is itself partly a function of statistical variance over time, one must ask what causes some of the substantial differences observed between the mean ACS PPH values and the mean model-based PPH values. For example, in 2001, the mean ACS PPH is 2.503 while the model-based mean is 2.627. This is a substantial difference, one likely beyond the scope of simple sampling error. Is this difference partly due to the ACS residency rule? After all, it is not the same as the Decennial Census residency rule, the one that is inherent in the model-based ACS PPH values. With a two-month rule, the ACS clearly will tend to have higher PPH values in areas in which seasonal migrants are currently residing than would be the case with the "majority of your time" rule used by the Decennial Census. This might explain in part the higher ACS PPH values found in Pima County, Arizona, However, if this were the case, one would expect that the ACS PPH values would consistently be higher than the model-based PPH values in Tulare County, California, but they are not.

The ACS as a Source of PPH Data

As described at the start of this section, the ACS provides annual PPH estimates that are subject to sample (and non-sample) error. This means that they can fluctuate from year to year in a given population, which reflects a "Statistical Perspective." Demographers, however, tend to view PPH as a population attribute that has demographic determinants. This implies that demographers view PPH as an attribute that changes systematically over time - the "Demographic Perspective." The comparisons suggest that the ACS PPH values exhibit too little systematic change over time for a given area to be usable by demographers and others preparing post-censal population estimates.

The finding that the ACS PPH values are not particularly usable for purposes of making HUM-based population estimates is preliminary in nature. More work needs to be done not only to confirm this finding, but also to figure out if the ACS PPH values can be modified so that they could be used if the finding is confirmed. With this in mind, our suggestions for further analysis include: (1) conducting a broader scale comparison, taking into account the full range of counties; (2) examining ACS PPH values that are not controlled; (3) consideration of a way to utilize sample error (i.e., confidence intervals) in determining ACS PPH changes over time; (4) an examination of 5-year ACS PPH values when at least five years of data become available; and (5) making adjustments to ACS PPH values (deriving model-based PPH values from the ACS) that may provide more temporal stability.

The ACS is a resource of high potential value to all stakeholders and ACS PPH values represent the same type of resource to demographers making population estimates and their stakeholders (See, e.g., Smith, 1998). The goal of our suggestions for further research is to see if the ACS PPH values can become usable in terms of the demographic perspective, especially as implemented in HUM-based estimates.

In conclusion, I note that differences between statisticians and demographers are stressed n this section. However, a demographic perspective is not incompatible with a statistical perspective. At one level, the demographic perspective can be viewed as a model-based approach, a perspective that is shared with statistics (Hill, 1990; Jiang and Lahiri, 2006). Further, as noted throughout this report, demographers view PPH as a variable that responds to demographic and related determinants. Thus, at another level, the demographic perspective described here represents 'causality.' This also is a perspective that is shared with statistics (Cox and Wermuth, 2004). Finally, at a third level, the demographic perspective is empirical, which also is a perspective that is shared with statistics – Stigler (1986: 1) observes, for example, that ""..Modern statistics provides a quantitative technology for empirical science;…"

In short, the view that PPH is a variable that responds to demographic and related determinants is not only worthy of consideration, but one that is compatible with statistics, broadly speaking. I have identified three shared commonalities - a model-based perspective, a causal perspective, and an empirical perspective - that support this argument.

Summary

This review and evaluation of data and methods suggests that the Census Bureau continue with its development of a comprehensive system based on the HUM for the annual estimates of county populations nationwide. Ideally, this system would include current and historical housing unit data from previous census counts and the MAF as well as: (1) VR and PPH data from at least the two most recent decennial census counts; (2) postal deliveries for the same census years and the years since the most recent census; (3) for the same census years and years and all other available years, IRS data, Medicare enrollment, covered employment, school enrollment, births and deaths and other data available for all counties that have consistent definitions nationwide; and (4) ACS data, such as PPH and VR. I suggest that the system be set up to handle HUM estimates by structure type, knowing that for some data elements the structure type data are not available (e.g., the IRS data). This system would also include the standard geographic identifiers used by the Census Bureau along with derived classifications such as metropolitan and non-metropolitan status. In addition, it could include markers for the presence of seasonal populations and special populations by type (military, college dormitory populations, prisons, etc.).

With these and other data (as determined in a more rigorous examination), the system could be set up so that it has default actions (e.g., carry the VR and PPH values forward from the previous census) that are subject to automated checks that would 'flag' areas if the indicators used I the checks suggested that the default procedures may not be optimal (e.g., since the last census the IRS data show substantial in-migration as do the school enrolment data and the MAF shows added single unit housing structures) . If the flags are triggered, then a series of subsequent flags could indicate which other data (and their underlying methods) might be better suited (e.g., use the change in USPS delivery data to modify VR). Having this capability, of course, implies that the various methods need to be actively implemented in the system. For example, if PPH for a given county is flagged such that holding it constant since the last census is not optimal, then the geometric method or a ratio–regression model should be ready to provide alternatives, one of which could be selected by an analyst using established procedures.

In addition to developing the internal data, methods, and administrative procedures for such a system, I suggest that the Census Bureau continue to pursue external links. The HUBERT program serves as an excellent starting point. It could be the case, for example, that some states in the FSCPE have developed excellent systems for capturing changes stock that the Bureau could use. I also suggest that the regional offices be considered as active partners. Here, I cam thinking of the possibility of doing windshield and other surveys to update data for counties in which all of the ‘flags’ and subsequent analyses suggest that changes are of a magnitude that field work is required to obtain good VR and PPH data. If this was implemented, field staff training would be needed relative to optimal data capture methods and the channels for sending these data back into Suitland.

As noted by Swanson(2006) in his congressional testimony, the Census Bureau and its state and local government partners (as well as the private sector) would all benefit from a universal HUM system employing a functional MAF. He also stated that be believed that the problems he identified in achieving this goal were solvable. His suggested approach requires new thinking, new arrangements, and new relationships, not only by the Census Bureau and its traditional partners in the public sector, but also by the private sector. In developing such a system, I suggest that in addition to the Bureau’s own experience with developing and maintaining such systems (e.g., Devine and Coleman, 2003; Judson, 2000; Long, 1993a, 1993b; Marquis, Wetrogan, and Palacios, 1996; Prevost, 1996; Prevost and Leggieri, 1999; Wetrogan, 2007) , it consider the ideas presented by Tayman (1986) in regard to the integrated system used by the San Diego Association of Governments for preparing census tract level estimates, among others (e.g., Alaska Department of Labor, 1981; Kimpel and Lowe, 2007; Lowe and Mohrman, 2003; Lowe, Pittenger, and Walker, 1977; Swanson, Baker, and Van Patten, 1983)

Finally, there is the ACS. In its current stage, the PPH data it generates are not suitable as inputs into an HUM system “as is.” Substantial massaging needs to be done to iron out the temporal instabilities and the large variances found for PPH values in many areas. I have not examined them, but I suspect that there high levels of variance associated with VR estimates from the ACS and, as noted earlier, it is much more difficult to sort out variance from change for VR than PPH because of the difference in the determinants underlying them. However, ACS data should be an element in the comprehensive system I am proposing. Using the VR, PPH, and other data it produces will require among other things, a good understanding of the effects of the differences in residency definitions between the decennial census and the ACS. For many counties, these differences may substantially interfere with the use of "raw" and even "controlled" ACS data. As the ACS matures along with the analysts who use data from it, the residency and other issues will likely become resolved. As this occurs, some of the methods I have proposed will become less likely to be used (e.g., holding VR constant since the last census) while others become normal (e.g., using regression based methods in conjunction with census and ACS data to update PPH).

In conclusion, I commend the Census Bureau for considering the HUM and trust that whatever decision it makes in regard to the methods and systems used in its annual estimates program will be well-considered.

Endnotes to the Report

1. The US Census Bureau document distributed at this conference uses the term “inter-censal estimate” in its title, while the document itself clearly makes reference to post-censal estimation work (U.S. Census Bureau, n.d.). I believe, however, that the distinction between inter-censal and post-censal is worth maintaining.

2. For the record, one can also construct estimates for a point in time that predates a census. I have not run across the term “pre-censal,” however and so do not use it here. Here it also is useful to note that there is a large body of literature on how to make estimates of populations and their characteristics for countries that lack censuses and good registration systems (Popoff and Judson, 2004). There are also methods developed for the estimation of wildlife populations that can be used with special populations such as the homeless – “capture-recapture” and “transit surveys,” for example (Williams, Nichols and Conroy, 2002). However, as is the case with the “statistical” tradition, I do not cover the estimation methods associated with “statistically underdeveloped areas” and wildlife populations

3. The MAF is already being used for “direct estimation” because it forms the sample frame for the Census Bureau’s “American Community Survey.”

4. The synthetic method of estimation is defined by Swanson and Stephan (2004: 776) as “a member of the family of ratio estimation methods used to estimate characteristics of a population in a sub-area (e. g., a county) by re-weighting ratios (e.g., prevalence rates or incidence rates) obtained from a survey or other data available at a higher level of geography (e.g., a state) that includes the sub-area in question.” As alluded to in the preceding definition, the synthetic method is usually viewed as belonging to the statistical tradition because of its frequent use with survey data. For a description of the synthetic method see Judson and Popoff (2004: 681-683). I also note that the “composite” method (Bryan, 2004b: 550-551) is a type of synthetic estimation.

5. While the United States lacks a national population registration system there are, as noted in the body of the report, administrative records in the private sector that contain information on people that is used for commercial purposes (e.g., credit reporting systems such as those operated by Equifax, Experian, and TransUnion). Experian also conducts consumer marketing activities (See endnote # 9). These systems can be used to generate population estimates. However, using them requires money and the accuracy of such estimates is hard to judge because of the proprietary nature of the data.

6. Although their discussion of such adjustments is in the context of making projections rather than estimates, Smith, Tayman, and Swanson (2001: 239-277) provide a comprehensive description that covers many of the same issues found in developing estimates.

7. This section is adopted from Swanson and Hough (2007).

References

Alaska Department of Community and Regional Affairs. 1981a. *The Housing Unit Method of Population Estimation: A Manual for Municipal Personnel Responsible for Annual Population Estimates.* Juneau, AK: Alaska Department of Community and Regional Affairs, Division of Local Government Assistance and Alaska Department of Labor, Research and Analysis Section.

Alaska Department of Community and Regional Affairs. 1981b. *Standards for Conducting a Population Census in Small Alaskan Cities.* Juneau, AK: Alaska Department of Community and Regional Affairs, Division of Local Government Assistance.

Alaska Department of Labor. 1981. *State of Alaska Census Enumerator's Manual*. Juneau, AK: Alaska Department of Labor, Research and Analysis Section.

Alaska Department of Labor. 1982. *State of Alaska Census Administrator's Manual*. Juneau, AK: Alaska Department of Labor, Research and Analysis Section.

Brass, W., A. 1968. *The Demography of Tropical Africa*. Princeton, NJ: Princeton University Press.

Bryan, T. 2004a. "Basic Sources of Statistics." pp. 9-41 in J. Siegel and D. Swanson (Eds.) *The Methods and Materials of Demography, 2nd Edition*. New York, NY: Elsevier Academic Press.

Bryan, T. 2004b. "Population Estimates." pp. 523-560 in J. Siegel and D. Swanson (Eds.) *The Methods and Materials of Demography, 2nd Edition*. New York, NY: Elsevier Academic Press.

Bryan, T. and R. Heuser. 2004. "Collection and Processing of Demographic Data." pp. 43-63 in J. Siegel and D. Swanson (Eds.) *The Methods and Materials of Demography, 2nd Edition*. New York, NY: Elsevier Academic Press.

Burch, T. 1970. "Some Demographic Determinants of Average Household Size: An Analytic Approach." *Demography* 7 (1): 61-69.

Burch, T. 1967. "The Size and Structure of Families: A Comparative Analysis of Census Data." *American Sociological Review* 32 (3): 347-363.

Burch, T., S. Halli, A. Madan, K. Thomas, and L. Wai. 1987. "Measures of household composition and headship based on aggregate routine census data." pp. 19-33 in J. Bongaarts, T. Burch and K. Wachter (eds.) *Family Demography: Methods and Their Application*. Oxford: Clarendon Press.

Byerly, E. 1990. "State and Local Agencies Preparing Population and Housing Estimates." *Current Population Reports, Series P-25-1063*. Washington, DC: U. S. Government Printing Office.

Chandrasekaran (Chandra Sekar), C. and W. E. Deming. 1949. "On a Method of Estimating Birth and Death rates and the Extent of Registration." *Journal of the American Statistical Association* 44: 101-115.

Citro, C. and G. Kalton (Eds.). 2007. *Using the American Community Survey: Benefits and Challenges.* Panel on the Functionality and Usability of Data from the American Community Survey. National Research Council. Washington, DC: The National Academies Press.

Coale, A. 1965. "Estimates of Average Size of Household." pp. 64-69 in A. Coale, L. Fallers, M. Levy, D. Schneider, and S. Tompkins (Eds.) *Aspects of the Analysis of Family Structure.* Princeton, NJ: Princeton University Press.

Cook, T. 1996. "When ERPs Aren't Enough: A Discussion of Issues Associated with Service Population Estimation." *Working Paper 96/4*. Demography Section, Australian Bureau of Statistics, Belconnen, ACT, Australia.

Cork, D., M. Cohen, and B. King (Eds.). 2004. *Reengineering the 2010 Census: Risks and Challenges*. Panel on Future Census Methods, National Research Council. Washington, DC: National Academies Press.

Cox, D. and N. Wermuth. 2006. "Causality: A Statistical View." *International Statistical Review* 72 (3): 285-305.

Devine, J. and C. Coleman. 2003. "People Might Move but Housing Units Don't: An Evaluation of the State and County Housing Unit Estimates." *Population Division Working Paper Series No. 71*. Washington, DC: U. S. Census Bureau (available online at http://www.census.gov/population/www/documentation/twps0071/twps0071.html., last accessed October, 2008).

Devlin, K. 2008. *The Unfinished Game: Pascal, Fermat, and the Seventeenth-Century Letter that Made the World Modern.* New York, NY: Perseus Publishers.

Edelstein, R. and D. Tang. 2007. "Dynamic Residential Housing Cycles Analysis." *The Journal of Real Estate Finance and Economics* 35 (3): 295-313.

Eldridge, H. 1947. "Problems and Methods of Estimating Post-censal Population." *Social Forces* 24: 41-46.

Erickson, E. 1974. "A Regression Method for Estimating Population Changes of Local Areas." *Journal of the American Statistical Association* 69: 867-875.

Ericksen, E. 1973. “A Method for Combining Sample Survey Data and Symptomatic Indicators to obtain Population Estimates for Local Areas.” *Demography* 10: 137-160.

Fay, R. 2005. “Model-assisted estimation for the American Community Survey.” pp. 3016-2023 in *Proceedings of the Joint Statistical Meetings, ASA Section on Survey Research Methods*. Alexandria, VA: American Statistical Association.

Fay, R. 2007. “Variance Reduction for Sub-County Estimates in the American Community Survey.” Paper presented at the Spring Meeting of the Census Advisory Committee of Professional Associations. Suitland, MD: April 19th.

Frisėn, C. 1958. "Effectiveness of Our Tools for Estimating Population Change in Small Areas," pp. 229-232 in *Proceedings of the American Statistical Association, Social Statistics Section*. Alexandria, VA: American Statistical Association.

Frisén, C. 1951. “Symptomatic Data in Population Estimates: Problems and Limitations.” Paper presented at the annual meeting of the Pacific Sociological Association, Berkeley, CA (published in *Research Studies of the State College of Washington*, vol. 19, pp. 107-110).

Gabriel, S. A., and F. E. Nothaft. 2001. “Rental Housing Markets, The Incidence and Duration of Vacancy, and the Natural Vacancy Rate.” *Journal of Urban Economics* 49 (1): 121-49.

Goldsmith, H. D. Jackson, and J. Shambaugh. 1982. “A Social Analysis Approach.” pp. 169-190 in E. Lee and H. Goldsmith (eds.) *Population Estimates: Methods for Small Area Estimates*. Beverly Hills, CA: Sage Publications.

Griffin, D. and P. J. Waite. 2006. “American Community Survey Overview and the Role of External Evaluations.” *Population Research and Policy Review* 25: 201-223.

Habermann, H. 2006. “Research to Improve Population Estimates” Part of a presentation by H. Habermann at the Spring (May 18th -19th) 2006 meeting of the Census Advisory Committee for Professional Associations.

Hauser, P., and B. Tepping. 1944. "Evaluation of Census Wartime Population Estimates and of Predictions of Postwar Population Prospects for Metropolitan Areas.” *American Sociological Review* 9: 473-480.

Hendershott, P. H., B. D. MacGregor and R. Y. C. Tse. 2002. "Estimation of The Rental Adjustment Process." *Real Estate Economics* 30 (2):165-183.

Hill, J. 1990. “A General Framework for Model-Based Statistics.” *Biometrika* 77 (1): 115-126.

Hogan, H. 2007. “Measuring Population Change using the American Community Survey.” pp. 13-30 in S. Murdock and D. Swanson (eds.) *Applied Demography in the 21st Century.* Dordrecht, The Netherlands: Springer.

Hough, G., and D. Swanson. 1998. "Towards an assessment of continuous measurement: A comparison of returns with 1990 census returns for the Portland test site," *Journal of Economic and Social Measurement* 24: 295-308.

Hough, G. and D. Swanson. 2006 "An Evaluation of the American Community Survey: Results from the Oregon Test Site." *Population Research and Policy Review* 25(3): 257-273.

Hsueh, L., H. Tseng, and C. Hsieh. 2007.. "Relationship Between the Housing Vacancy Rate, Housing Price, and the Moving Rate at the Township Level in Taiwan, in 1990 and 2000." *International Real Estate Review* 10 (1): 119-150.

Jiang, J. and P. Lahiri. 2006. "Mixed Model Prediction and Small Area Estimation." *Test* 15 (1): 1-96.

Judson, D. 2000. "The Statistical Administrative Records System and Administrative Records Experiment 2000." Paper presented at the National Institutes of Statistical Sciences Data Quality Workshop, Morristown, NJ, November 30th – December 1st .

Judson, D. and C. Popoff. 2004. "Selected General Methods." pp. 677-732 in J. Siegel and D. Swanson (Eds.) *The Methods and Materials of Demography, 2nd Edition*. New York, NY: Elsevier Academic Press.

Khor, A., Y Ming, and L. Yuan. 2000. "The Natural Vacancy Rate of the Singapore Office Market." *Journal of Property Research* 17 (4): 329-338.

Kimpel, T. and T. Lowe. 2007. "Estimating Household Size for Use in Population Estimates." *Population Estimates and Projections, Research Brief no. 47.* Olympia, WA: Washington State Office of Financial Management.

Kintner, H. and D. 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.

Kish, L. 1998. "Space/Time Variations and Rolling Samples." *Journal of Official Statistics* 14 (1): 31-46.

Kordos, J. (Ed.). 2000. "Special issue on Small Area Estimation." *Statistics in Transition: Journal of the Polish Statistical Association* 4 (4).

Korbin, F. 1976. "The Fall in Household Size and the Rise of the Primary Individual in the United States." *Demography* 13 (1): 127-138.

Long, J. 1993a. "An Overview of Population Estimation Methods." pp 20-1 to 20-5 in D. Bogue, E. Arriaga and D. Anderton (Eds.) *Readings in Population Research Methodology* (Vol. 5), Chicago: United Nations Population Funds and Social Development Center.

Long, J. 1993b. *Postcensal Population Estimates: States, Counties and Places*. Technical Working Paper No. 3. Washington DC: U.S. Bureau of the Census.

Lowe, T. 2009. Personal Communication.

Lowe, T. 2000a "Developing Trends in Household Size for use in Population Estimates." *Population Estimates and Projections Research Brief No. 10*. Olympia, WA: Washington State Office of Financial Management.

Lowe, T. 2000b "Occupied versus Rented: Census and Real Estate Occupancy Rates and Their Use in Population Estimates." *Population Estimates and Projections Research Brief No.* 9 (May). Olympia, WA: Washington Office of Financial Management.

Lowe, T., and M. Mohrman. 2003. "Use of Postal Delivery Data in the Population Estimate Process." *Population Estimates and Projections Research Brief No.* 18 (January). Olympia, WA: Washington Office of Financial Management.

Lowe, T., M. Mohrman and D. Brunink. 2003. "Developing Postal Delivery Data for Use in Population Estimates." *Population Estimates and Projections Research Brief No.* 17 (January). Olympia, WA: Washington Office of Financial Management.

Lowe, T., L. Weisser, and B. Myers. 1984. "A Special Consideration in Improving Housing Unit Estimates: The Interaction Effect. Paper presented at the annual meeting of the Population Association of America, Minneapolis, MN.

Lowe, T., D. Pittenger, and J. Walker. 1977. "Making the Housing Unit Method Work: A Progress Report. Paper presented at the Annual Meeting of the Population Association of America, St. Louis, MO.

Marquis, K., S. Wetrogan, and H. Palacios. 1996. "Towards a U.S. Population Database from Administrative Records." Unpublished Paper, U.S. Census Bureau (Available online, http://www.census.gov/srd/papers/pdf/km9601.pdf, last accessed May, 2008).

McDonald, J. and D. McMillen. 2000. "Residential Building Permits in Urban Counties: 1990–1997." *Journal of Housing Economics* 19 (3): 175-186.

McKibben, J. 2006. "School District Planning and the 2010 Census: Data Uses and Needs". *Journal of Economic and Social Measurement*, 31, (3) 221-232.

Murdock, S. and D. Ellis. 1991. *Applied Demography: An Introduction to Basic Concepts, Methods, and Data*. Boulder, CO: Westview Press.

Myers, D. and A. Doyle. 1990. “Age-specific Population-per-Household Ratios: Linking Population Age Structure with Housing Characteristics.” pp. 109-130 in D. Myers (Ed.) *Housing Demography: Linking Demographic Structure and Housing Markets*. Madison, WI: University of Wisconsin Press.

National Research Council. 2007. *Using the American Community Survey: Benefits and Challenges*. Washington, DC: National Academies Press.

National Research Council. 2006a. *Once, Only Once, and in the Right Place: Residence Rules in the Decennial Census*. Washington, DC: National Academies Press.

National Research Council. 2006b. *Expanding Access to Research Data: Reconciling Risks and Opportunities.* Washington, D.C: National Academies Press.

National Research Council. 2003. *Statistical issues in Allocating Funds by Formula*. Washington, DC: National Academies Press.

National Research Council. 1980. *Estimating Population and Income of Small Areas*. Washington, D.C.: National Academies Press.

Perrone, S. 2008. “Address Coverage and Improvement and Evaluation Program – 2005 National Estimate of Coverage of the Master Address File.” pp. 37-85 in S. Murdock and D. Swanson (Eds.) *Applied Demography in the 21st Century*. Dordrecht, The Netherlands, Springer.

Pittenger, D. 2004. “Use of Census Data and City Housing Reports in Creating Housing-Unit Demolition Rates.” *Population Estimates and Projections Research Brief No. 28*. Olympia, WA: Washington State Office of Financial Management.

Pittenger, D. 1982. “Critique of Administrative Record Procedures.” pp. 39-42 in E. S. Lee and H. F. Goldsmith (Eds.), *Population Estimates: Methods for Small Area Analysis*. Beverly Hill, CA: Sage Press.

Platek, R., J. Rao, C. Sarndal, and M. Singh (Eds.) (1987). *Small Area Statistics: An International Symposium*. New York, NY: John Wiley.

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

Popoff, C. and D. Judson. (2004). “Some Methods of Estimation for Statistically Underdeveloped Areas.” pp. 603-641 in J. Siegel and D. Swanson (Eds.) *The Methods and Materials of Demography, 2nd Edition*. New York, NY: Elsevier Academic Press.

Prevost, R. (1996). "Administrative Records and the New Statistical Era." Paper presented at the 1996 Annual meeting of the Population Association of America. New Orleans, LA, May 9th - 11th.

Prevost, R., and C. Leggieri. 1999. "Expansion of Administrative Records Uses at the Census Bureau: A Long-Range Research Plan." Paper presented at the conference of the Federal Committee on Statistical Methodology (available online, http://www.fcsm.gov/99papers/prevost.pdf, last accessed, May 2008).

Prevost, R. and D. Swanson. (1985). "A New Technique for Assessing Error in Ratio-Correlation Estimates of Population: A Preliminary Note." *Applied Demography* 1 (November): 1-4.

Purcell, N. J. and L. Kish. 1980. "Postcensal Estimates for Local Areas (or Domains)." *International Statistical Review* 48: 3-18.

Purcell, N. J. and L. Kish. 1979. "Estimation for Small Domains." *Biometrics* 35: 365-384.

Rao, J. (2002). *Small Area Estimation*. San Francisco, CA: Jossey-Bass.

Reese, A. J. 2006. "A Comparison of Housing Unit Estimates to the American Community Survey's Aggregated Master Address File." Paper Presented at the Annual meeting of the Southern Demographic Association, Durham, NC.

Rives, N.W. 1982. "Assessment of a Survey Approach." pp. 79-96 in E. Lee and H. Goldsmith (Eds.) *Population Estimates: Methods for Small Area Analysis*. Beverly Hills, CA: Sage Publications.

Robinson, J.G. and D. Dixon. 2009. "An Evaluation of the Consistency of Subcounty Estimates of Population Based on the American Community Survey and the Population Estimates Program." Paper (to be) presented at the annual meeting of the Population Association of America, Detroit, MI: May 2nd.

Roe, L, J. Carlson, and D. Swanson. 1992. "A Variation of the Housing Unit Method for Estimating the Population of Small, Rural Areas: A Case Study of the Local Expert Method." *Survey Methodology* 18(1): 155-163.

Serow, W. and N. Rives. 1995. "Small Area Analysis: Assessing the State of the Art." pp. 1-9 in N. Rives, W. Serow, A. Lee, H. Goldsmith, and P. Voss (eds.). *Basic Methods for Preparing Small Area Population Estimates.* Madison, WI: Applied Population Laboratory, Department of Rural Sociology, University of Wisconsin.

Shryock, H. 1938. "Methods of Estimating Post-censal Population." *American Journal of Public Health* 28: 1042-1047.

Shryock, H. and N. Lawrence. 1949. "The Current Status of State and Local Population Estimates in the Census Bureau." *Journal of the American Statistical Association* 44 (246): 157-173.

Scire, M. 2007. "2010 Census: Population Measures are Important for Federal Funding Allocations." *Testimony Before the Subcommittee on Information Policy, Census, and National Archives, Committee on Oversight and Government Reform, House of Representatives. Publication GAO-08-230T.* Washington, DC: General Accountability Office

Siegel, J. 2002. *Applied Demography: Applications to Business, Government, Law, and Public Policy.* San Diego, CA: Academic Press.

Smith, A. S. 1998. "The American Community Survey and Intercensal Population Estimates: Where are the Crossroads?" *Population Division Technical Working Paper No. 31*. U. S. Census Bureau.

Smith, S. K. 1994. "Estimating Temporary Populations: The Contributions of Robert C. Schmitt." *Applied Demography* 9 (1): 4-7.

Smith, S. 1986. "A Review and Evaluation of the Housing Unit Method of Population Estimation." *Journal of the American Statistical Association* 81: 287-296.

Smith, S., and S. Cody. 2004. "An Evaluation of Population Estimates in Florida: April 1, 2000." *Population Research and Policy Review* 23 (1): 1-24.

Smith, S. K., and M. House. 2007. "Temporary Migration: a Case Study of Florida." *Population Research and Policy Review* 26 (4): 437-454.

Smith, S., and B. Lewis, 1980. " Some New Techniques for Applying the Housing Unit Method of Local Population Estimation." *Demography* 17: 323-339.

Smith, S., and M. Mandell. 1984. "A Comparison of Population Estimation Methods: Housing Unit Versus Component II, Ratio Correlation, and Administrative Records." *Journal of the American Statistical Association* 79: 282-289.

Smith, S. K., J. Nogle, and S. Cody. 2002. "A Regression Approach to Estimating the Average Number of Persons per Household." *Demography* 39 (4): 697-712.

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

Statistics Finland. 2004. *Use of Register and Administrative Data Sources for Statistical Purposes: Best Practices of Statistics Finland*. Handbook Series, No. 45. Helsinki, Finland: Statistics Finland.

Stigler, S. 1986. *The History of Statistics: The Measurement of Uncertainty before 1990.* Cambridge, MA: Belknap Press of Harvard University Press.

Swanson, D. 2006. Statement before the Subcommittee on Federalism and the Census Oversight Hearing Committee on Government Reform, U.S. House of Representatives. Hearing Topic: "Two Plus Two Should Never Equal Three: Getting Intercensal Population Estimates Right the First Time." Wednesday, 6 September 2006 Room 2247 Rayburn House Office Building.

Swanson, D. 2004. "Advancing Methodological Knowledge within State and Local Demography: A Case Study." *Population Research and Policy Review* 23 (4): 379-398.

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. 1989. "Confidence Intervals for Post-censal Population Estimates: A Case Study for Local Areas." *Survey Methodology* 15: 271-280.

Swanson, D. 1981. "Allocation Accuracy in Population Estimates: An Overlooked Criterion With Fiscal Implications." pp. 13-21 in *Small Area Population Estimates and Their Accuracy, Series GE-41, No. 7.* U. S. Bureau of the Census. Washington, DC: US Government Printing Office.

Swanson, D. 1980. "Improving Accuracy in Multiple Regression Estimates of Population Using Principles from Causal Modeling." *Demography* 17: 413-428.

Swanson, D. 1977."An Evaluation of the 1977 Round of Municipal Vacancy Estimates for 1 & 2 Unit Structures." *Staff Document No. 42*. Office of Financial Management, Population, Enrollment and Economic Studies Division. Olympia, WA: Washington State Office of Financial Management.

Swanson, D. and G. Hough. 2007. "An Evaluation of Persons Per Household (PPH) Data Generated by the American Community Survey: A Demographic Perspective." Paper presented at the Annual Meeting of the Southern Demographic Association. Birmingham, AL. (Available online at http://www.census.gov/acs/www/AdvMeth/Multi_Year_Estimates/presentations.html, last accessed, October, 2008).

Swanson, D. and T. Lowe. 1979. "Comparative Analysis of Change in Average household Size with reference to IRS Data on Average Exemptions Per Return: Census Results from Selected Municipalities in Washington, 1970, 1977, and 1978." *Special Report No. 22.* Office of Financial Management: Olympia, WA. Washington State Office of Financial Management.

Swanson, D. and J. McKibben. 2009. "Developing Annual Population Data in the United States: New Possibilities for the 21st Century." (To be) Presented at the 2009 Conference of the International Union for the Scientific Study of Population, September 26-October 3, Marrakech, Morocco (Session 2405).

Swanson, D. and L. Pol. 2005. "Contemporary Developments in Applied Demography within the United States. *Journal of Applied Sociology* 21 (2): 26-56.

Swanson, D. and L. Pol. 2008. "Applied Demography: Its Business and Public Sector Components." in Yi Zeng (ed.) The Encyclopedia of Life Support Systems, Demography Volume. UNESCO-EOLSS Publishers. Oxford, England. (with L. Pol).(Online at http://www.eolss.net/).

Swanson, D. and G. E. Stephan. 2004."Glossary." pp. 751-778 in J. Siegel and D. Swanson (Eds.) *The Methods and Materials of Demography, 2nd Edition*. New York, NY: Elsevier Academic Press.

Swanson, D., and J. Tayman. 1995 "Between a Rock and a Hard Place: The Evaluation of Demographic Forecasts." *Population Research & Policy Review* 14: 233-249.

Swanson, D., and L. Tedrow. 1984. "Improving the Measurement of Temporal Change in Regression Models used for County Population Estimates." *Demography* 21: 361-372.

Swanson, D., B. Baker, and J. Van Patten. 1983. "Municipal Population Estimation: Practical and Conceptual Features of the Housing Unit Method." Paper Presented at the 1983 Annual Meeting of the Population Association of America, Pittsburgh, PA, April 14th - 16th. (Reprinted as pp. 99-105 in *Alaska Population Overview 1982*. Alaska Department of Labor, Juneau, AK.

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

United Nations.1969. *Methodology and Evaluation of Population Registers and Similar Systems. Studies in Methods, No. 15*. Sales No. E.69.XVII.15. New York City, NY: United Nations.

U. S. Census Bureau. (No Date). "The U.S. Census Bureau's Intercensal Population Estimates and Projections Program: Basic Underlying Principles." Unpublished document provided to participants of this conference (reproduced as Appendix A in this paper).

U. S. Census Bureau, Population Division. 2007. "County Housing Unit Estimates, April 1, 2000 to July 1, 2006." Washington, DC: Population Division, U. S. Census Bureau. (available online, http://www.census.gov/popest/housing/files/HU-EST2006_US.CSV, last accessed January, 2009).

U. S. Census Bureau. 2006. Methodology for County Housing Unit Estimates for Vintage 2006." Washington, DC: Population Division, U.S. Census Bureau(available online, http://www.census.gov/popest/topics/methodology/2006_hu_meth.html, last accessed January 2009).

U. S. Census Bureau. 2004a. "The Census Bureau's Master Address File (MAF): Census 2000 Address List Basics." Washington, DC. Geography Division, U. S. Census Bureau (available online http://www.census.gov/geo/mod/maf_basics.pdf , last accessed September, 2008).

U.S. Census Bureau. 2004b. Meeting 21st Century Demographic Needs: Implementing the American Community Survey." *Report 8: Comparisons of the American Community Survey Three-Year Averages with the Census Sample for a Sample of Counties and Tracts*. Washington, DC: U.S. Census Bureau.

U. S. Census Bureau. 2003. "American Community Survey Operations Plan. Release 1: March 2003." Washington, DC: U. S. Census Bureau.

U. S. Census Bureau. 2001a. "Meeting 21st Century Demographic Data Needs – Implementing the American Community Survey: July 2001." *Report 1: Demonstrating Operational Feasibility.* Washington, DC: U. S. Census Bureau.

U. S. Census Bureau. 2001b. The American Community Survey, Updated Information for America's Communities, November, 2001 (ACS/01-BLKT).

U.S. GAO. 2006. *2010 Census: Census Bureau Needs to take Prompt Actions to Resolve Long-Standing and Emerging Address and Mapping Challenges.* GAO-06-272. Washington, DC: U.s. Government Accountability Office (Note: Effective July 7, 2004, the GAO's legal name changed from the Government Accounting Office to the Government Accountability Office).

U.S. GAO. 2001.2000 *Census: Significant Increase in Cost Per Housing Unit Compared to 1990 Census.* GAO-02-31. Washington, DC: U.s. Government Accountability Office (Note: Effective July 7, 2004, the GAO's legal name changed from the Government Accounting Office to the Government Accountability Office).

U. S. Department of Commerce. 2002. "A Better Strategy is needed for Managing the Nation's Master Address File." *Inspection Report No. OSE-12065*. Washington, D. C.: U. S. Department of Commerce.

Van Auken, P., R. Hammer, P. Voss, and D. Veroff. 2006. "The American Community Survey in Counties with 'Seasonal' Populations." *Population Research and Policy Review* 25 (3): 275-292.

Voss, P., and H. Krebs. 1979. "The use of Federal Revenue Data for Improving Current Estimates of Average Household Size for Minor Civil Divisions: An Evaluation." *Technical Series 70-5.* Applied Population Laboratory. Madison, WI: University of Wisconsin.

Voss, P., C. Palit, B. Kale, and H. Krebs. 1995. "Censal Ratio Methods." pp. 70-88 in N. W. Rives, W. Serow, A. Lee, H. Goldsmith, and P. Voss (Eds). *Basic Methods for Preparing Small-Area Population Estimate*s. Madison, WI: Applied Population Laboratory, University of Wisconsin.

Walashek, P. and D. Swanson. 2006. The Roots of Conflicts over US Census Counts in the late 20th Century and Prospects for the 21st Century." *Journal of Economic and Social Measurement* 31 (3, 4): 185-206.

Waldrop, J. 1995. "Preface." pp. v-vi in N. Rives, W. Serow, A. Lee, H. Goldsmith, and P. Voss (eds.) *Basic Methods for Preparing Small-Area Population Estimates*. Madison, WI: Applied Population Laboratory, Department of Rural Sociology, University of Wisconsin.

Wang, C. 1999. "Development of National Accounting of Address and Housing Inventory: The Baseline Information for Post-censal Population estimates." Paper prepared for the Estimates Methods Conference, U.S. Bureau of the Census, Federal Office Building #3, Suitland, Maryland, June 8th. (available online at http://www.census.gov/population/www/coop/popconf/paper.html, last accessed 28 June 2006).

Washington Office of Financial Management. 1978. *The Housing Unit Method: A Manual for Municipal Personnel Responsible for Making Annual Population Estimates.* Olympia, WA: Washington Office of Financial Management.

Washington Office of Financial M anagement. 2007. "Developing Trends in Household Size for use in Population Estimates." *Population Estimates and Projections, Research Brief no. 10.* Olympia, WA: Washington State Office of Financial Management.

Weidman, L., R. Creecy, D. Malec, and J. Tsay. 2008. "Exploration of the Use of Empirical Bayes Procedures for Estimating Changes in Occupancy Rate and Persons Per Household." *Research Report Series (Statistics #2008-07*). Suitland, MD: U.S. Census Bureau.

Wetrogan, S. 2007. "Evaluating County Population Estimates Using a Housing Unit Based Approach." Paper presented at the Annual Meeting of the Southern Demographic Association. Birmingham, AL.

Williams, B., J. Nichols, and M. Conroy. 2002. *Analysis and Management of Wildlife Populations.* San Diego, CA. Academic Press.

Wilmoth, J. (2004). "Population Size." pp. 65-80 in J. Siegel and D. Swanson (Eds.) *The Methods and Materials of Demography, 2nd Edition*. New York, NY: Elsevier Academic Press.

Zeng, Y., K. Land, Z. Wang, and D. Gu. 2006. "U. S. family household momentum and dynamics: an extension and application of the ProFamy Method." *Population Research and Policy Review* 25: 1–41

Appendix A. Principles underlying the US Census Bureau's estimates and projections programs.

I. Background

The U.S. Census Bureau's Population Estimates and Projections program is designed to fulfill the mandates of Title 13, Section 181, of the U.S. Code.

> During the intervals between each census of population required under section 141 of this title, the Secretary, to the extent feasible, shall annually produce and publish for each State, county, and local unit of general purpose government which has a population of fifty thousand or more, current data on total population and population characteristics and, to the extent feasible, shall biennially produce and publish for other local units of general purpose government current data on total population. Such data shall be produced and published for each State, county, and other local unit of general purpose government for which data is compiled in the most recent census of population taken under section 141 of this title. Such data may be produced by means of sampling or other methods, which the Secretary determines will produce current, comprehensive, and reliable data.

A. To satisfy this mandate, the program of population estimates has grown over the years to produce the following products annually:

1. Monthly estimates of the national population of the United States by age, sex, race, and Hispanic origin
2. Annual estimates of the population of states by age, sex, race, and Hispanic origin
3. Annual estimates of the population of counties by age, sex, race, and Hispanic origin
4. Annual estimates of the total population of functioning governmental units
5. Annual estimates of the number of housing units for states and counties.

B. In addition to meeting the mandates of Title 13, these estimate products are used for a variety of purposes, including the following:

1. Controls for federally sponsored surveys, including the Current Population Survey (CPS) and the American Community Survey (ACS)
2. Allocation of federal dollars totaling over $200 billion annually
3. Denominators for various indicators, including vital statistics, per capita income, and cancer incidence rates
4. Calculation of the number of clerks the Senate hires
5. Requirements of the Federal Election Commission

6. Denominators for poverty rate estimation at selected levels of geography
7. Program planning by federal, state, local, and private entities

II. **Implicit Assumptions**

Implementation of the annual program of intercensal estimates is guided by several implicit assumptions.

A. Timely release of the annual products is critical

1. The maximum lag time between estimate date and dissemination of last data product is 12 months.
2. Annual national and state population totals must be released within 6 months of estimate date to meet requirements of IRS Bonding Authority.
3. State estimates of the population aged 18 and older must be available within 6 months of estimate date to satisfy requirements of the Federal Election Commission.
4. National and state population controls to be used for the new calendar year CPS must be available by late January of the new calendar year.
5. Estimates of state and county characteristics must be available within 9 months to meet requirements for use as population controls for the American Community Survey.
6. Estimates of functioning governmental units should be available within 12 months of estimate date for use by HUD in funds allocation.

B. Each annual production consists of a time series of estimates from the last decennial census date to the estimate date and is produced using the latest available data and the current approved methodology.

1. Current-year data products contain revisions to the prior year's estimates that are caused by incorporating:
   a. Improved methodology.
   b. New data inputs.
   c. Revisions to prior year data inputs.

2. The term "vintage" is used to refer to the reference date of an estimates cycle. Estimates released with a reference date of July 2005 are referred to as the "vintage 2005" set of population estimates and will include a consistent time series back to April 2000.

C. Within any vintage, all products use the same vintage of input data and must sum to the earlier released products of the same vintage for the same measurement.

1. Since the national and state population totals are the first to be released, all subsequent estimate products must sum to the national and state totals

that already appear for that vintage. This insures consistency within any vintage and means that the sum of the "parts" will always equal the previously released U.S., state, or county total.

2. Since the national population estimates tabulated by characteristics are the first characteristics to be released, the sum of the state and county characteristics must equal the national characteristics of the same vintage.

D. Only one consistent set of products and related materials is developed within a vintage. That set of products is intended to serve all customers' needs and uses.

1. The methodology and data inputs used to develop the population estimates used as denominators for vital statistics rates are consistent with those used to develop the population controls for the CPS and ACS.

2. Custom data products are consistent with the publicly released data products. For example, the annual race estimates for counties use a bridged race algorithm developed by NCHS. However, while the race data conform to the bridging algorithms developed by NCHS, the estimates of total populations and populations by age and sex generally agree with the publicly released data products.

E. The population estimates begin with the most recent decennial-census enumerated count updated to July 1 of each year, and as such, are based on the usual-residence concept used in the most recent decennial census.

1. The population estimates base for each estimate date is updated to include Count Question Resolution (CQR) changes to the decennial census base as well as geographic updates due to annexation and other geographic program changes.

2. The components of population change used to update the most recent census will be consistent with the best set of components available. Ongoing evaluation indicates that the coverage and the consistency of vital statistics and other administrative records data differ from those of decennial census data. Therefore, in the annual estimates, the size of the population based mainly on administrative records data differ from the size based mainly on census data.

F. States, counties, and units of local government have the right to challenge the population estimates prepared by the Census Bureau under the provisions of Title 15, The Code of Federal Regulations, Part 90. The results of accepted challenges will be incorporated into the following year's population estimates as long as the challenge is received by October 1 of the year in which the estimate was released.

III. Current Broad Methodological Assumptions

A. Prior to incorporating a new methodology or data set, it is desirable to thoroughly evaluate a set of estimates that use this new methodology or data set and compare it with the most recent decennial census results. When this is not possible, the methods are judged by the following criteria.

1. Soundness: The method should be based on solid reasoning – i.e., the formulas that embody the method should be mathematically valid and respect the attributes of the input data as they relate to the estimation task.

2. Integrity: A strategy that consistently applies the declared method is preferred to one that uses ad-hoc fixes to address particular challenges of the estimation task.

3. Parsimony: A simpler strategy is preferred to a more complex one.

4. Robustness: The method that produces the most reasonable estimates (defined below) across the full range of potential input-data values and in the presence of the random variation normally associated with those values while maintaining the orthodoxy and consistency of the estimates (also defined below) is preferred.

5. Adaptability: A technique that can be applied more broadly (e.g., across geographic summary levels), thus promoting the integration of the Census Bureau's estimates system, is preferred to a more product-specific remedy.

6. Transparency: A strategy that is more readily understandable and replicable by external parties is preferred. Moreover, a strategy that provides some explanatory information (i.e., how did the size or distribution of the population come to be this way) is preferred over one that is merely predictive.

7. Usability: The method must be executable along with all other current projects under current staffing levels in a way that allows the Census Bureau to meet current deadlines.

8. Flexibility: The preferred method will allow the production of estimates when a specific instance of the input data normally required by the method is unavailable or deemed unsuitable.

B. As a final test, the method should produce output data that have the following qualities.

1. Orthodoxy: The values of the population estimates should be appropriate (e.g., no negative population numbers, all population estimates in whole numbers).

2. Consistency: The values of the population estimates for all universes (e.g., resident, civilian, civilian non-institutionalized), geographies (e.g., national, state, county), and characteristics (e.g., age, sex, race, Hispanic origin) should not contradict one another.

3. Reasonableness: The values of the population estimates should approximate the real values as determined by the following assessments.
   a. Post-Censal Change: The reasonableness of the total change in the population since the last decennial census.
   b. Time-Series Change: The reasonableness of the annual change in the estimates since the last census.
   c. Demographic Appropriateness: The values of the estimates and the demographic rates they imply fall within acceptable limits when evaluated by general demographic principles (e.g., the appropriateness of the sex ratios, age progression, implied family size, life expectancies, total fertility rates, etc.).
   d. Comparability: The estimates appear realistic when compared with other indicators of the size and distribution of the population (e.g., Medicare enrollment, school enrollment, housing unit estimates, etc.).

C. A consistent method is used for entities at the same level of geographic aggregation.

1. The method adopted for state totals must be used for all states.

2. The method adopted for counties within a state must be used for all counties within that state.

D. The Census Bureau develops the basic estimates for the nation, states, and counties by disaggregated race groups in order to meet the various custom race aggregations needed by users.

E. The cohort-component method is the preferred method for development of the national, state, and county-level total population estimates and population estimates by characteristics.

F. The distributive housing-unit method is the preferred method for the development of the functioning subcounty governmental-unit-level estimates.

G. State total population estimates are not developed independently. National population estimates are first developed; then county total population estimates are developed and controlled to the national total population estimates. The state total population estimates are the sum of the “nationally controlled” county total population estimates for the state.

H. Data on vital statistics and group quarters provided by members of the Federal State Cooperative Program for Population Estimates (FSCPE) are included in the process of developing state and county population estimates.

I. Although state members of the FSCPE are provided the opportunity to review the state and county population totals prior to final production, they must follow strict criteria and provide objective evidence when requesting modifications.

IV. Current Specified Methodologies

A. National level estimates will use the cohort-component technique applied to data from the latest decennial census as the base, data on births and deaths provided by the National Center for Health Statistics, and estimates of net international migration derived from data from the American Community Survey (ACS) See the url

<http://www.census.gov/popest/topics/methodology/2003_nat_char_meth.html http://www.census.gov/popest/topics/methodology/v2005_nat_char_meth.html>

For a detailed discussion of the methodology used to develop the most recent set of national population estimates by demographic characteristics.

B. State and county population estimates are developed using a demographic procedure called an "administrative records component of population change" method. A major assumption underlying this approach is that the components of population change are closely tracked by administrative data in a demographic change model. In order to apply the model, Census Bureau demographers estimate each component of population change separately. For the population residing in households, the components of population change are births, deaths, and net migration, including net international migration. For the non-household population, change is represented by the net change in the population living in group-quarters facilities.

Each component in our model represents data that are symptomatic of an aspect of population change. For example, birth certificates indicate additions to the population resulting from births, so these data are used to estimate the birth component for a county. Other components are derived from death certificates, Internal Revenue Service data (IRS), Medicare enrollment records, Armed Forces data, group-quarters population data, and data from the American Community Survey.

For a more detailed discussion of the development of county population totals see
<http://www.census.gov/popest/topics/methodology/2003_st_co_meth.html
http://www.census.gov/popest/topics/methodology/2005_st_co_meth.html>

C. State population characteristics are currently developed in a two-stage process. Estimates by age and sex are developed first using a cohort-component procedure whereby estimates of net migration are developed using school enrollment data. These estimates are controlled both to the national-level estimates by age and sex as well as the previously developed state population totals.

The second step in the process distributes the state age and sex estimates into race by Hispanic origin categories. This is done by preparing an initial set of state estimates by age, sex, race, and Hispanic origin that are controlled to the state age and sex estimates prepared in the first step and to the previously developed national estimates by age, sex, race, and Hispanic origin.

For a more detailed discussion of the development of the state population characteristics by age, sex, race, and Hispanic origin see

<http://www.census.gov/popest/topics/methodology/2003_st_char_meth.html
http://www.census.gov/popest/topics/methodology/2004_st_char_meth.html>

D. County population characteristics are developed using a proportional distribution method beginning with previously developed resident county population estimates by age (0-64 and 65+) and resident state population estimates by age, sex, race, and Hispanic origin. Then county-level estimates of age, sex, race, and Hispanic origin distributions are developed using information about post-censal change in the corresponding populations. Third, these distributions are applied to the original county estimates by age and state characteristics.

A detailed discussion of this method is provided at

<http://www.census.gov/popest/topics/methodology/2004_co_char_meth.htm>

V. Enhancement Priorities

A. Improve estimates of net international migration

1. Provide up-to-date, useful statistics and methodologies on the size, characteristics, and demographic impact of international migration to and from the United States for use in policy-making decisions and demographic and economic research.

2. Goals of immigration research
   a. Produce annual estimates of international migration
   b. Improve current migration-related survey questions on the ACS.
   c. Conduct extensive evaluations to determine the best method to incorporate ACS data into the population estimates.

3. Activities

   a. Evaluate reasonableness of estimates of annual change in the foreign-born data from ACS at the national level.

   b. Produce revised estimates of net international migration at the national level.

   c. Produce new demographic and geographic distributions for migrants.

d. Construct algorithms to estimate the migrant status of the foreign-born populations.
e. Produce estimates of international migrants by migrant status (legal migrants, temporary migrants, quasi-legal migrants, unauthorized migrants, and emigrants).

E. Improve Estimates of Internal Migration

1. Improve the accuracy of the annual migration estimates by age, sex, race, and Hispanic origin for counties by maximizing the efficient use of available administrative data files, Census 2000 data, and the American Community Survey (ACS) data.

2. The ultimate goal is to implement a person-based migration model incorporating administrative data from files such as the IRS 1040 and 1099 records, Medicare records, a derived person-characteristic file developed from the Social Security Administrative NUMIDENT file, and other administrative data that can be merged into the database. The database will enable analysts to match administrative data with Census 2000 (100% and sample data), CPS, and ACS data in order to develop models that correct possible demographic and geographic biases inherent in the use of an administrative records database when estimating migration rates for counties.

F. Develop a new methodology for estimating subnational population characteristics

1. Replace the methodology that develops state estimates by age and sex based on school enrollment data with a method that is consistent with the best set of administrative data available and exploits the power of current computing capacity.

2. Develop a method that addresses current deficiencies in the age distributions of the population in selected states and counties, especially the age distribution of the population aged 18 to 24.

3. Develop a new method to estimate county population by age, sex, race, and Hispanic origin.

G. Develop procedures to systematically incorporate participation by State FSCPE Agencies in the production of state and county population estimates

1. Address issues of consistency

2. Establish criteria for incorporating state participation

VI. Other Enhancements

A. Improve the distributive housing unit approach at the subcounty level.
   1. Develop procedures to update Census 2000 measures of vacancy and numbers of people per household (PPH or the Person Per Household measure) used in the estimates process.

   2. Improve estimates of housing units.

   3. Address inconsistencies between estimates developed using the distributive housing unit approach and those developed using the component approach.
      a. Develop improved procedures to estimate housing unit loss.
      b. Integrate enhancements from the Master Address File.

B. Address inconsistencies between data from the decennial census base and data on components of change from administrative records databases.

   1. Address inconsistencies between Census 2000 data and NCHS data on race and Hispanic-origin characteristics.

   2. Address unreasonable results from pairing NCHS mortality data with decennial census data and estimate results.

VII. Administrative Constraints

A. The methods developed must be capable of being implemented with current resources and within the current time frame for estimate production.

B. Production of the complete set of estimates must continue during any development stages.

C. Methods must be Transparent and Reproducible

Appendix B. Annotated Bibliography and Glossary*

This annotated bibliography is organized into five sections. The first two sections deal with population estimation as a whole and the housing unit method, respectively. The following three sections deal with each of the three elements of the Housing Unit Method needed to generate a population estimate: (1) number of housing units; (2) vacancy rates; and (3) persons per household

Citations for sample-based methods are included in the first section, such as "capture-recapture" techniques. Also included in the first section are citations for population forecasting methods that also can be used to generate current estimates and some cases historical estimates (e.g., via backcasting) as well as some citations to the use of administrative records and techniques stemming from formal demography (e.g., quasi-stable population theory) that also can be used to generate current and historical estimates.

The glossary is focused on population estimation. By necessity, it covers demographic terms that are not exclusively found in the field of population estimation.

Table of Contents

Several entries appear in more than one section because they include work on more than one topic.

I. Population Estimation

Akkerman, A. 2000. "On the Leontief Structure of Household Populations." *Canadian Studies in Population* 27(1): 181-194. The author considers the age distribution of all persons in households and the age distribution of household-heads and shows that formal relationship holding between the two age distributions is equivalent to the input-output relationship in the Leontief model of the open economy. The notions of household composition and household accommodation that have emerged independently over the past two decades are shown to be formally linked within this relationship.

Alcantara, A. 1999. "Assessment of IRS Tax returns Migration Coverage in New Mexico." Paper presented the Estimates Methods Conference, U. S. Bureau of the Census, Suitland, MD. (available online, http://www.census.gov/population/www/coop/popconf/paper.html, last accessed June 2007). Using the housing unit method, the author evaluates how well the IRS tax return method estimates migration in the three counties in Metropolitan Albuquerque. She uses data on building permits, manufactured and mobile homes as well as occupancy rates by type of unit will be collected for the period between 1996 and 1997 for this test. Results of the estimates based on the housing unit method are compared with the migration flows implied by the IRS data for the same time period.

Alho, J. 1994. "Analysis of Sample-based Capture-Recapture Experiments." *Journal of Official Statistics* 10: 245-256. Heterogeneous capture probabilities and logistic regression are used to model the estimation of population totals in a sample based capture –recapture experiment. Some aspects of experiments that combine a census with a sample (such as the U.S. Post Enumeration Survey) are studied. The role of sampling weights in the estimation of capture probabilities is considered and the author shows how purely sample based estimators can be combined with the census estimates to reduce variance. Logistic catch effort models are applied to the optimal allocation of resources to the sampling part and the capture-recapture part of the experiments. The author shows with an analytical example that there can be a genuine trade-off between the two sources of error.

Alho, J. 1990. "Logistic Regression in Capture-Recapture Models." *Biometrics* 46: 623-635. The effect of population heterogeneity in capture-recapture, or dual registration, models is discussed. An estimator of the unknown population size based on a logistic regression model is introduced, which allows different capture probabilities across individuals and across capture times. The resulting population estimator is shown to be consistent and asymptotically normal. A variance estimator under population heterogeneity is derived. The finite-sample properties of the estimators are studied via simulation and an application is presented.

Alvey, W., and F. Scheuren. 1982. "Background for an Administrative Record Census." pp. 137-146 in 1982 *Proceedings of the Social Statistics Section*. Alexandria, VA: American Statistical Association. The paper examines in a general way the Federal administrative record systems that might be employed in conducting a population census. The focus of the description given is on the extent to which these systems could be used together to obtain population census counts. Sections I through V describe the various microdata files, touching on coverage and content issues. Section VI discusses the basic methodology proposed, raising some of the major linkage issues to be considered. Finally, in Section VII, an agenda for researching the proposal is suggested.

Anderton, D., J. Conaty, and T. Pullum. 1983. "Population Estimates from Longitudinal Records in Otherwise Data-Deficient Settings." *Demography* 20: 273-284. The authors present and evaluate models which derive population parameters for the population subgroup underlying such longitudinal data; using the distribution of individual times until 1st recorded event within a measurement interval, population parameters are estimated which provide basic denominator data for analyzing event occurrence. The use of the models is demonstrated and evaluated through an application to genealogical records for a 19th century population. The models evaluated in this paper are not robust and appear sensitive to both measurement errors and model assumptions regarding trapping probabilities.

Atchley, R. 1968. "A Shortcut Method for Estimating the Population of Metropolitan Areas." *Journal of the American Institute of Planners* 34: 259-262. In this article a variation of the vital rates method, the Age-Color-Specific Death Rate (ACSDR) Method, is offered as a short-cut method for providing an immediate estimate of the current population of a metropolitan area. Comparisons show that estimates made by this method tend to deviate only slightly from those made by other more laborious methods. The limitations of the ACSDR Method are noted, and it is concluded that this method can provide a quick and reasonably accurate provisional estimate, but that it is not suitable as a general substitute for more intensive methods of estimating metropolitan populations

Batutis, M. 1993. "Evaluation of 1990 Population Estimates and the Future of the Census Bureau Sub-national Estimates Program." *Proceedings of the Social Statistics Section*, Alexandria, VA: American Statistical Association. First, this paper establishes a historical context for the 1990 evaluation program. In so doing, the paper attempts to show how the philosophy toward population estimates, as explicated in official publications, has shaped the current evaluation program. Second, the paper attempts to show, in broad outline, the changes in philosophy that are required to move the population estimates program toward the year 2000 and beyond.

Batutis, M., and R. Scardamalia. 1982. "Population Estimates for New York Counties by the Component Method II." pp. 133-136 in 1982 *Proceedings of the Social Statistics Section*. Alexandria, VA: American Statistical Association. The use of the component method II model for estimating U.S. population at state and county levels is evaluated using 1980 census data for New York State. The component method II is shown to fail because of the failure of the key assumption underlying the method, namely unvarying

Residual Migration Adjustment Factors RMAF). The usefulness of the method is in serious doubt unless a way can be found to update the RMAF as part of the estimating process.

Becker, P. 1999. "Using the Master Address File to Estimate the Population for Small Areas." Paper presented the Estimates Methods Conference, U. S. Bureau of the Census, Suitland, MD. (available online, http://www.census.gov/population/www/coop/popconf/paper.html, last accessed June 2007). The author starts by observing that the Census Bureau has moved from population-based methods to housing unit methods for purposes of estimating sub-county population because population-based methods have suffered from seriously inadequate data sources and high levels of geocoding error. Housing units, already located on the ground, provide a better data set with little geocoding error. However, the author points out that housing-based methods are flawed as well. The author examines the use of the Master Address File to provide housing required for the housing unit method of population estimation and discusses issues that need to be resolved before it can be used as a source of population estimates.

Belcher, D. 1921 "A Method of Estimating the Population of Continental United States for Inter-censal Years." *Journal of the American Statistical Association* (then known as *Quarterly Publications of the American Statistical Association*) 17: 994-998. The author shows how one can use interpolative methods to estimate the US population between decennial census counts.

Bell, W. 1993. "Using Information from Demographic Analysis in Post-Enumeration Survey Estimation." *Journal of the American Statistical Association* 88 (423):1106-1118. Population estimates from the 1990 Post-Enumeration Survey (PES), used to measure decennial census undercount, were obtained from dual system estimates (DSE's) that assumed independence within strata defined by age-race-sex-geography and other variables. The author considers several alternative DSE's, and used DA results for 1990 to apply them to data from the 1990 U. S. Census and PES.

Besbeas, P., Freeman, S. N., Morgan, B. J. T., Catchpole, E. A. 2002. "Integrating Mark-Recapture-Recovery and Census Data to Estimate Animal Abundance and Demographic Parameters." *Biometrics* 58:540-547. The authors show how a state-space model for census data in combination with the usual multinomial-based models for ring-recovery data provide estimates of productivity not available from either type of data alone.

Bogue, D. 1950. "A Technique for Making Extensive Population Estimates." *Journal of the American Statistical Association* 45: 191-193. This paper introduces the vital statistics method for making population estimates. It shows how it can be employed at the sub-national level.

Bogue, D., and B. Duncan. 1959. "A Composite Method for Estimating Post-censal Population of Small Areas by Age, Sex and Color." *Vital Statistics Special*

Reports, Vol. 47, No. 6. This paper introduces a method of making total population estimates by first using the vital statistics method to estimate the numbers in age (sex and race) groups and then adding the numbers to get the total population.

Bourgeois- Pichat, J. 1971. "Stable, Semi-stable Populations, and Growth Potential." *Population Studies* 25: 235-254. Starting from the definition of a Malthusian population given by Alfred J. Lotka, the author recalls how the concept of stable population is introduced in demography, first as a particular case of stable populations, and secondly as a limit of a demographic evolutionary process in which female age-specific fertility rates and age-specific mortality rates remain constant. Then he defines a new concept: the semi-stable population which is a population with a constant age distribution. He shows that such a population coincides at any point of time with the stable population corresponding to the mortality and the fertility at this point of time. In the remaining part of the paper it is shown how the concept of a stable population can be used for defining a coefficient of inertia which measures the resistance of a population to modification of its course as a consequence of changing fertility and mortality.

Bousfield, M.V. 1977. "Inter-censal Estimation Using a Current Sample and Census Data." *Review of Public Data Use* 5: 6-15. The author was among the first to describe the use of raking to force the marginal totals of a two-way sample table to match census totals. She shows how it can be used to generate population estimates.

Bouvier, L. 1971. "Estimating Post-Censal Populations of Counties." *Journal of the American Institute of Planners* 37 (1): 45-46. The author applies the vita rates method to estimate the population of Rhode Island and its counties. He demonstrates that this inter-censal technique can produce relatively accurate results for states and small units.

Bouvier, L., D. Poston, and N. Zhai. 1991. "Population Growth Impacts of Zero Net International Migration." *International Migration Review* 31 (2): 294-311. This article examines the assumption that the effect of zero net international migration on the population of the United States. Examining the direct, indirect, total, and negative demographic impacts of zero net international migration through simulations with demographic data, the authors demonstrate that zero net international migration is not the same and therefore does not have the same demographic results and implications as zero international migration. They conclude that zero net international migration should not be confused with zero international migration.

Bracken, I. 1991. "A Surface Model Approach to Small Area Population Estimation." *Town Planning Review* 62(2): 225-37. The author applies a surface model approach to the estimation of small-area population and household characteristics and argues that the representation of population-related information by means of surface concepts offers a way to overcome many of the limitations of traditional, 'fixed' zone-based methods. The author suggests that the method can be applied to the estimation of population at local levels and gives ideas on how to accomplish this.

Brass, W. 1968. *The Demography of Tropical Africa*. Princeton, NJ: Princeton University Press. Divided into two major parts (basically data/methods and case studies), this book contains descriptions of model life table methods that can be used to estimate populations. Of particular note is the inclusion of Brass's two parameter model life table system.

Brass, W. 1975. *Methods for Estimating Fertility and Mortality from Limited and Defective Data*. Chapel Hill, NC: The University of North Carolina. This monograph contains a series of papers on indirect estimation methods that were otherwise difficult to obtain. The papers deal with a simple approximation for the time-location of estimates of child mortality from proportions dead by age of mother, further simplification of time location estimates for survivorship of adult relatives reported at a survey, the derivation of life tables from retrospective estimates of child and adult mortality, the relation between numbers of living mothers and numbers of living children, P-F synthesis and parity progression ratios, mortality in China using data from the 1982 census, and childhood mortality estimated from reports on previous births given by mothers at the time of a maternity.

Brown, H. 1955. "'A Technique for Estimating the Population of Counties." *Journal of the American Statistical Association* 50: 323-343. The author shows how school enrollment data can be used together with census correction factors to develop county population estimates.

Brown, W., and R. Scardamalia. 1991. "Estimating County Population by Age, Sex and Race: A State Perspective." *1990 Working Paper Series, No. 2.03.* Ithaca, NY: Department of Rural Sociology, Population and Development Program, Cornell University. The authors review a methodology used by the U.S. Census Bureau to prepare national estimates of county population by age, sex, and race. They then evaluate its applicability to preparing state estimates of county population, using data for New York State.

Brunsman, H. 1955. *The Estimation of Population Changes for New York City*. New York, Russell Sage Foundation. This monograph covers the recommendations made to the Mayor by the Committee on Statistical Programs for the City of New York on how to estimate and project the population of New York City and its sub-areas. It lists 21 data sources and finds that 15 of them are currently available. It discusses how to obtain the additional data needed to carry out the recommendations.

Bryan, T. 2004. "Population Estimates." pp. 523-560 in J. Siegel and D. Swanson (eds.) *The Methods and Materials of Demography, 2nd Edition*. New York City, NY: Elsevier Academic Press. This chapter provides an overview of the methods used to make population estimates in countries with well developed vita registration, census, and other systems. It provides the formulas and the data needed for these methods and gives examples.

Burghardt, J., and V. Geraci. 1980. “State and Local Annual Population Estimation Methods Employed by the Bureau of the Census.” *Review of Public Data Use* 8 (4): 339-354. This provides an overview of the population estimation methods used by the US Census and evaluates various population estimates prepared in the 1960s and 1970s using examples from Texas.

Byerly, E. 1984. “Estimates of the Population of States: 1970 to 1983.” *Current Population Reports, Series P-25, No. 957*, pp 4-7. Washington, DC: US Government Printing Office. This report contains revised estimates of the resident and civilian population of states for July 1, 1981, and July 1, 1982, and provisional estimates for July 1, 1983, together with provisional components of population change for the period April 1, 1980 to July 1, 1983. Revised annual inter-censal estimates for July 1, 1971, to 1979 are also presented, as well as components of population change for 1970 to 1980.

Byerly, E. and K. Deardorff. 1995. “National and State Population Estimates: 1990 to 1994.” *Current Population Reports, P25, No. 1127*, US Bureau of the Census. Washington, DC: US Government Printing Office. This report begins with a brief section highlighting 1990 to 1994 population trends, which is followed by a section describing methodology. The report concludes with detailed tables of U.S., region, division, and State population estimates for 1990 to 1994.

Cai, Q. 2007. “New Techniques in Small Area Population Estimates by Demographic Characteristics.” *Population Research and Policy Review* 26 (2): 203-218. This paper explores methods for population estimates by age, sex, race, and Hispanic origin at the census tract level for Multnomah County, Oregon, including the use of building permits to indirectly estimate migration and examining the changes in age/sex structure using the American Community Survey.

Carmen, A., and J. Somoza. 1965. “Survey Methods, Based on Periodically Repeated Interviews, Aimed at Determining Demographic Rates.” *Demography* 2: 289–301. The authors present the results of a special survey designed to permit data to be obtained in the less developed countries, estimating natality and mortality, calculated within a level of acceptable confidence. The advantages and limitations of the method are discussed.

Carrier, N., and J. Hobcraft. 1971. *Demographic Estimation for Developing Societies*. London, UK: Population Investigation Committee. London School of Economics. This manual is designed to detect and reduce errors in demographic data. It also serves as a guide for using age distributions to estimate fertility and mortality. It has an extensive set tables generated by the three-parameter stable model, which is the main purpose of the book.

Cavanaugh, F. 1981. "The Census Bureau's 1980 Census Test of Population Estimates," in *Small-Area Population Estimates -Methods and Their Accuracy and New Metropolitan Area Definitions and Their Impact on the Private and Public Sector, Series GE-41, No. 7.*

Washington, DC, Government Printing Office. The author provides results of the Census Bureau's tests of the accuracy of the methods it used for population estimation during the 1970s.

Cerone, P. 1987. "On Stable Population Theory with Immigration." *Demography* 24 (3): 431-438. The paper extends stable population theory to include a constant stream of immigration. It is shown that under a constant stream of immigration, the population will asymptotically tend toward a constant, linear, or exponential behavior, depending on whether the fertility behavior is below, equal to, or above replacement level. All of the parameters are determined in terms of the characteristics of the population at the origin.

Chamberlain, A. 2006. *Demography in Archaeology*. Cambridge, England: Cambridge University Press. After describing features of demography and its use in archaeology, the author covers methods that can be used to develop estimates of historical populations and their characteristics. He concludes the book with a discussion of the relevance of demography to archaeology and his views on future challenges.

Chandra Sekar, C. and W. E. Deming. 1949. "On a Method of Estimating Birth and Death rates and the Extent of Registration." *Journal of the American Statistical Association* 44: 101-115. A mathematical theory is presented which, when applied to a comparison of the registrar's lists of births and deaths with a list obtained in a house-to-house canvass, gives an estimate of the total number of events over an area in a specified period. It also gives the extent of registration. In a test of the method done in India in 1947, it was found that the estimated total number of events for the area is usually greater when the estimate is built up by summing the totals for individual groups than when it is computed at once for the aggregated population. This observation, according to the theory, confirms positive dependence and indicates that the greater figure is closer to the truth.

Chattopadhyay, M., P. Lahiri, M. Larsen, and J. Reimnitz. 1999. "Composite Estimation of Drug Prevalences for Sub-State Areas." *Survey Methodology* 25: 81–86. A hierarchical model is proposed to address problems found with variances in both synthetic and design-based survey estimates of drug use. The authors propose Empirical Bayes composite estimators, which incorporate survey weights, of drug use prevalence and jackknife estimators of their mean squared errors, to overcome these issues and illustrate the use of these estimators.

Chaudhuri A. and T. Christofides. 2006. "Item Count Technique in Estimating the Proportion of People with a Sensitive Feature." Journal of Statistical Planning and Inference 137 (2): 589 -593. In assessing the prevalence of a sensitive attribute like habitual heroin consumption in a community of people, indirect questioning is a necessity to extract truth on ensuring protection of privacy. The current literature seems to need supplementary specification of a relevant practical and theoretical justification for one possibility by what is called an Item Count Technique. This method can be easily incorporated in large scale sample surveys where the medium of collecting information is a structured questionnaire. This feature will make this technique attractive to social survey researchers. In this article, the authors present an amendment to the currently available technique rendering it well-equipped with a provision to protect privacy and also a sound theoretical foundation.

Childers, D., and H. Hogan. 1984. "The IRS/Census Direct Match Study–Final Report." *SRD Research Report No. Census/SRD/RR-84/11*. Statistical Research Division, Bureau of the Census. Washington, DC: U. S. Department of Commerce. This paper reports on a study to investigate the feasibility of using the Internal Revenue Service Individual Master File (IRWIMFI as a frame for matching to the census in order to estimate gross under-coverage in the census, and to examine the difficulties in tracing individuals to the census using the IRS/IMF address. The authors conclude that the study has demonstrated that the problems of post office boxes, rural routes, and business addresses can be overcome with proper follow-up procedures and that the potential of this sampling frame is immense.

Chu, S. 1974. "On the use of Regression Method in Estimating Regional Population." *International Statistical Review* 42 (1): 17-28. The author examines the mathematical foundation of the ratio-correlation model and finds it "peculiar," with such drawbacks as the interdependence of shares and the fact that its coefficients are constant between census counts. He then proposes using past growth rates instead of shares in the model and extends this through stratification. He tests his idea using 1960 census data for the US stats (using models constructed from 1940 and 1950 data) and finds that ratio-correlation model under-estimates fast growing states and over-estimates slow growing states, but that the overall reduction in the MAPE is only from 6.87 percent to 6.79 percent for the states.

Clogg, C., J. Schockey, and S. Eliason. 1990. "A General Statistical Framework for the Adjustment of Rates." *Sociological Methods and Research* 19 (2): 156-195. The authors present a general framework integrates standardization procedures common in demography, biometrics, and other areas with statistical methodology for the analysis of log-linear models. A family of rate-adjustment methods is derived from the log-linear model; the conventional method of direct standardization is a special case. Extensions of earlier methods include (a) adjustment for three-factor interaction, (b) adjustment for marginal association between composition and group, (c) adjustments that use a standard group, and (d) adjustments that control for both marginal composition-group interaction and three-factor interaction. Statistical inference for adjusted rates is facilitated in several ways: (a) by presenting key hypotheses that can be tested routinely with log-linear methods, (b) by efficient point and interval estimation of rates, (c) by assessing the sampling variability of absolute or relative comparisons of rates across groups, and (d) by smoothing the data. Examples illustrate the flexibility of the proposed framework.

Coale, A. 1984. "Construction of a life table from accurate enumeration of a closed population in two censuses." *Population Index* 50 (2). This paper outlines how census survival ratios can be used with assumptions to develop a life table.

Coale, A. 1965. "Estimates of Average Size of Household." pp. 64-69 in A. Coale, L. Fallers, M. Levy, D. Schneider, and S. Tompkins (Eds.) *Aspects of the Analysis of Family Structure*. Princeton, NJ: Princeton University Press. The author provides a demographic perspective on the determinants of average household size.

Coale, A. 1971. *The Growth and Structure of Human Populations: A Mathematical Investigation.* Princeton, NJ: Princeton University Press. This book considers the relationship between levels, age patterns, and time patterns of fertility and mortality and the growth and age composition of populations. Instead of simply providing demographers with the mechanical implements to calculate age distribution, birth and death rates, and rates of increase, the author attempts to explain how age structures are formed and vital rates determined.

Coale A., and P. Demeny. 1966. *Regional Model Life Tables and Stable Populations* Princeton, N.J.: Princeton University Press. This first edition provides model tables for which the terminal open-ended age group is 80 years and life expectancies that ended at age 77.5. This works well for populations with high mortality, but not so well for populations with low mortality. The tables presented in this book are in two principal forms: model life tables and model stable populations.

Coale, A., and N.W. Rives. 1973. "A Statistical Reconstruction of the Black Population of the United States 1880-1970: Estimates of True Numbers by Age and Sex, Birth Rates, and Total Fertility." *Population Index* 39 (1): 3-36. This paper describes new procedures that the authors have used to reconstruct the black population, distributed by age and sex, from 1880 to 1970. The authors take advantage of minimal international migration for this population and use the mechanics of the age structure and growth of a closed population to generate the estimates.

Coale, A., and J. Trussell. 1996. "The development and use of demographic models." *Population Studies* 50: 469-484. The authors examine two classical demographic models--conventional life tables and stable populations--and a modern generalization of stable population theory and discuss mathematical models of conception and birth. The authors examine the use of demographic models in forecasting future mortality, nuptiality, and fertility and in population projection. They conclude with observations about the purposes and uses of demographic models

Coale, A., P. Demeny, and B. Vaughan. 1983. *Regional Model Life Tables and Stable Populations*. Princeton, NJ: Princeton University Press. This second edition incorporates an extension in the range of the original tables that will be particularly useful in applications to populations with low mortality. In the second edition, "the life tables and stable populations are tabulated by five-year age intervals to age 100, in all but the highest mortality levels, in which the last five-year tabulation is from 90 to 95. The greatest expectation of life at birth is now 80 years (for females) rather than 77.5." As was the case in the first edition (1966), the tables are presented in two principal forms: model life tables and model stable populations

Cohen, M. 1991. “Statistical Matching and Microsimulation Models.” pp. 62-85 in C. F. Citro and E. Hanushek (Eds.) *Improving Information for Social Policy Decisions: The Uses of Microsimulation Modeling, Volume 1I: Technical Papers* Washington, DC: National Academy Press. The author discusses the strengths and weaknesses of ‘distance’ functions as a means of record-matching, as well as those of a matching method based on variance-covariance analysis.

Cohen, S. 1980. “A Comparative Study of Synthetic Estimation Strategies with Applications to Data from the National Health Expenditure Study.” *Proceedings of the American Statistical Association, Survey Research Methods Section,* pp 595-600. Alexandria, VA: American Statistical Association. The author examines alternative strategies for obtaining estimates for small areas by exploiting sample survey and other data. The methods include synthetic estimation, regression estimation, and composite estimation. He finds that reliable estimates for small areas are difficult to obtain, if not impossible and that the selection of a given strategy is a function of the symptomatic information available, the appropriateness of a specified prediction model, and the number of available observational units. He suggests that the use of a composite estimator with judiciously chosen weights will avoid the chance of selecting the least precise technique and will potentially yield an estimate of superior precison.

Coleman, C., and D. Swanson.2007 “On MAPE-R as a Measure of Cross-Sectional Estimation and Forecast Accuracy.’ *Journal of Economic and Social Measurement* 32 (4): 219-233. The authors show that MAPE-R can be calculated simply, thus overcoming the cumbersome calculation procedure used in its introduction and noted as a feature needing correction. They find this closed form expression for MAPE-R to be a member of the family of power mean-based accuracy measures. While further lines of research are called for, nothing in their examination of MAPE-R rules out its use.

Congdon, P. 1989. “An Analysis of Population and Social Change in London Wards in the 1980s.” *Transactions of the Institute of British Geographers* N.S. 14: 478-491. This paper discusses the estimation and projection of small area populations in London, and considers trends in inter-censal social and demographic indices which can be calculated using these estimates. Trends in spatial inequality of such indicators during the 1980s are analyzed and point to continuing wide differentials. A typology of population and social indicators gives an indication of the small area distribution of the recent population turnaround in inner London, and of its association with other social processes such as gentrification and ethnic concentration.

Cook, T. 1998. “Overnight Visitor Counts in Australia and Their Implications for Population Estimation.” *People and Place* 6 (1):60-70. The author points out that on the night of the 1996 Census, 5.4 per cent of the people counted in Australia were staying away from home (visitors), compared to 4.6 per cent in 1986. She goes on to note variations by state and comments on the measurement of overseas visitors.

Cook, T. 1996. “When ERPs Aren’t Enough: A Discussion of Issues Associated with Service Population Estimation.” *Working Paper 96/4*. Demography Section, Australian

Bureau of Statistics, Belconnen, ACT, Australia. The author discusses the shortcomings of the Estimates of Resident Population (ERP) for many planning purposes and infrastructure needs in Australia. She notes that while population estimates based on place of 'usual residence' are conceptually sound and are favored over 'place of enumeration' estimates by many international statistical agencies, the relevance of 'usual residence' based estimates to some users is limited by the level of population mobility hidden within these estimates. She covers the issues involved in estimating ‘service (de facto) populations for the Australian Bureau of Statistics.

Courgeau, D, V. Nedellec, and P. Empereur-Bissonnet. 2000. “Duration of Residence in the Same Dwelling. A Test of Measurement Using Electricity Utility Company Records.” *Population* 12: 335-342. The authors report on the use of accounts records maintained by France’s national electrical provider, EDF, to estimate moving and migration.

Cowan, C., and D. Malec. 1986. “Capture-recapture Models when Both Sources have Clustered Observations.” *Journal of the American Statistical Association* 81: 347-353. Capture-recapture models assume that individuals in the population are captured one at a time and independently of each other. There are often situations, however, where individuals are captured in small clusters or groups. This paper provides a model that allows individuals to be captured in groups; the EM algorithm is used to estimate parameters in the model that include capture probabilities and the size of the population under study.

Cox, P. 1970. *Demography, 4th Edition*. Cambridge, UK: Cambridge University Press. A text and reference book that covers major theories and methods used in demography. It includes a chapter on population projections.

Creech, J. W., and D. K. Sater. 1999. “Assessment of the Coverage of the Population by Exemptions Represented on Individual Income Tax Returns for Tax Year 1989.” Unpublished Paper. Population Division, Bureau of the Census. Washington, DC: U.S. Census Bureau. The authors find that about 85 to 90 percent of the population is covered by tax returns, but with significant geographic variation. In one state they find only 80 to 85 percent coverage and in many small counties, the coverage is under 70 percent.

Crosetti, A., and R. Schmitt. 1956. "A Method of Estimating the Inter-censal Population of Counties." *Journal of the American Statistical Association* 51: 587-590. This paper describes and applies the ratio-correlation method to the estimation of county populations in Washington State for the year 1940. It finds that the method is more accurate than four other methods.

Darroch, J. 1961. “The Two-Sample Capture-Recapture Census When Tagging and Sampling Are Stratified.” *Biometrika* 48, 241-260. The author starts by recalling the capture-recapture argument used for the simplest type of experiment with only two samples and negligible death and emigration rates and then proceeds to argue since that the probability of "capture" is not uniform over a region, stratification could be employed in both the first and second samples to improve estimation methods when the probability of being captured in the second sample does not hold.

Darroch, J. 1959. "The Multiple-recapture Census II. Estimation When There is Immigration or Death." *Biometrika* 46: 336-351. This paper treats the multiple-recapture census for which the population is not closed. The aims of this paper are to provide exact, fully stochastic models for the observed frequencies of individuals, to show how simply these frequencies naturally group themselves, and to obtain estimates of the unknown parameters. When there is immigration only or death only, the estimates are shown to be asymptotically efficient and their variances are found. In addition, a method of performing tests on the values of the parameters is given. When both immigration and death are operating, on the other hand, the complexity of the probability density prevents us from going further than obtaining the estimates and merely indicating how their variances can be found.

Darroch, J. 1958. "The Multiple-recapture Census I: Estimation of a Closed Population." *Biometrika* 45: 343-359. This paper treats the multiple-recapture census for which the population is closed both to augmentation from outside and departure from inside and the number of samples is fixed.

Davis, S. 1994. "Evaluation of Post-censal County Estimates for the 1980s." *Population Division Working Paper No. 5*. Washington, DC: U. S. Census Bureau. This paper examines various issues concerning the evaluation of the official population estimates prepared for U.S. counties. It groups counties by size and growth and identifies which types of counties had better estimates.

Davis, S. 1988. "Methodology for Experimental County Population Estimates for the 1980s." *Current Population Reports, P-23, No. 158*. Washington, DC: U. S. Government Printing Office. This report briefly describes the procedures used in developing county population estimates in the 1980s by the Census Bureau, and indicates changes to the methodology previously used. The methodology used is a cohort-component projection technique in which actual data for the components are substituted for projected values whenever possible and totals are controlled to available estimates by age, sex, race, and geographic area. The Census Bureau considers these estimates experimental because they had not been tested fully against a census.

D'Allesandro, F. and J. Tayman. 1980. "Ridge Regression for Population Estimation: Some Insights and Clarifications." *Staff Document No. 56*. Olympia, Washington, Office of Financial Management. An explication of ridge regression in conjunction with the ratio-correlation method with examples for Washington State counties.

DeGuilbert-Lantoine, C. 1987. "Estimating Intercensal Populations in French Departments." *Population* 42 (6): 881-909. The author explains that the population in each French département has been estimated in each intercensal year from exact information on natural population movement, together with estimates of net migration, which were based on migration trends during the past. She notes that the quality of the results varied: a comparison between the results obtained and the populations enumerated in the Census of 1982 showed significant discrepancies for certain départements, particularly the Ile-de-France. Having considered the accuracy of the French estimates, estimation methods used in the United States, she studies, and an attempt is made to evaluate the extent to which the French estimates could be improved by applying some of the American methods. The method she selected consisted of estimating total migration from school enrolment data. She assumed that the relation between school enrolment and total migration observed during the previous intercensal period remained unchanged in each département, and net migration is then estimated. The author found that pplying this method to all French départements for the period 1975-82 tended to produce more accurate results than the old method. The method can, she argues, be considered as a useful complement to the estimation methods used by INSEE.

Deming, W.E., and N. Keyfitz. 1967. "Theory of Surveys to Estimate Total Population," pp. 141-144 in *Proceedings of the World Population Conference, Belgrade, 1965, Vol. 3.* New York City, NY: United Nations. This paper discusses some of the statistical problems encountered in estimating by sampling the total number of a population, without benefit of a previous census, and presents a device for this purpose which may have other uses as well. The authors consider two kinds of situations: (a) the population is fixed, each person being nominally attached in some recognizable manner to a fixed location, such as a dwelling unit; (b) the population is mobile – here today, somewhere else tomorrow. Some theory for the moving population is introduced.

Diffendal, G., et al. 1987. "Comparison of Adjustment Methods for Census Undercount in Small Areas." pp. 56-65 in *Proceedings of the Social Statistics Section*. Alexandria, VA: American Statistical Association. The authors compare synthetic and regression techniques for adjusting small area data. Examples are provided using data from California, Mississippi, New Jersey, and North Dakota.

Diffendal, G., S. Ogunwole, and A. S. Smith. 1999. "Applying Data from the American Community Survey (ACS) and the Master Address File (MAF) to the Inter-censal Population Estimates Program." Paper presented at the Estimates Methods Conference, U. S. Bureau of the Census, Suitland, MD. (available online, http://www.census.gov/population/www/coop/popconf/paper.html, last accessed June 2007). This paper addresses some of the significant issues involved in integrating the Inter-censal Population Estimates program and the American Community Survey. It uses two cases studies will detail research underway since the introduction of the ACS with four communities in 1996. In the first case study, the ACS results are compared to county population estimates by age, sex, race and Hispanic origin. In the second, ACS results are with sub-county population estimates. Both case studies conclude with a discussion of how the Inter-censal Population Estimates program and the American Community Survey can inform one another in the future.

Dundon, T. 1961. "A Method for Estimating the Population of Missouri Counties." *Business and Economic Review* 2 (1): 10-15. (no information available).

Eberle, G. 1938. "Population Estimates of Local Communities and Economic Planning." *Journal of the American Statistical Association* 33: 701-703. The author classifies the population by permanent and non-permanent residence and provides six governing rules for making estimates and determining population changes in urban communities for these types of population.

Eldridge, H. 1947. "Problems and Methods of Estimating Post-censal Population." *Social Forces* 24: 41-46. The author identifies three types of estimates that need to be made for the U.S., De jure, De facto, and troops overseas. She then discusses the methods and data – including ration books – that could be used to make these estimates.

Ericksen, E. 1973. "A Method for Combining Sample Survey Data and Symptomatic Indicators to obtain Population Estimates for Local Areas." *Demography* 10: 137-160. In this paper, a new method of post-censal estimation is presented in which the symptomatic information is combined with sample data by means of a regression format. Combining symptomatic information on births, deaths, and school enrollment with sample data from the Current Population Survey, county estimates of population growth were computed by means of the new method for the post-censal period after 1960. These estimates were tested for accuracy by comparison with a set of special censuses which were conducted between 1964 and 1967 in 75 counties. The results of this test are promising, but not conclusive.

Erickson, E. 1974. "A Regression Method for Estimating Population Changes of Local Areas." *Journal of the American Statistical Association* 69: 867-875. A regression method is presented in which current sample data and symptomatic information are combined to estimate post-censal populations for local areas. This procedure was tested for counties and states using 1970 Census data, and the resulting estimates were found to be more accurate than estimates computed by standard demographic procedures for the same period. The ratio-correlation estimates were the most accurate series of standard estimates. When this series was added to the set of symptomatic information used in the regression method, further increases in accuracy were obtained.

Ericksen, E., and J. Kadune. 1985. "Estimating the Population in a Census Year: 1980 and Beyond." *Journal of the American Statistical Association* 80 (389): 98-109. The authors argue that decennial census results should not be viewed as counts to be reported directly, but as data to be used in estimating the population and its characteristics. They propose methods by which the results of the 1980 census could be so analyzed using both other nationally collected information currently available at the Census Bureau and locally collected information especially likely to be needed in areas where undercount rates are high.

Espenshade, T. and J. Tayman. 1982. "Confidence Intervals for Post-censal State Population Estimates." *Demography* 19 (2): 191-210. This paper develops a methodology for constructing confidence intervals around post-censal state population estimates. Using regression equations, forecast intervals are derived around the average age-specific death rates over the post-censal estimation period. These results, combined with the number of post-censal deaths and the most current census counts, are translated into confidence intervals for the age structure. 2 approaches are offered for constructing total population confidence intervals. 1 examines a simulated distribution while the other focuses on the mathematical derivation of population means and variances. The methodology is illustrated by deriving statistically desirable confidence intervals around the July 1, 1975 population of Florida.

Espenshade, T., F. Hobbs, and L. Pol. 1981. "An Experiment in Estimating Post-censal Age Distributions of State Populations from Death Registration Data." *Review of Public Data Use* 9: 97-114. This article sets out a method of disaggregating post-censal estimates of total state population into five-year age categories. The central assumption is that the age pattern of temporal mortality change in a particular state is the same as that in the U.S. as a whole. Data on the number of deaths in each age interval, combined with assumptions about death rates, then indicate the size of the corresponding base population. The method is applied to the State of Florida and comparisons with the 1970 Census for Florida reveal good agreement, especially for ages 15 and older.

Espenshade, T., L. Bouvier, and W. B. Arthur. 1982. "Immigration and the stable population model." *Demography* 19 (1): 125-133. The authors extend stable population theory to include immigration. The central finding is that, as long as fertility is below replacement, a constant number and age distribution of immigrants (with fixed fertility and mortality schedules) leads to a stationary population. Neither the level of the net reproduction rate nor the size of the annual immigration affects this conclusion; a stationary population eventually emerges.

Ewbank, D. 1981. *Age Misreporting and Age-Selective Underenumeration: Sources, Patterns, and Consequences for Demographic Analysis*. Washington, DC: National Academy Press. This is a comprehensive review of the literature dealing with age misreporting and age-selective under-enumeration in demographic data collection. It describes the errors and how they can be identified and handled.

Fay, R. 1987. "Application of Multivariate Regression to Small Domain Estimation." pp. 91-102 in R. Platek, J.N.K. Rao, C.E. Sarndal, and M.P. Singh (eds.) *Small Area Statistics*. New York: Wiley. The estimates of the small area parameters can profitably 'borrow strength' from data on related multiple characteristics and/or auxiliary variables from other neighboring areas through appropriate models. The author proposes doing this by using multivariate regression for small area estimation of multiple characteristics.

Fay, R. 1974. *Statistical Considerations in Estimating the Current Population of the United States*. Ph.D. Dissertation, Department of Statistics. Chicago, IL: University of Chicago. This dissertation assesses Demographic Analysis as technique for estimating the U.S. population.

Fay, R., and R. Herriot. 1979. "Estimates of Income for Small Places: An Application of James Stein Procedures to Census Data." *Journal of the American Statistical Association* 74: 269-277. An adaptation of the James-Stein estimator is applied to sample estimates of income for small places (i.e., population less than 1,000) from the 1970 Census of Population and Housing. The adaptation incorporates linear regression in the context of unequal variances. Evidence is presented that the resulting estimates have smaller average error than either the sample estimates or an alternate procedure of using county averages.

Feeney, G. 1990. "Untilting Age Distributions: A Transformation for Graphical Analysis." *Asian and Pacific Population Forum* 4 (3): 13-20. This article presents a new approach to the plotting of age distribution data. 'Untilting' is a way of transforming data that vary systematically from very high to very low values so as to show local variation more clearly. The article derives an untilting transformation from the formal structure of age distributions. The transformation turns out to be closely related to two familiar demographic techniques, reverse-survival estimation of births and birth rates, and comparison of observed with stable age distributions. The ideas are illustrated by application to age distributions from the 1979 and 1989 censuses of Vietnam.

Feeney, G. 1970. "Stable Age by Region Distributions." *Demography* 7(3): 341-348. This paper is aimed at exploring the nature of stable compositions, the interrelation among their various components, and their relation to fertility, mortality, and mobility patterns that generate them. If the pattern of fertility, mortality and interregional migration exhibited by the United States population during 1950-60 were to continue in the future, the proportions of persons in the various age groups and regions would fluctuate from decade to decade. These fluctuations would become less marked with time, however, and eventually all the proportions would stabilize at certain fixed values. This collection of values may be called a stable age by region composition corresponding to the given schedule of fertility, mortality and migration. The same phenomenon may be observed when individuals move between socioeconomic categories as, for example, socioeconomic status or educational attainment levels. The substantial differences between these various situations conceal remarkable similarities. In each case the continued operation of schedules of fertility, mortality and mobility between categories may result in a stable composition.

Feeney, D., J. Hibbs., and R. Gillaspy. 1995. "Ratio-correlation Method." pp. 118-136 in N. W. Rives, W. Serow, A. Lee, H. Goldsmith, and P. Voss (Eds). *Basic Methods for Preparing Small-Area Population Estimate*s. Madison, WI: Applied Population Laboratory, University of Wisconsin. The paper provides an overview of the ratio-correlation method and gives examples of its use in Minnesota.

Fellegi, I. and A. Sunter. 1969. "A Theory for Record Linkage." *Journal of the American Statistical Association* 64: 1183-1210. The authors develop a mathematical model to provide a theoretical framework for a computer-oriented solution to the problem of recognizing those records in two files which represent identical persons, objects or events (said to be matched). A theorem describing the construction and properties an optimal linkage rule and two corollaries to the theorem which make it a practical working tool are given.

Felton, G. 1986. "Evaluation of Population Estimation Procedures for Counties: 1980." *Current Population Reports P-25, No. 984*. Washington, DC: U.S. Census Bureau. This is the last in a series of reports presenting the results of the evaluation of population estimation procedures used by the US Census Bureau during the 1970s. It covers a summary of earlier county population evaluations, a discussion of problems with the 1980 estimates evaluation, and suggests changes that should be made for the 1980 decade.

Folsom, R., B. Greenberg, D. Horvitz, and J. Abernathy. 1973. "The Two Alternate Questions Randomized Response Model for Human Surveys." *Journal of the American Statistical Association* 68 (343): 525-530. The authors introduce a model that is applicable when the population proportion with a non-sensitive attribute is not known in advance and two samples are required to estimate the population proportion with a sensitive attribute. It consists of using the sensitive question and two non-sensitive alternate questions in each sample, yielding two unbiased estimates.

Fonseca, L., and J. Tayman. 1989. "Post-censal Estimates of Household Income Distribution." *Demography* 26(1): 149-159. This method of estimating household income distributions uses the lognormal probability curve as a model. The lognormal curve is calibrated to census data and post-censal estimates of median income and income standard deviation are used to solve for proportionate income distribution. The paper concludes by discussing the method's utility and information sources which can be used to update median income and income standard deviation during the post-censal period

Frisėn, C. 1958. "Effectiveness of Our Tools for Estimating Population Change in Small Areas," pp. 229-232 in *Proceedings of the American Statistical Association, Social Statistics Section*. Alexandria, VA: American Statistical Association. The author provides a two part classification for assessing how functional population estimation methods are from the perspective of government: (1) a framework for examining symptomatic data; and (2) the estimation techniques themselves. He uses this to examine the vital rates method (by age race and sex) for age estimation and utility hookups for use in the housing unit method

Frisén, C. 1951. "Symptomatic Data in Population Estimates: Problems and Limitations." Paper presented at the annual meeting of the Pacific Sociological Association, Berkeley, CA (published in *Research Studies of the State College of Washington*, vol. 19, pp. 107-110) (No summary available).

Fuller, W. 1991. "Small Area Estimation as a Measurement Error Problem." pp. 333-352 in T. Kaul and J. Sengupta (Eds.) *Economic Models, Estimation, and Socioeconomic Systems: Essays in Honor of Karl A. Fox*. Amsterdam, The Netherlands: Elsevier Science Publishers. A discussion of how to combine model based approaches and probability (sample) based approaches to reduce variance and gain information thorugh "borrowed strength."

Galdi, D. 1985 "Evaluation of 1980 Subcounty Population Estimates." *Current Population Reports. P25, No. 963*. Washington, DC: U. S. Government Printing Office. This report presents an evaluation of estimates of total population for approximately 36,000 subcounty areas in the U.S using 1980 decennial census counts.

Galvez, J., and C. McLarty. 1996. "Measurement of Florida Temporary Residents Using a Telephone Survey." *Journal of Economic and Social Measurement* 22 (1): 25-42. This paper examines methodologies reported in the literature for estimating temporary migrants and suggests an `add-on' telephone survey design that is replicable. Results from a cross section of a Florida telephone survey are compared to results obtained using a similar instrument in Arizona. The telephone survey promises both cost effectiveness and a design that can be repeated in different locations.

Ghosh, M., and J. N. K. Rao. 1994. "Small Area Estimation: an Appraisal." *Statistical Science* 9: 55-93. The present article is largely an appraisal of some of the methods that "borrow strength" from related areas to find more accurate estimates for a given area or sets of areas. The methods include synthetic, sample size dependent, empirical best linear unbiased prediction, empirical Bayes and hierarchical Bayes estimation. The performance of these methods is also evaluated using some synthetic data resembling a business population. Empirical best linear unbiased prediction, as well as empirical and hierarchical Bayes, for most purposes, seem to have distinct advantage over other methods.

Gibson, C. 1988. "Estimates of Household, for Counties: July 1, 1985." *Current Population Reports, Series P-23, No. 156*. Washington, DC: U.S. Bureau of the Census. The author describes the methodology used by the Census Bureau to estimate households at the county level. He discusses group quarters and the adult population in households. Tables provide estimates of the number of households and the average number of persons per household by county and county equivalent. An appendix contains corrections to the 1980 census counts of population and housing units for selected states and counties.

Gibson, C. 1986. "Post-censal estimates of households by size and type for states and of total households for counties." Paper presented at the Population Association of America Annual Meeting, San Francisco. This paper examines a demographic approach to sub-national household estimates which assumes that national trends in the average adult population per household and in the distribution of households by type and by size occur at the sub-national level. The assumption is tested for the 1970-1980 decade using US census data. Results show that, for states, the mean absolute error is lowest for the 20 years and up population (.69%), followed by the 18 years and up population (.76%), the 15 years and up population (1.05%), and all ages (1.61%). Counties show a lower mean error for the 18 and up group (1.47%), followed by the 20 and up (1.59%), the 15 and up (1.7%), and all ages (2.95%).

Goldberg, D., V. R. Rao, and N. K. Namboodiri. 1964. "A Test of the Accuracy of Ratio-Correlation Population Estimates." *Land Economics* 40 (1): 100-102. This paper reports the results of testing the ratio-correlation method for the 83 counties of Michigan using 1960 census data. The results suggest that the method performs well.

Goldstein, S. 1954. "City Directories as Sources of Migration Data." *American Journal of Sociology* 69 (2): 169-176. The city directories of Norristown, Pennsylvania, were analyzed to determine their usefulness for the study of migration and occupational mobility. Tests showed that they provided a complete enumeration of the city's population and its occupational composition. Death certificates identify persons who disappeared from directory listings through death, and birth certificates and school records identify those who first appeared in the listings upon arriving at the minimum age for inclusion. Then, by the method of residues, the remainder were classified as either out-migrants or in-migrants. Thus, through corroborative use of diverse sources, American demographers have a valid substitute for the system of continuous population registers found in several European countries.

Gonzalez, M., and C. Hoza. 1978. "Small Area Estimation with Applications to Unemployment and Housing Estimates." *Journal of the American Statistical Association* 79: 7-15. The purpose of this study is to investigate methodologies for constructing inter-censal estimates of various characteristics of the population for small areas. The proposed methodology is illustrated mainly in the context of unemployment estimates, with one section utilizing dilapidated housing estimates. Alternative synthetic estimates of unemployment based on the 1970 Census 20-percent sample are investigated and their relative error is analyzed. The reliability of the synthetic estimates is discussed in the context of dilapidated housing estimates. Two types of regression models are studied, and the improvements obtained by excluding outliers from the regression are discussed.

Greenberg, B., A. Abdul-Ela, W. Simmons, and D. Horvitz. 1969. "The Unrelated Question Randomized Response Model: Theoretical Framework." *Journal of the American Statistical Association* 64 (326): 520-539. This paper develops a theoretical framework for the unrelated question randomized response model suggested by W. R. Simmons. It provides a method for allocating the total sample to each of the two sub-samples required by the method and recommendations on choosing values of the parameters that can be assigned by the researcher using this method

Grieco, E. 2002. "An Evaluation of Bridging Methods using Race Data from Census 2000." *Population Research and Policy Review* 21 (1-2): 91–107. As the author notes, the question on race from Census 2000 was different from previous censuses because it allowed respondents to select one or more races to indicate their racial identities. Because of this change, the race data from Census 2000 are not directly comparable with data from earlier censuses. She continues that researchers can use `bridging' methods to assign more than one race respondents to single race categories to maximize the comparability of Census 2000 race data with earlier censuses. In this paper, she uses several bridging methods to generate race population estimates and analyzes the variability in those estimates across six single race groups.

Habermann, H. 2006. “Research to Improve Population Estimates” Part of a presentation by H. Habermann at the Spring (May 18th -19th) 2006 meeting of the Census Advisory Committee for Professional Associations (CACPA). An overview of research on population estimation methods underway at the Census Bureau, with criteria for developing methods.

Hamilton, C. H. 1967. “The Vital Statistics Method of Estimating Net Migration by Age Cohorts.” *Demography* 4 (2): 464-487. The focus of this paper is the development and testing of a method of estimating deaths which occur during a decade of aging birth and death cohorts The results obtained by using the Vital Statistics (VS) method for age cohorts show that (1) the average census survival rate (CSR) method generally yields algebraically lower estimates of net migration than does the VS method; but (2) there are some striking exceptions which are apparently associated with errors in census enumeration by age, sex, and color. The exclusive use of the VS method in estimating net migration for age cohorts may lead to substantial error. The author notes that the magnitude of these errors in estimating net migration, as well as in census enumeration, can be roughly approximated if it is assumed that the use of the CSR method yields reasonably accurate estimates of net migration.

Hamilton, C. H. 1964. “Comparison of two formulas in making population estimates.” *Rural Sociology* 29: 426-431. The author compares the geometric and exponential methods of extrapolation for purposes of population projection or estimation. He notes that George Barclay’s portrayal of them can lead to confusion.

Hamilton, C. H., and J. Perry. 1962. “A Short Method for Projecting Population by Age from One Decennial Census to Another.” *Social Forces* 41 (December): 163-170. The logical basis and mathematical character of a short method for making population projections by age from one decennial census to another is discussed. The method is based on the assumption that age-specific vital rates and migration rates of the recent past will continue unchanged from one decade to the next. The method includes techniques for making projections of the population under 10 years of age at the end of the decade ahead.

Hamilton, C. H., and F. Henderson. 1944. “Use of the Survival Rate Method in Measuring Net Migration.” *Journal of the American Statistical Association* 39: 197-206. This paper seeks to demonstrate the superiority of the census survival method over life table survival methods in estimating net migration over inter-censal periods and to show how census survival rates can be adjusted for states and counties.

Hammer, R., P. Voss, and R. Blakely. 1999. “Spatially Arrayed Growth Forces and Small Area Population Estimates Methodology.” Paper presented at the Estimates Methods Conference, U. S. Bureau of the Census, Suitland, MD: (available online, http://www.census.gov/population/www/coop/popconf/paper.html, last accessed June, 2007). This paper attempts to quantify the spatial growth forces acting upon minor civil

divisions, as well as other small areas such as census tracts and block groups, and incorporate these forces into population estimation methods.
Happel, S., and T. Hogan. 1987. "Estimating the Winter Resident Population of the Phoenix Area" *Applied Demography* 3: 7-8. This paper described a survey-based method for estimating de facto populations that is applied to seasonal residents of Phoenix, Arizona.

Hauser, P., and B. Tepping. 1944. "Evaluation of Census Wartime Population Estimates and of Predictions of Postwar Population Prospects for Metropolitan Areas." *American Sociological Review* 9: 473-480. This paper describes the use of war ration books to estimate the population for US metropolitan areas during World War II.

Healy, M. 1982. "Using Administrative Records: Introduction and Basic Procedures." pp. 27 –37 in in E. S. Lee and H. F. Goldsmith (Eds.), *Population Estimates: Methods for Small Area Analysis*. Beverly Hill, CA: Sage Press. The author describes the component technique used by the U.S. Census Bureau to estimate population, one that uses 'administrative records' (IRS files) to estimate the migration component. She focuses on small areas - places and census tracts - and concludes that the production of 'administrative records' based population estimates for these types of geographies was unfeasible at the time of her writing.

Heer, D., and P. Herman. 1990. "Estimating the population of Los Angeles County Census Tracts by Ethnicity." pp. 83-88 in *1990 Proceedings of the Social Statistics Section*. Alexandria, VA: American Statistical Association. This study presents a methodology for estimation of population under 15 and 15-44 years, 45-64 years, and 64 and over years by sex and by census tract for 20 different ethnic groups. The innovation from traditional methodologies that would be inaccurate at the tract level involved expected values from a linear regression equation relating births (or deaths) in each year between 1980 and 1986 to time as the independent variable. Sample error is diminished by not recording the actual increase. Population estimates, the absolute change in population, and percent change in population are specified, as well as comparisons to the US Bureau of census estimates and the 1986 Test Census for Central Los Angeles. The US Bureau of Census estimates differed markedly for the Hispanic Population. The author's explanation is that the difference is due primarily to the assumption in the US Census that the rate of immigration was the same between 1975 and 1980, and 1980 and 1985. The difference between the Test Census, 56 tracts and 5 ethnic groups, is possibly related to differences in estimation of the Hispanic undercount. After a walking tour of tract 5309, the authors concluded that there was substantial evidence of undercount in the Test Census due to an undercount of housing units in backyards. With a Test Census showing an Hispanic neighborhood decline yet a 32% Hispanic County population increase, the author believes their figures indicate a low-income housing shortage and population increase in predominately Hispanic tracts and only a 19% countywide population increase

Heeringa, S. 1982. "Statistical Models for Small Area Estimation." pp. 126-132 in *Proceedings of the Social Statistics Section*. Alexandria, VA: American Statistical

Association. The author describes how small area estimates can be developed from sample survey data.

Hepburn, G., T. Mayor, and J. Stafford. 1976. "Estimation of Market Area Population from Residential Electrical Utility Data." *Journal of Marketing Research* 13 (3): 230-236. The authors note that an ability to project current population in small market trading areas is critical to many types of market research. The use of selected utility data to estimate current population in small geographic areas is evaluated.

Hill, K., H. Zlotnik, and J. Trussell. 1981. *Demographic Estimation: A Manual on Indirect Techniques*. Washington, DC: National Academy of Sciences Press. This is a manual published by the Committee on Population and Demography of the National Research Council on various technologies for estimating fertility and mortality from incomplete or inadequate data. Each chapter describes the basic method of estimation, assumptions on which it is based, the data required for its application, the method of application, and a sample application of the method to an actual set of data from a developing country. The following methods of estimation are covered: 1) estimation of fertility from information on children ever born; 2) estimation of childhood mortality from information on children ever born and children surviving; 3) estimation of adult survivorship from information on orphanhood and widowhood data; 4) estimation of adult mortality from distribution of deaths by age data; 5) life table data; 6) age distributions used for fertility or mortality estimations; 7) reverse survival information; and 8) successive census age distributions.

Hogan, H. 2008. "Measuring Population Change Using the American Community Survey." pp. 13-30 in S. Murdock and D. Swanson (Eds.). *Applied Demography in the 21st Century*. Dordrecht, The Netherlands, Springer. This chapter provides an overview of ways that the American Community Survey could be used in conjunction with population estimates to measure population characteristics and their change.

Hogan, H., and C. Cowan. 1980. "Imputations, Response Errors, and Matching in Dual System Estimation." pp. 263-268 in *Proceedings of the Section on Survey Research Methods*. Alexandria, VA: American Statistical Association. The authors propose a simple solution when imputations create problems for matching that is needed in dual system estimation. Their proposed solution - one should determine the number of non-matchable cases and subtract them from the counts of both systems. The paper discusses this proposed solution.

Hollman, W. and W.N. Rasmussen. 1976. "A Comparison of Four Methods of Estimating County Population with Mid-Decade Census Results, Selected California Counties." Paper presented at the Annual Meeting of the Population Association of America, Montreal, Quebec, Canada.

Hoque, N. 2008. "An Evaluation of Population Estimates for Counties and Places in Texas for 2000." pp. 125-148 in S. Murdock and D. Swanson (Eds.) *Applied Demography in the 21st Century*. Dordrecht, The Netherlands, Springer. The author

compares the estimates done in 2000 for Texas counties and cities by the Texas Demographer's office with Census 2000 results. The estimates were generated by three methods, Component Method II, ratio-correlation, and the Housing Unit Method. The author also provides descriptions of the methods and the context in which the estimates were done. The results suggest that averaged estimates perform better than estimates done by one method separately.

Hough, G., and D. Swanson. 2006. "An Evaluation of the American Community Survey: Results from the Oregon Test Site." *Population Research and Policy Review* 25(3): 257-273. Using a loss function analysis and other tools, this paper reports preliminary findings from a comparison of ACS and Census 2000 results in Multnomah County, Oregon, one of five national "local expert" test sites set up to compare ACS data collected at the time of Census 2000. The preliminary findings suggest that there are notable differences between some of the corresponding variables found in the ACS and Census LF that require more detailed examination. For example, the loss function analysis reveals notable differences for race and disability variables. In other comparisons of corresponding variables between ACS and Census 2000, differences are found within each of the four major areas of interest: (1) demographic characteristics, (2) social characteristics, (3) economic characteristics, and (4) housing characteristics, with housing characteristics showing the least similarity overall. These results also suggest that more detailed examinations are needed to understand differences between corresponding variables collected by ACS and the Census.

Hough, G., and D. Swanson 1998 "Toward an Assessment of Continuous Measurement: A Comparison of Returns with 1990 Census Returns for the Portland Test Site." *Journal of Economic and Social Measurement*. 24: 295-308. This paper is part of the initiation of an empirically-based discussion of the capability of the American Community Survey (ACS) to provide small area data comparable in quality to that provided by the census long form, the current gold standard for detailed, small area data. The authors compare mail return rates of the 1996 ACS to the 1990 mail return rates of the census long form for tracts in the test site and find that overall return rates are virtually the same--69 percent. However, ACS return rates are higher than those of the 1990 long form in that quartile of the tracts were the return rates for the long form were the lowest and lower in that quartile of the tracts where the long form rates were the highest.

Howe, A. 1999. "Assessing the Accuracy of Australia's Small-Area Population Estimates." *Journal of the Australian Population Association* 16(1-2): 47-63. This paper describes the process and methods used by the Australian Bureau of Statistics to assess the accuracy of its after population estimates. It notes, however, there appears to be no straightforward method of assessing these estimates. Errors that occur with population estimates can be attributed to several factors, both broad and specific to individual areas. These factors include inherent characteristics of the region, such as population size and growth rate; changes in the geographic boundaries; quality of input data; estimation method; and adjustments to control totals (state populations).

Ingram, D., J. O'Hare, F. Scheuren, and J. Turek. 2000. "Statistical Matching: A New Validation Case Study." Paper presented at the 2000 Joint Statistical Meetings, Indianapolis, IN, August 5–11. This paper employs data from the National Survey of American Families (NSAF) to assess the effect of violations of the conditional independence assumption on associations of health related measures and income measures observed in a data set constructed by statistically matching the National Health Interview Survey (NHIS) and the Current Population Survey (CPS). Contingency tables constructed from NSAF are compared with similar tables from the statistically-matched CPS/NHIS data sets. Chi-squared tests of independence applied both to the NSAF and CPS/NHIS contingency tables assess the extent to which associations between health and income variables were preserved in the statistically matched file. Implications for the use of the statistically matched data sets are explored and conjectures about the importance of failures in the conditional independence assumption in practical micro-simulation modeling settings are made.

Irwin, R. 1985. "County Inter-censal Estimates by Age, Sex, and Race: 1970-1980." *Current Population Reports P-23, No. 139.* Washington, DC: US Government Printing Office. This report describes the methodology used to develop a computer file of estimates of the population of U.S. counties by age, sex, and race for each year from 1970 to 1979. The report includes an evaluation of the quality of the estimates. The estimates presented on the tape are for 5-year age groups to age 85 and over for the total, White, and Black populations by sex. Census detail in both 1970 and 1980 was modified to permit consistent estimates throughout the decade. The age and sex detail in the 1980 census was not changed, but nearly 6 million persons of Spanish origin were moved out of the category "Other" (race not specified), the vast majority being transferred to White. A similar but much smaller modification was made in 1970. The Black population was not affected by the 1970 procedures and was increased only slightly in 1980. In addition to the inter-censal estimates from 1970 to 1979, the computer file contains the 1970 and 1980 census detail as modified, along with an extrapolation to July 1, 1980.

Irwin, R. 1976 "National Census Survival Rates in Population Projections for Small Areas." Paper presented at the annual meeting of the Population Association of America, Montreal Quebec, Canada. This long and highly specialized study demonstrates that if migration estimates obtained with national census survival rates are utilized in population projections done with life table survival rates, a substantial and very obvious misrepresentation will derive in calculating the age distribution of the projected population. Such misrepresentation can be avoided by using, instead of the life table survival rates, the national census survival rates adjusted according to future changes in expected mortality. The study demonstrates this theory by using adjusted national census survival rates for population projections for the years 1970-80.

Isaki, C. and L. Shultz. 1987. "Report on Demographic Analysis Synthetic Estimation for Small Areas." *SRD Research Report Number Census/SRD/RR-87/03*. Washington, DC: U.S. Bureau of the Census. (Available online, http://www.census2010.gov/srd/papers/pdf/rr87-03.pdf, last accessed May 2008). This report summarizes the work of the Census Undercount Adjustment for Small Area group as it pertains to Demographic Analysis (DA) Synthetic estimation. The papers include comparisons of the performance of the Demographic Analysis synthetic estimator with other estimators. In this report, the authors restrict discussion to the DA synthetic estimator and comparisons to the Census using various "measures of improvement" and "truths". The "measures of improvement" include mean absolute relative error as well as other indicators of performance such as errors in apportionment when compared to the "truth". The sources used to represent "truth" were the results of the 1980 Post-Enumeration Program (PEP) as well as constructed populations based on the 1980 Census.

Isaki, C. and L. Shultz, P. Smith, and G. Diffendal. 1985. “Small Area Estimation Research for Census Undercount – Progress Report.” *SRD Research Report Number Census/SRD/RR-85/07*. Washington, DC: U.S. Bureau of the Census. (Available online, http://www.census2010.gov/srd/papers/pdf/rr85-07.pdf, last accessed May 2008). This report summarizes work to-date on the examination of adjusting the total population count for undercount (because it is the first characteristic that is to be produced from the decennial census). The focus is in three general directions: (1) results of the 1980 PEP (Post Enumeration Program); (2) demographic analysis results for 1980; and (3) using the 1980 census data to simulate and evaluate the performance of potential adjustment methodologies. The remainder of the paper describes the limitations of the data tools previously mentioned, describes the adjustment methodologies being investigated and provides the results of work to-date.

Isserman, A. 1993. “The Right People, the Right Rates: Making Population Estimates and Forecasts with an Interregional Cohort-Component Model.” *Journal of the American Planning Association* 59 (1): 45-64. This article offers suggestions for making the technical choices required in a given forecast choices in order to create a more thoughtful population forecasting process. Emphasis is on an interregional cohort-component approach made possible by newly available in-migration and out-migration data for all U.S. counties. The interregional approach avoids conceptual problems and biases of conventional net migration approaches and can be used to make county population estimates, projections, and forecasts. Sample spreadsheets with formulas demonstrate the procedures for calculating rates and making projections. Population projections for fifty-five counties illustrate the effects of several methodological choices.

Jarosz, B. 2008. “Using Assessor Parcel Data to Maintain Housing Unit Counts for Small Area Population Estimates.” pp. 89-102 in S. Murdock and D. Swanson (Eds.) *Applied Demography in the 21st Century*. Dordrecht, The Netherlands, Springer. This paper describes the reasons why the San Diego Association of Governments switched from census and permit data to county assessor records to build housing unit files for purposes of estimating population by the Housing Unit Method. It describes the advantages of using the county assessor data as well as the disadvantages and difficulties they entail. It concludes with a description of the uses to which the data will be put, notably in the micro model known by the acronym of PECAS.

Judson, D. 2000. “The Statistical Administrative Records System and Administrative Records Experiment 2000.” Paper presented at the National Institutes of Statistical Sciences Data Quality Workshop, Morristown, NJ, November 30th – December 1st . This paper provides some general principles for using administrative records properly And gives an overview of the US Census Bureau’s StARS/AREX program, along with its history, goals and design. It concludes with some applications and evaluations of StARS 1999 and StARS 2000 versus Census 2000.

Judson, D., and M. Bauder. 2002. “Evaluating the Ability of Administrative Records Databases to Replicate Census 2000 Results at the Household Level.” Paper presented at the Annual Meeting of the American Statistical Association, New York City, NY, August 11th - 15th. The authors report on the results of the AREX experiment, which was designed to gain information about the feasibility of an administrative records (AR) census that was conducted in two areas of the United States in 2000. The authors report that 81.4 % of the census addresses linked on a one-to-one basis with an AREX address. They recommend that the Census Bureau: continue improving computerized record linkage, develop methods to reduce the time lag of AR data; test AR data for substitution and imputation purposes in future census tests; continue to improve race and Hispanic ethnicity modeling and imputation, and explore new uses of modeling.

Judson, D., and C. Popoff. 2004. “Appendix C, Selected General Methods.” Pp. 677-732 in . Siegel and D. Swanson (eds.) *The Methods and Materials of Demography, 2nd Edition*. Amsterdam, The Netherlands: Elsevier Academic Press. This describes a range of topics associated with measurement and analysis, none of which is strictly demographic. It covers sampling, interpolation, adjustment of data to marginal distributions, the role of computer models in demographic research, the use of matrix methods in demography, micro-simulation, and system dynamic models, regional demographic- economic models.

Judson, D., C. Popoff, and M. Batutis. 2001. “An Evaluation of the Accuracy of U.S. Census Bureau County Population Estimation Methods.” *Statistics in Transition* 5: 185-215. The authors focus on biases in the administrative records, which are used to estimate population changes, as clues to finding the possible reasons for over- and under-estimated counties. They examine, in detail, the sources of data used in the Census Bureau's methodology for reasons why there might be a systematic bias based on the circumstances under which data are gathered - either by the agency responsible for the administrative record used or by the person who will report. They test for these biases by using indicative economic and demographic data contained in the Bureau of the Census USA Counties 1994 release and identify causes of discrepancies in estimates that are systematic to the methodology, which suggest the direction and likely magnitude of the discrepancy. They report that their results are completely consistent with a prior hypotheses.

Kadane, J., and E. Ericksen. 1984. “Revised Estimates of State and Central City Populations on Census Day, 1980.” pp. 208-210 in *1984 Proceedings of the Social*

Statistics Section. Alexandria, VA: American Statistical Association. This paper uses information gathered by the U.S. Census Bureau for evaluation of the 1980 Census in order to give improved estimates of the population for 16 central cities and 50 states and remainders of states. The hierarchical model used to estimate the undercount is reviewed

Katzoff, M., and D. Malec. 1983. "Application of Some Common Spatial Models to Population Data." pp. 522-524 in *1983 Proceedings of the Social Statistics Section*. Alexandria, VA: American Statistical Association. This paper demonstrates the application of some results of a study of models and methods for assessing spatial autocorrelation in the context of small-area post-censal estimation to the problem of estimating county population sizes in the years following a U.S. Decennial Census.

Kilss, B., and W. Alvey. 1984 (Eds.) *Statistical Uses of Administrative Records: Recent Research and Present Prospects Vol I and Vol. II.* Washington, DC: U. S. Department of Treasury, Internal Revenue Service. This two volume set can be viewed as a handbook on the statistical use of administrative records. It provides six goals for statistical uses of administrative records, discusses developments up to the 1980s that were relevant to these recommendations, and describes elements of a strategy designed to achieve the six goals.

Kintner, H., and D. Swanson. 1993 "Measurement Errors in Census Counts and Estimates of Inter-censal Net Migration." *Journal of Economic and Social Measurement* 19 (2): 97-120. The authors present a method for generating confidence intervals around estimates of inter-censal net migration, made using the life table survival method, that incorporate estimates of census measurement errors. The life table survival method applies a life table to a census count to project survivors at some past or future time points. Net migration is then estimated as the difference between the projected number of survivors and the enumerated population at that time. The authors present confidence intervals based on mean square error, the sum of the variance, and squared bias. It is assumed that random variation in the number of net migrants in an age-sex group is due to random variation in mortality rates and to measurement errors in census counts. The authors illustrate the technique using data from a small area in Alaska.

Kitagawa, E. 1980. *Estimating Population and Income of Small Areas.* Washington, DC: National Academy Press. This monograph represents the work of the Panel on Small Area Estimates of Population and Income, which was charged with evaluating the Census Bureau's procedures for making post-censal estimates of population and per capita income for local areas. The Panel found that the methods used by the Census Bureau wee generally sound, but found that directionally biased estimates and large random errors for some areas. The Panel made eight recommendations.

Kitagawa, E., and B. Spencer. 1981. "On Estimating Population and Income for Local Areas." *Statistical Reporter* 81 (8): 377-381. The authors summarize the highlights of a 1980 report prepared by the Panel on Small-Area Estimates of Population and Income to evaluate the procedures used by the U.S. Bureau of the Census in making post-censal estimates of population and per capita income for local areas

Krótki, K. 1978. (Ed.) *Developments in Dual System Estimation of Population Size and Growth*. Edmonton, Alberta, Canada: University of Alberta Press. This book explores the collection of vital statistics and the estimation of population size by two independent systems, and comparing the results on a name-by-name basis. This book discusses a number of theoretical issues related to dual-systems of data collection, practical problems that arise in carrying out such systems, reports in detail on selected surveys (particularly in Africa where vital statistics systems are notably weak), and summarizes actual surveys as well as the state of the art.

Lahiri, S. 1999. "Future Population Estimates in Destable Populations." Paper presented the Estimates Methods Conference, U. S. Bureau of the Census, Suitland, MD. (available online, http://www.census.gov/population/www/coop/popconf/paper.html, last accessed June 2007). In using the Hamiltion-Perry method, one needs life table survival ratios and an assumption of the equality between the census survival ratios and the corresponding life table survival ratios. The author has shown elsewhere that such an equality holds good only when the population under study is either stationary or stable and therefore proposes a method that neither requires the knowledge of life table survival ratios nor the assumption of equality between the census survival ratios and the corresponding life table survival ratios. The technique proposed here assumes that the population under study is closed to migration, and it further requires the knowledge of age-specific growth rates, which can be estimated using earlier census enumerations over two or three inter-censal periods.

Land, K., and G. Hough. 1986. "Improving the Accuracy of Inter-censal Estimates and Post-censal Projections of the Civilian Noninstitutional Population: a Parameterization of Institutional Prevalence Rates." *Journal of the American Statistical Association* 81 (393): 62-74. This article examines the empirical validity of the assumption that the age-, sex-, and race-specific proportions of the population that are institutionalized--as estimated by the last census--remain constant until the next census by using data from the decennial censuses for 1940-1980 and, in light of substantial decade to decade changes in the age patterns of the institutional proportions for sex- and race-specific populations, seeks to develop alternative methods. To pursue the latter objective, parametric curves are fit to the age-specific institutional proportions for each population for each decade. A study of the observed historical variation in the parameters of these curves then leads to some suggestions about how their shapes can be estimated between censuses and projected beyond the latest available census to provide more accurate estimates and projections of the civilian non-institutional population.

Lee, E., and H. Goldsmith (Eds.). 1982. *Population Estimates: Methods for Small Area Analysis*. Beverly Hills, CA: Sage Publications. This monograph compiles papers presented at the Small Area Estimation Conference sponsored by the U.S. National Institute of Mental Health and held in Annapolis, Maryland, in November 1978. Ten papers by various authors are included, together with a series of evaluations and synopses by the editors. The overall theme of the papers is exploring new ways to estimate population numbers and characteristics. The following procedures are introduced and discussed: use of administrative records, synthetic approaches, surveys, simulations, and area cohort studies. The geographical focus of the studies is on the United States.

Long, J. 1993. "An Overview of Population Estimation Methods." pp 20-1 to 20-5 in D. Bogue, E. Arriaga, and D. Anderton (Eds.) *Readings in Population Research Methodology* (Vol. 5), Chicago: United Nations Population Funds and Social Development Center. This paper notes that population estimates require choosing appropriate benchmarks (or auxiliary data) in estimating the population change since the last census. It classifies methods into two general types, flow and stock. Flow methods, which are also known as component methods, because they require estimation of each component of population change since the last census. Stock methods relate changes in population size since the last census to changes in other measured variables: the number of housing units, automobile registrations, total number of deaths (and births), and tax returns. A special case is known as the regression or ratio-correlation method. Another commonly used population stock method at the sub-national level is the housing unit method. Finally, the paper notes that estimation methods are used in combination.

Long, J. 1993. "Post-censal Population Estimates: States, Counties, and Places." *Population Division Working Paper Number 3*. Washington, DC: U.S. Bureau of the Census. (Available online, http://www.census.gov/population/www/documentation/twps0003.html, last accessed, March, 2008).This paper described the population estimates program of the U.S. Census Bureau and its history. It describes methods and their data sources and concludes that population estimates are an integral part of the U.S. statistical system. It also contains a short bibliography.

Lunn, D., S. N. Simpson, I. Diamond, and L. Middleton.1998. "The Accuracy of Age-specific Population Estimates for Small Areas in Britain." *Population Studies* 52 (3): 327-344. This paper presents a multi-level statistical analysis of the accuracy of age-specific population estimates made for British local authorities in 1991. The aim of this work was to identify the factors that influence accuracy, and to investigate how these influences interact. The analyses showed that the following area characteristics are key factors: true population size; inter-censal population change; and percentages of unemployed residents, armed forces residents, and students. In addition, the authors found that the overall type of method used to calculate estimates is important, and that its effect varies both with area characteristics and with age-group.

Mandell, M., and J. Tayman. 1982. "Measuring Temporal Stability in Regression Models of Population Estimation." *Demography* 19 (1): 135-146. This paper introduces an empirical indicator designed to measure the temporal stability of regression models used to produce sub-national population estimates. An analysis of 67 counties in Florida centers on 1970 total population estimates generated from ratio-correlation and difference-correlation models. Comparisons are made between 8 different regression specifications and employ a quantitative measure of relative estimate accuracy. The major findings are that variable measurement and type are important determinants of estimate accuracy, and although temporal stability of the coefficients impacts estimation errors, the influence is not as pervasive as is suggested in the literature.

Marker, D. 1999. "Organization of Small Area Estimators Using a Generalized Linear Regression Framework." *Journal of Official Statistics* 15 (1): 1-24. In this article existing small area estimators are described, including Bayesian ones that have been proposed. A literature review is conducted for the estimators. The estimators are then organized from a general linear regression perspective, summarizing and showing where certain methods can be viewed as minor variations or generalizations of others. This includes a derivation of the conditions under which it is possible to view synthetic estimation as a form of regression. The article is aimed at pulling together the wide range of approaches that have been used for small area estimation.

Marks, E., W. Seltzer and K. Krotki. 1974. *Population Growth Estimation: A Handbook of Vital Statistics Estimation.* New York City, NY: The Population Council. The purpose of this handbook is to provide: an explanation of what the PGE technique is, some information on experiences around the world in its use, guidance on the general planning and the detailed design of a "Population Growth Estimation (PGE) study, including questions of cost, examples of procedures that may serve as models (even though imperfect ones) for the preparation of actual procedures, and a methodology for dealing with the inevitable weaknesses in the procedures used and in the estimates prepared. The PGE approach as used in the measurement or evaluation of vital statistics has three distinct features: the collection of reports of vital events by two quasi-independent data gathering procedures; the case-by-case matching of the reports in the two systems to determine which events are reported by both systems, and the preparation of an estimate of the number of events adjusted for omissions, or an estimate of the relative completeness of either system, on the basis of the match rates obtained. All three factors must be present for the study to be classified as one using the PGE approach.

Marquis, K., S. Wetrogan, and H. Palacios. 1996. "Towards a U.S. Population Database from Administrative Records." Unpublished Paper, U.S. Census Bureau (Available online, http://www.census.gov/srd/papers/pdf/km9601.pdf, last accessed May, 2008). This paper raises some orientation issues concerning a national database, provides some early results related to some of these issues, and discusses next steps for moving ahead with database building.

Martin, J., and W. Serow. "Estimating Characteristics using the Ratio-Correlation Method." *Demography* 15 (2): 223 - 233. Studies of the relative accuracy of methods of estimating the population at the sub-state level have generally found that the greatest degree of accuracy is provided by the ratio-correlation, or regression, method. This paper reports on research aimed at evaluating the effectiveness of this method in estimating the age and race composition of populations at the sub-state level. In addition to the basic multiple regression equation, variants such as stratification and the averaging of estimates from simple regression equations are also tested. For Virginia localities the most satisfactory results are generated by the non-stratified multiple regression equation.

Mason, A. 1996. "Population and Housing." *Population Research and Policy Review* 15 (6): 419-435. This paper assesses how population growth affects the housing sector and, in turn, economic development. Among other questions, the author examines how population growth increases the demand for residential land, housing, and urban infrastructure? Demographic methods were critical to answering the questions, especially assessing the impact of population growth on the demand for housing.

Massey, D., and R. Zenteno. 1999. "The Dynamics of Mass Migration." *Proceedings of the National Academy of Sciences*. 96 (Social Sciences): 5328-5335. The authors specify a set of equations defining a dynamic model of international migration and estimate its parameters by using data specially collected in Mexico. They then use it to project the a hypothetical Mexican community population forward in time. The authors argue that the model quantifies the mechanisms of cumulative causation predicted by social capital theory and illustrates the shortcomings of standard projection methodologies. The failure to model dynamically changing migration schedules yields a 5% overstatement of the projected size of the Mexican population after 50 years, an 11% understatement of the total number of U.S. migrants, a 15% understatement of the prevalence of U.S. migratory experience in the Mexican population, and an 85% understatement of the size of the Mexican population living in the United States.

McCarthy, K., A. Abrahamse, and C. Hubay. 1982. *The Changing Geographic Distribution of the Elderly: Estimating Net Migration Rates with Social Security Data*. Santa Monica, CA: RAND. This report presents a procedure that state and local planning agencies can use to monitor the movement of elderly persons into and out of individual U.S. counties. The procedure uses readily available Social Security data to produce accurate estimates of elderly net migration rates. The report describes the procedure's method and rationale, documents the formal estimation model, explains how to apply the model, and furnishes an illustration to guide the user.

McHenry, J. 1999. "A Local Alternative to Demographic Data Vendor Population and Household Estimates – Applying a Modified Housing Unit Approach to Individual Households." Paper presented the Estimates Methods Conference, U. S. Bureau of the Census, Suitland, MD. (available online, http://www.census.gov/population/www/coop/popconf/paper.html, last accessed June 2007). The author proposes a solution to problems encountered with small area data purchased from vendors that uses readily available local data combined with the housing method. Purchasers of small-area demographic data from national data vendors confront three interrelated problems. He tests his proposed solution using 1998 Miami-Dade County tax assessor records and complementary data and argues that the resulting population and household estimates compare very favorably to estimates sold by national demographic data vendors.

McHugh, K. 1990. "Seasonal Migration as a Substitute for, or Precursor to, Permanent Migration." *Research on Aging* 12 (2): 229-245. This article examines conditions under which seasonal movement serves as a substitute for, or precursor to, permanent migration among winter visitors to recreational vehicle (RV) parks in the Phoenix, Arizona area. Ties to the home community, ties to the seasonal residence, demographic characteristics, and commitment to a mobile lifestyle are specified as determinants of expectations of moving to Phoenix on a permanent basis. The model is tested using survey data for a sample of 1,001 winter visitors in Phoenix RV parks. Results of a discriminant analysis indicate that place ties and position in the life cycle condition expectations of permanently migrating to a seasonal residence.

McKibben, J. 2007. "The Use of School Enrollment Data to Estimate Undercount in Small Areas." Paper presented at the Applied Demography Conference, San Antonio, TX, January 8th. This paper shows how enrollment data can be used in an age-sex pyramid to estimate census coverage by examining ratios of school age children to adult age groups in which their parents are located.

McKibben, J. 1988. "Evaluating the Accuracy of County Population Estimates Using the Reverse Demographic Accounting Method." Paper presented at the annual meeting of the Population Association of America, New Orleans, LA. The author examines the accuracy of the U.S. Bureau of the Census's 1975 county population estimates for Indiana by comparing them with "Expected Census" figures generated by the reverse demographic method. This method develops "expected" 1975 census figures by algebraically subtracting the reported number of net migrants for the period 1975-1980 from the reported 1980 census count and adding to this figure reported deaths and subtracting reported births for the same period. The results reveal that the Bureau's population estimates tend to underestimate rural and overestimate urban counties in Indiana for 1975. Furthermore, only 20 of Indiana's 92 counties had less than a 5 percent absolute error, while 80 percent of the counties with errors in excess of 5 percent absolute error were rural.

McKibben, J., and D. Swanson. 1997. "Linking Substance and Practice: A Case Study of the Relationship between Socio-economic Structure and Population Estimation." *Journal of Economic and Social Measurement* 24 (2): 135-147. The authors argue that at least some of the shortcomings in accuracy of population estimates would be better understood by linking these methods with the substantive socio-economic and demographic dynamics that clearly must be underlying the changes in population that the methods are designed to measure. They provide a case study of Indiana over two periods, 1970-1980 and 1980-1990, which was selected because a common population estimation method exhibits a common problem over the two periods: its coefficients change. The authors link these changes to Indiana's transition to a post-industrial economy and describe how this transition operated through demographic dynamics that ultimately affected the estimation model.

McVey, W. 1974. *An Empirical Assessment of a Modified Censal Ratio Estimation Technique*. Ph.D. Dissertation, Department of Sociology, University of Alberta, Edmonton, Alberta, Canada. The author examines the censal ratio technique in the context of broader models and suggests modifications that his data support as improvements in accuracy and efficiency.

Morrison, P. 1982." Different Approaches to Monitoring Local Demographic Change." *RAND paper P-6743*. Santa Monica, CA: RAND Corporation. This paper is concerned with the problem of how to update demographic variables for small areas in the United States in the years following the decennial census. As background to considering the merits of a newly-proposed survey-based procedure, the author reviews various other conventional approaches to post-censal estimation--trend extrapolation, component analysis, and the use of symptomatic data. He then considers the survey approach as a complement to these other approaches

Mulry, M., and B. Spencer. 1993. "Accuracy of the 1990 Census and Undercount Adjustments." *Journal of the American Statistical Association* 88 (423): 1080-1091. This article describes the total error analysis and loss function analysis done by the U.S. Census Bureau when it recommended to its parent agency, the U.S. Department of Commerce that the 1990 census be adjusted for net undercount. In its decision not to adjust the census, the Department of Commerce cited different criteria than aggregate loss functions. Those criteria are identified and discussed.

Mulry, M., and B. Spencer.1991. " Total Error in PES Estimates of Population." *Journal of the American Statistical Association* 86 (416): 839-63. The authors describe a methodology for estimating the accuracy of dual systems estimates (DSE's) of population, census estimates of population, and estimates of undercount in the census. The DSE's are based on the census and a post-enumeration survey (PES). They apply the methodology to the 1988 dress rehearsal census of St. Louis and east-central Missouri and discuss its applicability to the 1990 U.S. census and PES. The methodology is based on decompositions of the total (or net) error into components, such as sampling error, matching error, and other non-sampling errors. Limited information about the accuracy of certain components of error, notably failure of assumptions in the 'capture-recapture' model, but others as well, lead the authors to offer tentative estimates of the errors of the census, DSE, and undercount.

Murdock, S., and D. Swanson (Eds). 2008 *Applied Demography in the 21st Century*. Dordrecht, The Netherlands: Springer. A compendium of selected refereed papers presented at the 2007 Applied Demography Conference. The book offers 21 chapters organized into five major sections: (1) data use and measurement issues; (2) population estimates and projections; (3) applied demography and health; (4) the breadth of applied demography; and educating the applied demographer.

Murdock, S., and D. Ellis. 1991. *Applied Demography*. Boulder, CO. Westview Press. This is an introductory text designed to show the application of demographic techniques in such areas as real estate, marketing, and regional and services planning. It includes chapters on demographic concepts and trends, data sources and the principles of data use, basic methods and measures of applied demography, and methods for estimating and projecting populations. The geographical focus is on the United States.

Murdock, S., H., S. Hwang, and R. Hamm. 1995. “Component Methods.” pp. 10-53 in N. W. Rives Jr., W. J. Serow, A. S. Lee, H. F. Goldsmith, and P. R. Voss (Eds.) *Basic Methods for Preparing Small-area Population Estimates*. Madison, WI: Applied Population Laboratory, University of Wisconsin. This chapter provides an overview of component methods. It is divided into three sections: (1) the conceptual framework of component method estimation, its sources of data, widely used methods, and key assumptions; (2) general guidelines for using component methods; and (3) an evaluation of the relative strengths and weaknesses of component procedures and the common problems encountered in their use.

Myers, D. (Ed.). 1990. *Housing Demography: Linking Demographic Structure and Housing Markets.* Madison, WI: University of Wisconsin Press. This book consists of 13 papers selected to bridge the gap between the separate analysis of population data and housing data in the United States. The approach is interdisciplinary, involving demography, economics, geography, sociology, and urban planning. A major objective of the book is theory building. The papers are grouped under three main headings: linking housing characteristics with household composition, life course and cohort models of housing choice, and housing consumption among the elderly.

Myers, D., and A. Doyle. 1990. “Age-specific Population-per-Household Ratios: Linking Population Age Structure with Housing Characteristics.” pp. 109-130 in D. Myers (ed.) *Housing Demography: Linking Demographic Structure and Housing Markets*. Madison, WI: University of Wisconsin Press. The authors develop a demographic framework for determining Persons per household and propose models for estimating these values with the age structure of the household population.

Namboodiri, N. K. 1972. "On the Ratio Correlation and Related Methods of Sub-national Population Estimation" *Demography* 9: 443-453. This paper recognizes that thee ratio-correlation method is one of several methods currently being used in the United States for making post-censal estimates of sub-national units such as states and counties. The author conjectures that it seems to have been commonly presumed by those writing on the subject that the working of the ratio-correlation method can be understood simply on the basis of the multiple regression theory. That this common notion may sometimes be untenable is demonstrated in this paper. The author points out that the ratio-correlation method of sub-national population estimation has certain characteristic features that make its application fall sometimes outside the usual contexts in which the conventional multiple regression theory is applicable. A number of alternatives to the ratio-correlation method are suggested. Some of the alternatives suggested are shown to yield relatively more accurate results when used for estimating post-censal populations of North Carolina counties.

Namboodiri, N. K., and N. Lalu. 1971. “The Average of Several Simple Regression Estimates as an Alternative to the Multiple Regression Estimate in Post-censal and Inter-censal Population Estimation: A Case Study.” *Rural Sociology* 36: 187-194. The authors find that averaging the population estimates produced by several simple regression models increases accuracy.

Newell, C. 1988. *Methods and Models in Demography*. London, UK: Belhaven Press. This book outlines the methods used to study population structure and change by presenting the major descriptive and analytical models developed by demographers to investigate the interrelationships between fertility, age, structure, and mortality. With illustrations, tables, and data drawn from a wide range of countries in both the developed and developing world, it explicates the potential uses and limitations of the current models for population analysis, estimation, and forecasting.

Norris, D., M. Britton, and R. Verma. 1982. “The use of Administrative Records for Estimating Migration and Population.” pp. 245-259 in *1982 Proceedings of the Social Statistics Section*. Alexandria, VA: American Statistical Association. This paper gives an overview of the current methods and data sources used to prepare the Canadian sub-national population estimates, with particular emphasis on the migration components. The differences between the June 1981 post-censal population estimates and the 1981 Census Population counts are presented, with similar comparisons for 1976 and 1971. Then the results of the examination of some possible refinements to the current methodology are outlined, followed by a discussion of some of the issues involved and future work planned.

Office of Management and Budget. Statistical Policy Office. 1993. “Indirect Estimators in Federal Programs.” *Statistical Policy Working Paper 21*. Federal Committee on Statistical Methodology. Subcommittee on Small Area Estimation. Washington, DC: Office of Management and Budget. This monograph comprises the eight reports which from a 1991 meeting by the Federal Committee on Statistical methodology that describe the use of indirect estimators and they are based on case studies from a variety of federal programs.

O’Hare, W. 1980. “A Note on the use of Regression Methods in Population Estimates.” *Demography* 17 (3): 341-343. Evidence which has emerged in the past few years indicates that the relative accuracy of population estimates derived from the ratio-correlation method and the difference-correlation method varies from state to state. In assessing the possible reasons why neither technique is uniformly more accurate, attention is focused on the temporal instability of the statistical relationships between symptomatic indicators and population change. The author concludes that further improvement in population estimates based on regression techniques is likely to be limited until demographers derive means of measuring and adjusting for these temporal changes.

O’Hare, W. 1976. “Report on a Multiple Regression Method for Making Population Estimates.” *Demography* 13 (3): 369-379. This paper re-introduces the ‘difference-correlation method of making population estimates first reported by Schmitt and Grier in 1966. The author uses it to produce population estimates for the counties of Michigan and finds that it produces estimates with a smaller mean percentage error than estimates produced by the ratio-correlation method. The author notes that differences of proportions which are used in the difference-correlation method will always have means of zero, while the ratios used in the ratio-correlation method have means which vary. Higher inter-correlations among the variables and increased temporal stability of the inter-correlations are two advantages of using differences rather than ratios

Oshungade, I. 1986. “Use of Percentage Change in Small Area Statistics.” *The Statistician* 35: 531-545. This paper presents a method for predicting the post-censal estimates for small area statistics. The method regressed the percentage changes in the dependent variable on the percentage changes in the symptomatic (independent) variables for a sample of some areas over two censuses to estimate the post-censal values for all the areas. The method can use raw and transformed data and has been used to estimate population values for 83 enumeration areas of Colchester District in 1981.

Palit, C., P. Voss, H. Krebs, and B. Kale. 1982. “Population Estimation from Administrative Records.” pp. 261-277 in *1982 Proceedings of the Social Statistics Section*. Alexandria, VA: American Statistical Association. The authors discuss the estimation of U.S. population from administration records, which are defined as "data which are routinely collected by some other agency for non-demographic purposes." Possible sources of error are examined, with a focus on lack of quality control and on variation in quality over time. The use of various smoothing or filtering procedures in order to reduce error is explained.

Perrone, S. 2008. “Address Coverage and Improvement and Evaluation Program – 2005 National Estimate of Coverage of the Master Address File.” pp. 37-85 in S. Murdock and D. Swanson (Eds.) *Applied Demography in the 21st Century*. Dordrecht, The Netherlands, Springer. The author reports on the Master Address File (MAF) portion of the U.S. Census Bureau’s ‘Address Coverage Improvement and Evaluation Program.’ She provides annual calculations, 2000-2005, of the national coverage of housing unit counts and updates the 2000 baseline accordingly. She compares the estimated true housing unit count to annual housing counts from the MAF. She concludes that maintaining an accurate list of addresses continues to be a challenge.

Pitkin, J. 1992. “A Comparison of Vendor Estimates of Population and Households With 1990 Census Counts in California.” *Applied Demography* 7(1):5-8. This study compares the 1990 U.S. population estimates of four national demographic data vendors with the actual counts of the 1990 Census. It was performed for a consortium of three public utility companies in California. In addition to total population, this study evaluates the estimates of three other demographic variables: (1) number of households; (2) population of Spanish origin; and ((3) white population. Differences in accuracy are found among the vendors and between variables. On average, the estimates were most accurate for total population, considerably less so for race (white population) and households and by far the least accurate for Hispanic population.

Pitken, J. 2008. “U. S. Immigration in the Rear View Mirror.” Paper presented at the annual meeting of the Population Association of America, New Orleans, LA. This paper uses reverse survival for estimating annual immigration prior to 2000 on a basis that is consistent with 2000 census counts. This new method is based on responses to the census question on Year of Entry that was asked of foreign-born persons in 2000. Although a similar question had been asked in earlier (1980, 1990, 2000) censuses, the 2000 census questionnaire for the first time asked for a response in terms of an exact year rather than a period of years. The advantage of this method is that it is, by definition, consistent with 2000 population counts.

Pittenger, D. 1982. "Critique of Administrative Record Procedures." pp. 39-42 in E. S. Lee and H. F. Goldsmith (Eds.), *Population Estimates: Methods for Small Area Analysis*. Beverly Hill, CA: Sage Press. The author extends a critique by M. Healy in this same volume by citing additional problems with using the 'administrative records method' (IRS files) to estimate internal migration in the U.S. He points out that a properly done Housing Unit Method done in conjunction with this technique could overcome some of these problems.

Pittenger, D. 1977. "Population Forecasting Standards: Some Considerations Concerning Their Necessity and Content." *Demography* 14 (3): 363-368. This paper discusses the nature of population forecasting and provides guidelines for standards that are in line with good professional practice yet do not stifle creativity or technical advances.

Plane, D., and P. Rogerson. 1994. *The Geographical Analysis of Population, with Applications to Planning and Business*. New York City, NY: John Wiley and Sons. This textbook concentrates on both applied demographic and planning techniques which rely upon geographical aspects of population data. It describes methods used to assess the impact of population change on facility demand, school enrollment, changes in product market, transportation and recreation demand forecasting. It contains problems for training students and has solutions to these problems with actual data.

Platek, R., J. Rao, C. Särndal, and M. Singh. 1987. *Small Area Statistics: An International Symposium*. New York City, NY: John Wiley and Sons. This book contains the papers from the symposium. It is divided into five sections: (1) policy issues; (2) population estimation for small areas; (3) theoretical developments; (4) the use of small area estimation by agencies and other organizations; and (5) a panel discussion. The second section contains discussion of regression and other models for small area population estimation.

Pollard, A., F. Yusuf, and G. Pollard.1991. *Demographic Techniques, 3rd Edition*. Sydney, Australia: Pergamon Press. This monograph is an introduction to demographic techniques for students who have a general interest in population studies. Exercises at the end of each chapter allow students to attempt practical problems. Sources of demographic data and statistics, basic demographic measures, the use of life tables, and the applications of stationary population models are reviewed. All aspects of mortality (cause of death, death rates, infant and maternal mortality), and fertility (family size, natural increase, fertility schedules, and differential fertility) are covered in detail). Other topics include stable population and population models, population projections, demographic sample surveys, multiple-decrement tables, testing the accuracy of demographic data, and estimating demographic measures from incomplete data.

Popoff, C., and D. Judson. 2004. "Some Methods of Estimation for Statistically Underdeveloped Areas." Pp. 603-641 in J. Siegel and D. Swanson (eds.) *The Methods and Materials of Demography, 2nd Edition*. Amsterdam, The Netherlands: Elsevier Academic Press. This chapter covers techniques for making estimates from deficient data. The methods include stationary, stable and quasi-stable population techniques, samples, and dual system estimators.

Pozzi, F., Small, C., and G. Yetman. 2003. "Modeling the distribution of human population with night-time satellite imagery and gridded population of the world." *Earth Observation Magazine* 12(4): 24–30. The authors use the "World Stable Lights" dataset as a potential means to refine the spatial detail of the population dataset. They compared the Log10 of population density to the nighttime light frequency for sample of regions of the world with spatially detailed administrative data and found a consistent relationship between population density and light frequency. Based on this relationship, they developed a transfer function to relate light frequency to population density and a mass-conserving algorithm that relocates fractions of populations within large administrative units to locations of lighted settlements.

Prasad, N., and J. N. K. Rao. 1990. "The Estimation of the Mean Squared Error of Small-Area Estimators." *Journal of the American Statistical Association* 85 (409): 163-171. The authors investigate three small area models that are all special cases of a mixed linear model involving fixed and random effects. Using Henderson's general theory a 2-stage estimator of a small-area mean under each model is found. Results of a Monte Carlo study on the efficiency of 2-stage estimators and the accuracy of the proposed approximation to Mean Square Error (MSE) and its estimator are summarized. The authors find that the MSE approximation provides a reliable measure of uncertainty associated with a 2-stage estimator as well as asymptotically valid confidence intervals on a small area mean as the number of small areas tends to infinity.

Preston, S. 1993. "The Contours of Demography: Estimates and Projections." *Demography* 30 (4): 593-606. The author discusses the scope of the field of demography and notes that the techniques for performing research in demography was improving more rapidly than in other areas of social science. The author identifies several areas of research in which demographers can be expected to be asked for answers from various clients, including voluntary international migration – a prescient inclusion.

Preston, S., P. Heuveline, and M. Guillot. 2001. *Demography: Measuring and Modeling Population Processes*. Oxford, UK: Blackwell. Besides the usual materials and methods, found in a demography textbook or reference book, this one contains three chapters of interest: (1) evaluating data quality; (2) stable population theory, and (3) indirect estimation methods.

Prevost, R. 1999. "Design Alternatives for Building Block Estimates." Paper presented the Estimates Methods Conference, U. S. Bureau of the Census, Suitland, MD. (available online, http://www.census.gov/population/www/coop/popconf/paper.html, last accessed

June 2007). This paper described the prototyping done by the Census Bureau of a Statistical Administrative Records System (StARS) that will combine several major federal administrative records systems such as the IRS individual tax returns and information returns, HCFA Medicare enrollment data, SSA 'Numident' information, HUD Tract Rental Assistance Certification System, Indian Health Service, and Selective Service files. This paper provides background information on StARS and proposes design alternatives for the creation of building block (census block and tract) population and housing estimates after 2000.

Prevost, R. 1996. "Administrative Records and the New Statistical Era." Paper presented at the 1996 Annual meeting of the Population Association of America. New Orleans, LA, May 9th -11th. The author notes that the Census Bureau could reinvent and revolutionize data collection and processing operations in the next millennium by record integration, research, and maximizing the utility of administrative records.

Prevost, R. 1994. "State Housing Unit and Household Estimates: April 1, 1980, to July 1, 1993. *Current Population Reports, Series P-25, No.* 1123. Washington, D: US Government Printing Office. A description of methods, data, procedures, and results for the state level estimates of housing units done the period from 1980 to 1993.

Prevost, R., and C. Leggieri. 1999. "Expansion of Administrative Records Uses at the Census Bureau: A Long-Range Research Plan." Paper presented at the conference of the Federal Committee on Statistical Methodology (available online, http://www.fcsm.gov/99papers/prevost.pdf, last accessed, May 2008). The authors present a vision and strategy for the process through which the Census Bureau could reinvent and revolutionize data collection and processing operations in the next millennium. Their proposed strategy is designed to draw strength from current operations through integration, research, and maximizing the utility of administrative records. The authors argue that this strategy will provide the ability to expand annual inter-censal products, and alternative approaches for statistically representing the United States with cost reductions in 2010. A major component of this strategy is a complete assessment of how effectively the Census Bureau is able to address customers and stakeholder needs. Thus, they include a needs assessment in their long-term strategy, one with three basic goals: (1) Reduced direct data collection cost; (2) Increased data quality; and (3) Reduced respondent burden.

Prevost, R., and D. Swanson. 1986. "A New Technique for Assessing Error in Ratio-correlation Estimates of Population: A Preliminary Note." *Applied Demography* 1(November): 1-4. Among other things, this paper provides a proof of the algebraically equivalency of a series of weighted separate synthetic estimators and the ratio-correlation regression method. The authors find that the coefficients in the ratio-correlation serve as weights, including the intercept term, which serves as a weight for an estimator formed by multiplying the proportion of the population of the area in question (e.g., a county in a given state) at the last census by the current population (estimate) for the aggregated set of areas (e.g., the state).

Price, D. 1955. “Examination of Two Sources of Error in the Estimation of Net Internal Migration.” *Journal of the American Statistical Association* 50 (271): 689-700. Relative to using the survival rate method to estimate net migration, the author exams errors stemming from two sources in this method: (1) survival rates; and (2) (net) undercounting of the population. Using data from all (48 at the time) states, he examines the errors found by in using a single set of survival rates and state-specific life tables in ten states and concludes that the median error resulting from survival rates is about 14 percent of the estimate of net migration. Using census survival rates, he estimates about one third of net migration estimates are in error by 25 percent or more due to the effects of (net) under-enumeration.

Pursell, D. 1970. "Improving Population Estimates with the use of Dummy Variables" *Demography* 7 (1): 87-91. In a case study of West Virginia counties, he author finds that the ratio-correlation model of estimating county population can be improved with the use of dummy variables and stratification to represent “type” of county, such as region or rate of development.

Purcell, N. and L. Kish. 1980. “Post-censal Estimates for Local Areas (or Domains). *International Statistical Review* 48 (1): 3-18. The authors review methods for making small area estimates and present a classification for them, along with an extensive bibliography. They propose adding Iterative Proportional Fitting to these methods and justify both the model and the method within a broad theoretical framework. They provide some empirical results for their model that show improvements over other methods.

Purcell, N. J. and L. Kish. 1979. “Estimation for Small Domains.” *Biometrics* 35: 365-384. This paper describes the use of Iterative Proportionate Fitting (ITF) for fitting sample data with iterative techniques to a flexible non-linear model that preserves specified relationships found in an earlier census. The authors find that their Structure Preserving Estimates (SPREE) provide substantial accuracy improvements over symptomatic accounting techniques, to include synthetic estimation, regression-symptomatic indicators, sample-regression methods, and James-Stein, Bayesian and Shrinkage estimates.

Qiu, F., K. Woller, and R. Briggs. 2003. “Modeling Urban Population Growth from Remotely Sensed Imagery and TIGER GIS Road Data.” *Photogrammetric Engineering and Remote Sensing* 69: 1031–1042. The authors modeled population growth from 1990 to 2000 in the north Dallas-Fort Worth Metroplex using two different methods: (1) a conventional model based on remote sensing land-use change detection; and (2) a newly devised approach using GIS-derived road development measurements. These methods were applied at both city and census-tract levels and were evaluated against the actual population growth. It was found that accurate population growth estimates are achieved by both methods. At the census-tract level, their models yielded a comparable result with that obtained from a more complex commercial demographics model. At both city and census-tract levels, models using road development were better than those using

land-use change detection. In addition to being efficient in cost and time, the authors argue that their models provide direct visualization of the distribution of the actual population growth within cities and census tracts when compared to commercial s demographic model.

Rasmussen, W.N. 1975. "The Use of Driver License Address Change Records for Estimating Interstate and Intercounty Migration." pp. 16-22 in *Inter-censal Estimates for Small Areas and Public Data Files for Research, Small Area Statistics Papers, Series GE-41, No. 1*. U. S. Bureau of the Census, Washington, DC: US Government Printing Office. This paper introduces the Driver License Address Change (DLAC) method, which was developed in the 1970s by the state of California to estimate migration in a component model sued to estimate state population. It was subsequently modified to provide estimates for California's counties. In broad outline, the method estimates changes in county population proportion as a function of changes in various data series for three age groups: the population under 18; population, ages 18-64; and the population 65 years and over.

Raymer, J. and A. Rogers. 2007. "Using Age and Spatial Flow Structures in the Indirect Estimation of Migration Streams." *Demography* 44 (2): 199-223. This article outlines a formal model-based approach for inferring interregional age-specific migration streams in settings where such data are incomplete, inadequate, or unavailable. The estimation approach relies heavily on log-linear models, using them to impose some of the regularities exhibited by past age and spatial structures or to combine and borrow information drawn from other sources. The approach is illustrated using data from the 1990 and 2000 U.S. and Mexico censuses.

Raymondo, J. 1992. *Population Estimation and Projection: Methods for Marketing, Demographic, and Planning Personnel*. New York, NY: Quorum Books. The author provides a step-by-step guide for dealing with the issues associated with the methods of producing population estimates and projections. It begins by describing the general nature and differences of population estimates and projections and then covers several methodological issues that provide a foundation for the basic concepts of demographic analysis related to the production of population estimates and projections. Census concepts, including geography, are covered and the book includes a general discussion on major sources of data for the estimation and projection process.

Redfern, P. 1989. "Population Registers: Some Administrative and Statistical Pros and Cons." *Journal of the Royal Statistical Society, Series A* 152: 1-41. The author discusses the advantages and disadvantages of introducing a central population register in the context of the United Kingdom and describes the use of population registers in other European countries. He covers the statistical implications of better population registers, the various population records that do exist in the United Kingdom, and the proposal to introduce identity cards in Australia. A summary of a discussion that took place following presentation of the paper at the Royal Statistical Society is included.

Rees, P. 1994. "Estimating and Projecting the Populations of Urban Communities." *Environment and Planning* 26 (11): 1671-97. The author describes a model for estimating

and projecting the populations of communities living in small areas within cities. The model provides a means of updating the demographic inputs needed for projection between censuses and means of developing scenarios of demographic change and housing development. Illustrative projections are discussed and interpreted for the northern English city of Bradford.

Rees, P., and R. Woods 1986. "Demographic Estimation: Problems, Methods, and Examples." pp. 301-343 in R. Woods and P. Rees (Eds). *Population Structures and Models*. London, UK: Allen and Unwin. In this chapter, the authors outline some of the issues the demographic researcher has to confront when investigating populations that vary over space, as well as some of the techniques that may be employed in estimating the data needed for such investigations. The elements considered by the authors include "(a) the concepts used in defining the population stocks and flows; (b) the age-time frameworks adopted; (c) the spatial frameworks used; and (d) the temporal and accounting frameworks employed. These concepts are applied to two examples.

Rees, P., P. Norman, and D. Brown. 2004. "A Framework for Progressively Improving Small Area Population Estimates." *Journal of the Royal Statistical Society, Series A* 167: 5–36. The paper presents a framework for small area population estimation that enables users to select a method that is fit for the purpose. The adjustments to input data that are needed before use are outlined, with emphasis on developing consistent time series of inputs. The authors show how geographical harmonization of small areas, which is crucial to comparisons over time, can be achieved. For two study regions, the East of England and Yorkshire and the Humber, the differences in output and consequences of adopting different methods are illustrated. The paper concludes with a discussion of how data, on stream since 1998, might be included in future small area estimates.

Reese, A. J. 2006. "A Comparison of Housing Unit Estimates to the American Community Survey's Aggregated Master Address File." Paper Presented at the Annual meeting of the Southern Demographic Association, Durham, NC. This report presents a comparison of the estimates of housing units, produced by the Population Division of the U.S. Census Bureau, to the number of valid units in the Master Address File (MAF), at the national, state, and county levels of geography for 2002 through 2005. Geographic patterns based on differences between the two datasets were detected and discussed in an attempt to gain a deeper level of understanding of the differences between the two sources of data on the housing stock.

Reibel, M., and A. Agrawal, A. 2007. "Areal Interpolation of Population Counts using Pre-classified Land Cover Data." *Population Research and Policy Review* 26 (5–6): 619-633. To combine population counts across incompatible tract geographies corresponding to successive census enumerations, the authors propose dasymetric areal interpolation using a pre-classified urban land cover data layer. A test of the interpolation technique using the National Land Cover Dataset (NLCD) shows significant error reduction over area weighted interpolation of the same data. The authors find that the NLCD compares favorably with other common techniques when considered on the basis of accuracy, precision and ease of use.

Reibel, M., and M. Bufalino. 2005. “A Test of Street Weighted Areal Interpolation using Geographic Information Systems.” *Environment and Planning A*, 37: 127–139. The authors test a technique of areal interpolation using Geographic Information Systems (GIS) that employs a digital map layer representing streets and roads to derive varying density weights for small areas within aggregation zones. They find that the technique reduces errors in estimation when compared to estimates derived using the commonly applied area weighting technique, with its assumption of uniform density. They argue that the street weighting technique is much easier to use than other interpolation techniques that have also been shown to reduce error when compared to area based weighting.

Rives, N.W. 1982. “Assessment of a Survey Approach.” pp. 79-96 in E. Lee and H. Goldsmith (Eds.) *Population Estimates: Methods for Small Area Analysis*. Beverly Hills, CA: Sage Publications. The author examines survey procedures as sources of population estimates and suggests that they be considered in areas where sampling frames exist (Tiger/Line files) that allow for clusters samples to be easily drawn. He discusses costs and related implementation obstacles and offers suggestions for overcoming them.

Rives, N.W., and W. Serow. 1984. *Introduction to Applied Demography: Data Sources and Estimation Techniques*. Beverly Hills, CA: Sage Publications. This was the first book published on applied demography. It identifies kinds and sources of demographic data and then explains how to use this information to determine demographic trends and their consequences.

Rives, N.W., W.J. Serow, A. Lee, H. Goldsmith and P. Voss. 1995. *Basic Methods for Preparing Small-Area Population Estimates*. Madison, WI: Applied Population Laboratory, University of Wisconsin. The chapters in this book largely were complied from papers given at a workshop on small area population estimation organized for the 1984 meeting of the Southern Demographic Association. Including an afterword, there are eight chapters and an annotated bibliography. Component methods are discussed in the book along with the administrative records method, censal-ratio methods. The housing unit method, and the ratio-correlation method. One chapter provides an assessment of these methods.

Robinson, J.G., P. Das Gupta and B. Ahmed. 1990. “A Case Study in the Investigation of Errors in Estimates of Coverage Based on Demographic Analysis: Black Adults Age 30-64 in 1980.” pp. 11-20 in *Proceedings of the Social Statistics Section, American Statistical Association*. Alexandria, VA: American Statistical Association. This paper discusses the possibility of error in some specific components that are used to develop the demographic estimates of population and coverage for U.S Blacks--namely, births, deaths, and base population. The effect of classification error has also been considered. It has been shown that the net effect of these various sources of error is to overstate the current estimates of percent net undercount for Blacks." The focus is on the 1980 census.

Rogers, A. 1975. *Introduction to Multiregional Mathematical Demography*. New York City, NY: Wiley Interscience. This textbook is about the modeling, analysis and projection of populations using concepts from multi-state demography. It is focused on migration and regional analysis, particularly in terms of the growth and distribution of multiregional population systems.

Rogers, A. 1995. *Multiregional Demography: Principles, Methods and Extensions*. New York City, NY: John Wiley and Sons. This treatise is primarily designed as reference book on multiregional demography. It combines and updates the author's previous work on this subject and is designed for the readers familiar with the mathematics of uni-regional demographic analysis. Particular attention is paid to the analysis of migration and its contribution to projections of the growth and distribution of multiregional population systems. The concepts are illustrated with observed data from many different countries.

Rogers, A., and L. Castro. 1986. "Migration." pp. 157–210 in A. Rogers and F. Willikens (Eds.). *Migration and Settlement: A Multiregional Comparative Study*. Dordrecht, Netherlands: D. Reidel Publishing Company. In 1976, the International Institute for Applied Systems Analysis initiated a study of migration and population distribution patterns in its seventeen member nations. Much of this study appears in this book. In this chapter, the authors provide a conceptual and methodological overview of the study of migration and catalogue sources of data that can be used in its study, especially sources found in the member nations of IIASA.

Rosenberg, H. 1968. "Improving Current Population Estimates Through Stratification" *Land Economics* 44: 331-338. The author suggests ways in which regions can be stratified, such as level of economic development and how these stratification schemes can be incorporated into standard methods of population, especially regression-based ones such as the ratio-correlation technique.

Rowntree, J. 1990. "Population Estimates and Projections." *Population Trends* 60 (Summer): 33-34. This article describes briefly the different series of official population estimates and projections prepared by the United Kingdom's Office of Population Censuses and Surveys. It also refers to the 'extrapolated estimates' that were introduced at the time this article was written.

Sailar, P., and M. Weber. 1998. "The IRS population count: An update." *1998 Proceedings of the Section on Survey Research Methods*. Alexandria, Va.: American Statistical Association. This is a continuation of the authors' study on the problems involved in using IRS administrative records to prepare estimates of the U.S. population size. "Organizationally, this paper is divided into four sections. They demonstrate how administrative records can be used to compute a population estimate and then discuss the reliability of this estimate. Next, they compare estimates from their data base, classified by age, sex, and state, to population data published by the Census Bureau. And finally, they summarize their conclusions and make some recommendations for further research."

Scardamalia, R. 1992. "Estimates Base Adjustment." *Applied Demography* 6 (3): 2-3. This brief paper raises questions about the possible effects of net undercount adjustments on the process of making post-censal population estimates.

Scheuren, F. 1999. "Administrative Records and Census Taking." *Survey Methodology* 25(2): 151–160. The author considers the possible uses of administrative records to enhance and improve population censuses. After reviewing previous uses of administrative records in an international context, he puts forward several proposals for research and development towards increased use of administrative records in the American statistical system.

Schmitt, R. 1956. "Estimating Daytime Populations." *Journal of the American Planning Association* 22 (2): 83-85. The author notes the difficulty in estimating de facto populations because of the paucity of data and makes suggestions on how these short-comings might be overcome in combination with selected population estimating methods.

Schmitt, R. 1954. "A Method of Projecting the Population of Census Tracts." *Journal of the American Planning Association* 20 (2): 102. Using 1930 and 1940 data, the author constructs a ratio-correlation method for the 84 census tracts comprising the Honolulu SMSA shows how it can be used to project the population of these tracts by substituting 1950 data and 1940 data for the 1930 and 1940 data, respectively.

Schmitt, R. 1954. "An Application of Multiple Correlation to Population Forecasting." *Land Economics* 30 (3): 277-279. This paper is a variation of the paper published by the author in 1954 in the *Journal of the American Planning Association*, although it largely relies on growth rates.

Schmitt, R. 1953. "A New Method of Forecasting City Population." *Journal of the American Planning Association* 19 (1): 40-42. The author presents a regression model with two independent variables, density and prior growth, for estimating/forecasting population and illustrates its use for Seattle Washington. The model was calibrated using 1930 and 1940 data for thirty-seven (large) cities. He finds that the model performs well for these same cities when applied to 1950 and tested against 1950 census counts.

Schmitt, R. 1952. "Short-Cut Methods of Estimating County Population." *Journal of the American Statistical Association* 47 (258): 232-238. Using preliminary 1950 census data for the 39 counties of Washington as a case study, the author tests the accuracy of a series of 20 short-cut methods of estimate. He includes various weighting methods in this study and finds that properly weighted estimates combined from several methods show promise.

Schmitt, R. 1952. "Estimating Current Populations of Census Tracts." *Sociology and Social Research*, 37, 12–15. The author tests the accuracy of different estimation methods using 1950 census data for the 29 census tracts of Honolulu, Hawai'i.

Schmitt, R., and A. Crosetti. 1954. “Accuracy of the Ratio-correlation Method for Estimating Post-censal Population.” *Land Economics* 30: 279-281. This paper introduces the ratio-correlation method and tests it using the 39 counties in Washington State.

Schmitt, R., and J. Grier. 1966. “A Method of Estimating the Population of Minor Civil Divisions.” *Rural Sociology* 31: 355-361. The authors introduce the “difference-correlation’ regression method using the populations and symptomatic indicators of minor civil divisions in Hawai’i and test its accuracy with 1960 census data. They find that it produced more accurate results than other methods included in the test.

Schoen, R. 1988. *Modeling Multigroup Populations*. New York City, NY: Plenum Press. This three-part book makes contributions to two areas of formal demography and extend their applications: (1) Multistate methods; and (2) the “two-sex” problem. The first part of the book covers life tables and stable population theory. The second part deals with multistate population models and the third part with the two-sex problem.

Schroeder, E., and D. Pittenger. 1983. “Improving the Accuracy of Migration Age Detail in Multiple-Area Population Forecasts.” *Demography* 20 (2): 235-248. Population projections are often required for many geographical areas, and must be prepared with maximal computer and minimal analytical effort. At the same time, realistic age detail forecasts require a flexible means of treating age-specific net migration. This report presents a migration projection technique compatible with these constraints. A simplified version of Pittenger's model is used, where future migration patterns are automatically assigned from characteristics of historical patterns. A comparative test of age pattern accuracy for 1970-1980 indicates that this technique is superior to the commonly used plus-minus adjustment to historical rates.

Serow, W., and N. W. Rives. 1995. “Small Area Analysis: Assessing the State of the Art.” pp. 1-9 in N. W. Rives, W. Serow, A. Lee, H. Goldsmith, and P. Voss (Eds). *Basic Methods for Preparing Small-Area Population Estimate*s. Madison, WI: Applied Population Laboratory, University of Wisconsin. The authors introduce the book they helped to edit in this chapter and discuss the problem of small-area population estimates in general.

Shahidullah, M. 2003. "Uses, Methods, Input Data, and Evaluation of Population Estimates: Results from a State Survey." *Applied Demography* 16: 5-7. The reports the results of a 1999 survey of the states engaged in the Federal-State Cooperative Program for Population Estimates that was aimed at assessing the uses, methods, input data, and testing and evaluation of population estimates for each state and Puerto Rico. The authors report that 48 states and Puerto Rico returned completed questionnaires.

Shahidullah, M., and M. Flotow. 2005. “Criteria for selecting a suitable method for producing post-2000 county population estimates: A case study of population estimates in Illinois.” *Population Research and Policy Review* 24 (3): 215-229. The authors compare 2000 county population estimates for Illinois against 2000 census counts. Administrative records (ADREC) and ratio correlation (Ratio-CORR) methods were used to produce two sets of controlled county estimates for 2000; a third set represented an average of the estimates reached using these methods. Another set using the ADREC method was not controlled to any estimate. ADREC estimates were more accurate than estimates from the Ratio-CORR or Average method in terms of Mean Absolute Percent (MAPE) or weighted MAPE. Undercount adjustment in general improved the accuracy of the estimates for all three methods. A top-down or bottom-up approach worked equally well. As a single method, ADREC performed best.

Shyrock, H. 1959. "Annotated List of Estimates of Net Migration: 1940 to 1950, by the Residual Method." *Population Index* 25 (1): 16-24. This report provides an annotated bibliography of methods for estimating net migration and actual estimates by thematic area and state.

Shryock, H. 1950. "Forecasts of Population in the United States." *Population Studies* 3(4): 406-412. The author describes how the Census Bureau published its first modern forecasts in 1947 as a cooperative venture with the Scripps Foundation and outlines organizational, research, and production plans for the future.

Shryock, H. 1938. "Methods of Estimating Post-censal Population." *American Journal of Public Health* 28: 1042-1047. A review of methods of population estimates, similar to his 1936 publication described in the following citation.

Shryock, H. 1936. "Population Estimates in Post-censal Years." *Annals of the American Academy of Political and Social Science* 188:167-176. The author traces the start population estimates done by the US Census Bureau to its establishment as a permanent organization in 1902 and describes the methods used over the years in estimating national, state and city populations. He describes why methods changed and summarizes the results of then-recent research by the Census Bureau on various methods, including the use of school enrollment data.

Shryock, H., and N. Lawrence. 1949. "The Current Status of State and Local Population Estimates in the Census Bureau." *Journal of the American Statistical Association* 44 (246): 157-173. This paper provides a brief history of population estimates done by the U.S. Census Bureau and a description of the methods being used at the time this paper was written. It compares the accuracy and general efficacy of these methods and discusses their strengths and weaknesses. The methods include Component Method I, Component Method II, arithmetic extrapolation, and apportionment.

Siegel, J. 2002. *Applied Demography: Applications to Business, Government, Law, and Public Policy.* San Diego, CA: Academic Press. This fourteen-chapter book can be used for instruction or reference. It is focused on the U.S. Chapters 9 and 10 cover estimates and projections of the population by age, sex and race, paying particular attention to the sub-national level. Chapter 11 focuses on estimates and projections of households, labor force, school enrolment, educational attainment and health. All three chapters discuss a wide range of methods and include lengthy sections on the evaluation of accuracy and utility. Each chapter comes with a categorized list of suggested readings.

Siegel, J., and C. H. Hamilton. 1952. "Some Considerations in the use of the Residual Method of Estimating Net Migration." *Journal of the American Statistical Association* 47 (259): 475-599. The authors compare Census net migration data with estimates of net migration obtained by the residual method, representing the difference between total population change and natural change during a period, and discuss some general problems in the use of the residual method . Several residual methods - the vital statistics method and the forward, reverse, and average survival rate methods- are described, compared, and evaluated. On the basis of a symbolic model representing the population in an age group in terms of migration cohorts, the authors show how the various survival rate formulas, unlike the vital statistics method, fail to make an accurate allowance for the net migration of persons who die during the migration period, except under very restricted conditions of migration. They then develop maximum theoretical errors in the use of the various survival rate formulas, resulting from the inability of survival rates to measure deaths occurring in an area exactly, and the theoretical errors under different conditions of timing of migration, and make suggestions on this how the problem may be handled.

Siegel, J., H. Shryock, and B. Greenberg. 1954. "Accuracy of Post-censal Estimates of Population for States and Cities." *American Sociological Review* 19: 440-446. The authors note that census survival rates, which measure inter-censal state and county net migration, by age, color, and sex, have certain technical advantages over other residual methods, specifically the exclusive virtue of a built-in technique which corrects for net census undercount by age, including under-enumeration and misstatements of age. Using three sets of estimated "closed" native populations and reported United States deaths as standards, they evaluate three sets of national census survival rates, but note that the unavailability of appropriate figures make it impossible to determine precisely which is most accurate. They analyze the effects of adjusting census survival rates for national-state mortality differentials in measuring inter-censal net migration and deaths for states. Next, they examine differences in state net migration and mortality arising from the use of the forward, reverse, and average formulas. They find insignificant differences between adjusted and unadjusted survival rates and among the forward, reverse, and average formulas. Finally, they provide estimates of net interstate migration for native whites, total whites, and nonwhites, by age and sex, for the 1940-50 decade, using census survival rates, and compare the results with the more correct estimates for all ages obtained with vital statistics data. By using the formula that most accurately duplicates "vital statistics" deaths over all ages, they argue that one can obtain the most precise inter-censal state net migration estimates by age.

Simpson, S. 1999. "Population Estimates for Small Areas in the U.K: Performance and Promise." Paper presented the Estimates Methods Conference, U. S. Bureau of the Census, Suitland, MD. (available online, http://www.census.gov/population/www/coop/popconf/paper.html, last accessed June 2007). This paper reports on "The Estimating with Confidence" project 1993-1996 which evaluated the accuracy of existing methods to produce population estimates for areas with 1,000 -30,000 people. A mid-term enumeration- a local census - was found to be the most accurate, but a number of cheaper desk-top methods employing

administrative data gave reasonable results in most areas. The accuracy of these methods has been tabulated and attributed to characteristics of the areas estimated and the estimation strategies employed.

Simpson, S., I. Diamond, P. Tonkin, and R. Tye. 1996. “Updating Small Area Population Estimates in England and Wales.” *Journal of the Royal Statistical Society, Series A*, 159 (part 2): 235–247. The authors state by observing that population estimates have important implications for resource allocation within government and commerce, and are often assumed to be without error. They note that currently, central government provides annual population estimates for all of the local and health authority districts in Britain, but estimates are needed for smaller areas, typically for electoral wards and postal sectors. Small area estimates are provided by some local authorities and commercial organizations, using different methods; the accuracy of these estimates is modeled by the authors within a multilevel framework. Certain characteristics of the small area and of the method of estimation are included as explanatory variables. Results show that the method of estimation used is of great importance.

Simpson, S., R. Cossey, and I. Diamond. 1997 “1991 Population Estimates for Areas Smaller than Districts.” *Population Trends* 90: 31-39. This article describes the construction of population estimates for mid-1991 for electoral wards in England and Wales and postal sectors in Scotland. It shows how earlier work adjusting 1991 census figures at national and local authority level for undercount and other factors has been extended to smaller areas in a way that produces estimates which are consistent with the estimates for larger areas. Estimates for smaller areas are needed to calculate employment, health and other indices, and as a starting point for population estimates between census years.

Simpson, S., E. Middleton, I. Diamond, and D. Lunn. 1997. “Small-area Population Estimates: A Review of Methods used in Britain in the 1990s.” *International Journal of Population Geography* 3 (3): 265-280. The authors note that population estimates are usually produced by local government administrations in Britain, for each small area within their authority. Increasing interest has been shown by commerce and by central government. They identify five main methods: apportionment, ratio change, additive change, cohort survival, and local censuses. They also note that estimation strategies also vary according to available data, the detail in which a population is estimated, and the precise combination of elements chosen from one or more of the main methods. They survey the use of methods at the beginning of the 1990s and provide examples. The author conclude by observing that the accuracy of each main method is quantified from empirical data collated by the ‘Estimating with Confidence’ project and provide a review of likely developments towards the end of the 1990s.

Singh, M., M. Hidiroglou, J. Gambino and M. Kovac□ević. 2001. “Estimation Methods and Related Systems at Statistics Canada.” *International Statistical Review* 69 (3): 461-485. This paper provides an overview of research in estimation techniques, their application, and the development of generalized estimation systems at Statistics Canada. The paper briefly discusses estimation for longitudinal data and a weighting approach

developed for cross-sectional data from these surveys. For cross-sectional household and business surveys, as well as the census of population, appropriate calibration estimators developed for each situation are briefly discussed. In addition, regression composite estimation, a method developed to improve the quality of cross-sectional estimates from rotating panel surveys such as the Canadian Labor Force Survey, is presented. With regard to more detailed cross-sectional estimates at sub-provincial levels, different approaches to small area estimation developed for various programs are also presented. The paper summarizes the various modules developed for the Generalized Estimation System and includes a discussion of new developments within the system such as two-phase estimation as well as the estimation of variance for a number of imputation procedures. The paper concludes with a brief review the status of current estimation research on selected topics and a discussion on the direction of future research.

Sink, L. 2000. "Estimates of the Population of Counties by Age, Sex, Race, and Hispanic Origin: 1990-1999." Unpublished report. Washington, DC. U.S. Census Bureau. (Available online, http://www.census.gov/popest/archives/methodology/90s-co-meth.txt, last accessed, May 2008). This is a description of the methods and data used to generate these estimates.

Sink, L. 1996. "Estimates of the Population of Counties by Age, Sex, Race, and Hispanic Origin: 1990-1994." *Release PE-48 (Methodology).* Washington, DC: U.S. Census Bureau. This is a description of the methods and data used to generate these estimates.

Sirken, M., M. Graubard, and R. La Valley. 1978. "Evaluation of Census Population Coverage by Network Surveys." pp. 239-244 in *Proceedings of the Section on Survey Research Methods*. Alexandria, VA: American Statistical Association. The authors investigate the use of dual system estimators in the context of two Post-enumeration survey methods, the Post-enumeration survey (PES) and the post-enumeration multiplicity survey (PEMS) that were being tested for use with the 1980 census. They note that except for the counting rule, the design features of both surveys are virtually the same. The PES adopts a de jure residence rule, and the PEMS adopts a multiplicity counting rule. The de jure residence rule specifies that people are eligible to be enumerated only at their usual places of residence. On the other hand, the multiplicity counting rule adopted by PEMS specifies that people are eligible to be enumerated at the households of specified close relatives as well as at their own de jure residences. Dual system estimators are investigated under both approaches with a focus on when PEMS would have a smaller correlation bias than PES.

Sivamurthy, M. 1969. "Errors in the Estimation of Net Migration Rate in the Studies of Internal Migration." *Journal of the American Statistical Association* 64 (328): 1434-1438. The author examines the effect of relaxing the conditions regarding the equality of the mortality rates and of the errors of enumeration between the states and the nation, on the error in the estimation of the net migration rate by the census survival ratio method. He shows that when the national population is closed, what the necessary and sufficient conditions are for this error to become zero. He does this by finding that the relative degree of enumeration for any age x of a given sex in the i^{th} state, as compared with the degree of enumeration in the nation at the first census, bears the same ratio to the relative degree of enumeration for the same cohort at age (x + n) at the second census, as the survival ratio of that cohort from age x to (x + n) in the i^{th} state bears to the corresponding survival ratio in the nation as a whole.

Smith, D. 1992. *Formal Demography*. New York City, NY: Plenum Press. This book is intended as a relatively non-technical introduction to current demographic methods. This book begins with an overview of demographic concepts and measures, including population pyramids and the Lexis diagram, to introduce readers to usual population configurations. Chapter 2 reviews data adjustment techniques that are widely used in demography, and includes elementary formulas for curve fitting, osculatory interpolation, and a selection of parametric distributions which find applications in fertility analysis. The chapter also introduces integral and derivative fittings for polynomial distributions, used in conjunction with the life table. Data adjustment by direct and indirect standardization is treated separately in Chapter 3. Chapters 4-6 focus on life table methodology. The later chapters of the book discuss fertility analysis, population projections and migration, and stable population theory.

Smith, S. K. 1996. "Demography of Disaster: Population Estimates after Hurricane Andrew." *Population Research and Policy Review* 15 (5-6): 459-477. This article describes the estimation problems created by Hurricane Andrew and how those problems were resolved through the use of existing data sources and the collection of new types of data. It concludes with a discussion of several conceptual, methodological and procedural issues that will have to be faced in virtually any attempt to estimate the demographic consequences of natural disasters.

Smith, S. K. 1994. "Estimating Temporary Populations: The Contributions of Robert C. Schmitt." *Applied Demography* 9 (1): 4-7. This paper reviews the contributions to methods and data sources made by R. Schmitt in estimating temporary (De facto) populations over his career as the Hawaii State Statistician.

Smith, S. K. 1989. "Toward a Methodology for Estimating Temporary Residents." *Journal of the American Statistical Association* 84: 430-436. The author notes that most population statistics for states, counties, and cities refer to permanent residents, or persons who spend most of their time in an area, but that at certain times, many states and local areas have large numbers of temporary residents who exert a significant impact on the area's economy, physical environment, and quality of life. He notes that typically, very little is known about the number, timing, and characteristics of these temporary residents. The author discusses the problems of defining and estimating temporary residents, focusing on the strengths and weaknesses of a number of data sources and estimation techniques. He closes with an assessment of the potential usefulness of developing methods to estimate temporary residents.

Smith, S. K., and M. House. 2007. "Temporary Migration: a Case Study of Florida." *Population Research and Policy Review* 26 (4): 437-454. Using several types of survey data, the authors examine the characteristics of non-Floridians who spend part of the year

in Florida and Floridians who spend part of the year elsewhere. They develop estimates of the number, timing, and duration of temporary moves and the origins, destinations, characteristics, and motivations of temporary migrants. This study presents the most comprehensive analysis yet of temporary migration in Florida and provides a model that can be used in other places. It also points to a serious shortcoming in the US statistical system, namely, the lack of information on temporary migration streams, which might be addressed via the American Community Survey in the future.

Smith, S. K., and P. Morrison. 2005. "Small-Area and Business Demography." pp. 761-786 in M. Micklin and D. Poston, Jr. (Eds.) *Handbook of Population*. New York City, NY: Kluwer Academic/Plenum. In providing an overview of the subject matter, the authors briefly discuss estimation and forecasting method.

Smith, S. K., and S. Cody. 2004. "An Evaluation of Population Estimates in Florida: April 1, 2000." *Population Research and Policy Review* 23(1): 1–24. The authors investigate the influence of differences in population size and growth rate on estimation errors; compare the accuracy of several alternative techniques for estimating each of the major components of the Housing Unit Method (HUM); compare the accuracy of 2000 estimates with that of estimates produced in 1980 and 1990; compare the accuracy of HUM population estimates with that of estimates derived from other estimation methods; consider the role of professional judgment and the use of averaging in the construction of population estimates; and explore the impact of controlling one set of estimates to another. Their results confirm a number of findings that have been reported before and provide empirical evidence on several issues that have received little attention in the literature. They conclude with several observations regarding future directions in population estimation research.

Smith, S. K., and M. Mandell. 1984. "A Comparison of Population Estimation Methods: Housing Unit versus Component II, Ratio-Correlation, and Administrative Records." *Journal of the American Statistical Association* 79 (386): 282-289. The authors observe that the Housing Unit Method (HUM) is often characterized as inferior to other methods for estimating the population of states and local areas. They challenge this characterization and in this paper evaluate population estimates produced by the housing unit method and by three other commonly used methods: component II, ratio correlation, and administrative records. Basing their analysis on 1980 census data from 67 counties in Florida and testing for precision, bias, and the distribution of errors, the authors find that their application of the HUM performs at least as well as the more highly acclaimed methods of local population estimation.

Smith, S. K., and C. McCarty. 1996. "Demographic Effects of Natural Disasters: A Case Study of Hurricane Andrew." *Demography* 33 (2): 265-275. The authors use information from a sample survey and other sources to evaluate the effects of Hurricane Andrew on the housing stock and population of Dade County, Florida. They estimate that more than half of the homes were damaged and that 330,000 people left their homes, at least temporarily, and that 40,000 left permanently.

Smith, S. K., and J. Nogle. 2004. "An Evaluation of Hispanic Population Estimates." *Social Science Quarterly* 85 (3): 731-745. Using a variety of techniques, the authors develop Hispanic population estimates for counties in Florida and evaluate the accuracy of those estimates by comparing them with 2000 Census counts. They find that Hispanic population estimates have larger errors than estimates of total population; errors vary considerably by population size and growth rate; some techniques perform better than others in places with particular population characteristics; and averages often perform better than individual techniques. They conclude that in many circumstances, symptomatic data series can provide more accurate estimates of the Hispanic population than more commonly used techniques.

Smith, S. K., and J. Nogle. 1997. "An Experimental Methodology for Estimating Hispanic Residents for States and Counties." *Journal of Social and Economic Measurement* 23 (2) 263 – 275. In this article, the authors describe an experimental methodology for estimating the Hispanic population of states and counties. They use post-censal data on births, deaths, and school enrollment for estimates of the total Hispanic population and data from the two most recent decennial censuses for estimates of the age, sex, and race distribution of that population. They discuss the strengths and weaknesses of this methodology and illustrate its application by making estimates of the Hispanic population for counties in Florida.

Smith, S. K., and S. Cody.1994. "Evaluating the Housing Unit Method: A Case Study of 1990 Population Estimates in Florida." *Journal of the American Planning Association*. 60: 209-221. The authors observe that the housing unit method (HUM) is the most commonly used approach to making small-area population estimates in the US. They evaluate the accuracy and bias of HUM-based population estimates produced for counties and sub-county areas in Florida for April 1, 1990. They find that population size has a negative effect on estimation errors (disregarding sign) but no effect on bias; growth rates have a U-shaped effect on estimation errors (disregarding sign) and a negative effect on bias; electric customer data provide more accurate household estimates than building permit data; errors in household estimates contribute more to population estimation error than do errors in estimates of average household size or group quarters population; and the application of professional judgment improves the accuracy of purely mechanical techniques. The authors argue the HUM offers a number of advantages over other population estimation methods and provides planners and demographers with a powerful tool for small-area analysis.

Smith, S. K., and M. House. 2007. "Temporary Migration: a Case Study of Florida." *Population Research and Policy Review* 26 (4): 437-454. In this paper, the authors analyze temporary migration streams in Florida, focusing on moves that include an extended stay. Using several types of survey data, they examine the characteristics of non-Floridians who spend part of the year in Florida and Floridians who spend part of the year elsewhere. They develop estimates of the number, timing, and duration of temporary moves and the origins, destinations, characteristics, and motivations of temporary migrants. They argue that this study presents the most comprehensive analysis yet of temporary migration in Florida and provides a model that can be used in other places.

They also point to a serious shortcoming in the US statistical system, namely, the lack of information on temporary migration streams. They observe that the American Community Survey provides an opportunity to remedy this problem.

Smith, S. K., and B. Lewis. 1983. "Some New Techniques for Applying the Housing Unit Method of Local Population Estimation: Further Evidence." *Demography* 20 (3): 407-413. The authors note that the housing unit method (HUM) of population estimation is often characterized as being imprecise and having an upward bias. In an earlier paper, the authors argued that the method itself cannot be properly characterized by a particular level of precision or direction of bias. Only specific techniques of applying the new method can have such characteristics. In that paper, the authors present several new techniques for estimating the number of households and average number of persons per household (PPH). However, the testing of these new techniques was limited by the lack of census results against which the estimates could be compared. Complete census data on population, households, and PPH are available and can be used to test alternate estimation techniques. The tests reported in the earlier paper using 1980 census data for Florida's 67 counties are replicated and the authors argue that this provide further evidence that the new techniques produce more precise, less biased estimates than previously used techniques.

Smith, S. K., and B. Lewis. 1980. "Some New Techniques for Applying the Housing Unit Method of Local Population Estimation." *Demography* 17 (3): 323-339. The authors argue that the housing unit method of population estimation is often characterized as being imprecise and having an upward bias. They believe that the method itself cannot properly be characterized by a particular level of precision or direction of bias. Only specific techniques of applying the method can have such characteristics. In this paper, the authors discuss several new techniques they have developed for estimating households and the average number of persons per household. The compare estimates produced by these techniques to estimates produced by several other techniques and find that special census results from Florida provide preliminary evidence that the new techniques produce more precise, less biased estimates than the other techniques.

Smith, S. K., and D. Swanson. 1998 "In Defense of the Net Migrant." *Journal of Economic and Social Measurement.* 24 (3-4): 155-170. This paper discusses the benefits and weaknesses of net migration as a theoretical concept and as a measure of population movement in the United States. It notes that net migration is useful in population estimation, forecasting, and analysis, but that it is faulted for not being as true a measure of population movement as gross migration because net migration is not a process itself, but the difference between two processes. The authors note that this same characterization could apply to the concept of "profit," the difference between gross income and gross expenditures. They go on to note that like profit, net migration is a useful concept even if it is not a "true' measure of population movement.

Smith, S. K., J. Nogle, and S. Cody. 2002. "A Regression Approach to Estimating the Average Number of Persons per Household." *Demography* 39(4): 697-712. The authors develop several regression models in which PPH estimates were based on symptomatic indicators of PPH change. They tested these estimates using county level data in four states and found them to be more precise and less biased than estimates based on more commonly used methods.

Smith, S. K., J. Tayman and D. Swanson. 2001. *State and Local Population Projections: Methodology and Analysis.* New York: Kluwer Academic/Plenum Publishers. This book focuses on the methodology and analysis of state and local population projections. It describes the most commonly used data sources and application techniques within each of three classes of projection methods (cohort-component, trend extrapolation, and structural models) and covers the components of population growth, the formation of assumptions, the development of evaluation criteria, and the determinants of forecast accuracy. It considers the strengths and weaknesses of various projection methods, paying special attention to the unique problems of making projections for small areas, and closes with an examination of technological and methodological changes affecting the production of small-area population projections.

Snow, E.C. 1911. "The application of the method of multiple correlation to the estimation of post-censal populations." *Journal of the Royal Statistical Society* 74 (part 6): 575-620. This paper represents the first published description of the use of multiple correlation (regression) in the estimation of population. It discusses a number of other methods (e.g., "The American Method"), pointing out their strengths and weaknesses, then describes the model framework and the data used, and applies the method to districts in the U. K., noting a "false start" from a first attempt and how it was revised.

Spar, M. and J. Martin. 1979. "Refinements to regression-based estimates of post-censal population characteristics." *Review of Public Data Use* 7: 16-22. The authors find that the ratio-correlation method is more accurate than others in estimating the populations of Virginia counties by race and age.

Spencer, B., and C. Lee. 1980. "Post-censal population estimation methods of the Census Bureau." pp. 131-187 in *Estimating Population and Income of Small Areas*. Panel on Small-Area Estimates of Population and Income, Committee on National Statistics, National Research Council. Washington, DC: National Academy Press. This paper described the methods used by the U.S. Census Bureau to estimate populations at the national and sub-national levels.

Starsinic, D. 1973. "Development of Populatiion Estimates for Revenue Sharing Areas." pp. 4- 7 in *Statistical Methodology of Revenue Sharing and Related Estimate Studies*, *Census Tract Papers Series GE-40, No. 10.* U.S. Bureau of the Census. Washington, DC: U.S. Bureau of the Census. The author outlines the challenges facing the Census Bureau in regard to developing a small area estimation system to be used with revenue sharing. He describes the experimental approach being examined by the Census Bureau to meet these challenges. It consists of using vital statistics and an estimate of migration based on IRS files. The author also describes various adjustments needed to make the system function.

Starsinic, D., and R. Forstall. 1989. "Patterns of Metropolitan Area and County Population Growth: 1980 to 1987." *Current Population Reports, P-25, No. 1039.* Washington, DC: US Government Printing Office. This report contains population estimates for July 1, 1987, revised estimates for July 1, 1986, and components of population change from 1980 to 1987 for individual metropolitan statistical areas (MSA's) and their component counties....The report also contains census counts for April 1, 1980, annual summary estimates since 1980 for the metropolitan and non-metropolitan population by State, a discussion of trends since 1980, and comparisons with the 1970's and 1960's. It also shows 1986 population estimates for MSA central cities. Summaries for the United States, regions, divisions, and States are presented for the 1987 population inside and outside metropolitan areas and for the 1986 population inside and outside central cities. Selected tables also include data for groupings of metropolitan areas, such as the Northeastern Megalopolis, for cities of 100,000 and over, and for counties.

Starsinic, D., and M. Zitter. 1968. "Accuracy of the Housing Unit Method in Preparing Population Estimates for Cities." *Demography* 5: 475-484. This paper reports the results of a test of the relative accuracy of the housing unit method in the estimation of the population of cities. Estimates were prepared for 47 cities in excess of 50,000 population in which special censuses were conducted during the years 1964-66. The test points up five features of the housing unit method for the estimation of the population of cities. (1) In general, the method yields estimates on the high side. When building-permit data were used as a basis for estimates, the deviations were positive in about 30 of the 47 cases. Of the cities in which the deviations were negative, about one-half had had substantial annexation after 1960. (2) The use of utility data instead of building-permit data generally reduces the size of errors, although here too there are substantially more positive than negative deviations. (3) Deviations are smaller when the average size of household is extrapolated than when the 1960 values are used. (4) When either building-permit or utility data were used, the average error in the estimate of the number of households was high. (5) The estimate of the number of households is a greater contributor to errors in estimates of population than is the estimate of the average size of household (for this test, extrapolated from 1950-60 values or assuming no change since 1960). Although the scope of the test was limited, the method made a relatively creditable showing, with average errors of 3.6 to 5.8 percent, excluding areas that are experiencing large annexations.

Starsinic, D., A. Lee, H. Goldsmith, and M. Spar. 1995. "The Census Bureau's Administrative Records Method." pp. 54-69 in N. W. Rives, W. Serow, A. Lee, H. Goldsmith, and P. Voss (Eds). *Basic Methods for Preparing Small-Area Population Estimate*s. Madison, WI: Applied Population Laboratory, University of Wisconsin. The authors begin this chapter by noting that the so-called Administrative Records Method (ARM) s similar to Component Method II (CM II) in that it is a component-based method. They note that the two methods differ in their approach to estimating the migration component in that the ARM uses tax returns and CM II uses school enrollment. They describe the details of the ARM and illustrate its use.

Starsinic, M. 2003. “Small Area Modeling for the American Community Survey.” Paper presented at the 2003 American Statistical Association Conference, August 3–7. The author investigates whether small area estimation techniques can improve the overall estimates and estimates of specific characteristics for counties and census tracts. The author uses EBLUP and hierarchical Bayes estimators and states that by analyzing data from the 2000 American Community Survey Comparison Test, he will be able to compare his estimators against "true" population values from Census 2000.

Statistics Canada. 1987. *Population Estimation Methods, Canada.* Ottawa, Ontario, Canada: Ministry of Supply and Services. The methods used to produce official population estimates for Canada are discussed and evaluated in this monograph. The chapters are organized around three recurring themes: (1) method; (2) data sources, and; (3) quality evaluation so that an attempt is made in each chapter to answer the corresponding questions: (1) how are the estimates produced; (2) what are the data and their sources, and; (3) how reliable are the estimates in light of various validation criteria?" Separate chapters are included on post-censal population estimates for the total population and for the population by sex, age, and marital status, inter-censal estimates, internal migration, emigration, local area population, and family estimates.

Statistics Finland. 2004. *Use of Registers and Administrative Data Sources for Statistical Purposes.: Best practices of Statistics Finland*. Helsinki, Finland: Statistics Finland. This monograph documents the development, workings, and uses of the Finnish Statistical System, including the generation of population data.

Stone, G. 2008. “Methods for Measuring the Population After a Disaster: Household Population Surveys in Post-Katrina New Orleans, October 2005 – October, 2006.” pp. 113-123 in S. Murdock and D. Swanson (Eds.). *Applied Demography in the 21st Century*. Dordrecht, The Netherlands, Springer. The author describes the use of three difference sampling designs to conduct household surveys in Post-Katrina New Orleans. He discusses their limitations, cost and time requirements and the usefulness of the estimates they produced.

Summers, A. and B. Wolfe. 1978. “Estimation of Household Income from Location.” *Journal of the American Statistical Association* 73: 288-292. The authors describe a Block Income Estimating Procedure (BIEP) for estimating the average income of the households of a Census block, using block housing information and Census tract distributions relating to housing and household income. They observe that BIEP also can be used to estimate the average income of geographic areas that do not coincide with Census tracts or postal zip code areas, the smallest spatial units for which official data are available. BIEP's efficacy is tested using a special Census tabulation for Philadelphia school feeder areas, and it is compared with its closest rival, a tracts-average procedure. The authors find that BIEP works well relative to the other procedure.

Sutton, G. and G, Danziger. 1989. "Are Accuracy and Quality the Same Thing in Projections of Local Area Populations?" pp. 186-191 in 1989 *Proceedings of the*

American Statistical Association, Social Statistics Section. Alexandria, VA: American Statistical Association. This paper offers a method that the authors believe will increase the acceptability of population projections as forecasts by drawing on ecological models of growth and by calling forth theories of age and family life cycle patterns in residential preferences. Sum results from the study of the differences in the shape of age distributions among communities are presented. The method is illustrated with data for Massachusetts from 1970 to 1980.

Swanson, D. 2008. "Measuring Uncertainty in Population Data Generated By the Cohort-Component Method: A Report on Research in Progress." pp. 165- 189 in S. Murdock and D. Swanson (Eds.). *Applied Demography in the 21st Century*. Dordrecht, The Netherlands: Springer. Building on earlier work, the author provides an ex post facto test of a procedure for generating a formal measure of uncertainty for short-term population forecasts made sing the cohort-component method using a 1990 forecast launch date and 2000 census data as the benchmark for a small population in a Nevada County. The procedure uses a Mean-Square Error Confidence Interval that is subject to limitations. The author finds that the technique shows promise, but needs more work to be useful.

Swanson, D. 2008. "The Demographic Effects of Hurricane Katrina on the Mississippi Gulf Coast: An Analysis by Zip Code." Presented at the Conference of the Mississippi Academy of Sciences, 20-22 February, Olive Branch, Mississippi. This paper provides an estimate of the effects of Hurricane Katrina on the population of 20 selected zip code areas in Hancock, Harrison and Jackson counties, Mississippi, that were at or near the epicenter of Hurricane Katrina. The effects are examined by using 1990 and 2000 census data, information from a special data collection funded by the National Science Foundation, and special county-level "Katrina impact" 2006 population estimates prepared by the U. S. Census Bureau. The Cohort Change Ratio Method is applied to 1990 and 2000 census data to generate 2007 population estimates in the absence of Katrina. These estimates are then adjusted to take Katrina's effects into account. By comparing the adjusted to the unadjusted estimates an idea of the absolute and relative impact of Katrina is gained. The comparison suggests that Katrina's demographic effects are profound and not only likely to affect the 2010 census counts in these areas, but that they may persist well beyond. Given the long-lasting demographic effects of such disasters, the author suggests that these methods be used in the future and provide specific recommendations on how this can be accomplished.

Swanson, D. 2006. Statement before the Subcommittee on Federalism and the Census Oversight Hearing Committee on Government Reform, U.S. House of Representatives Hearing Topic: "Two Plus Two Should Never Equal Three: Getting Intercensal Population Estimates Right the First Time." Wednesday, 6 September 2006, Room 2247 Rayburn House Office Building. Swanson's testimony covers three areas in regard to the Census Bureau's estimates program: (1) the major challenge faced by the Census Bureau in providing timely, accurate, and cost-effective estimates; (2) A suggestion for dealing with this challenge; and (3) Issues presented by his suggestion that need to be resolved. Swanson identifies the major challenge facing the estimates program as the atmosphere of conflict between the Bureau and many of its users, particularly state and local

governments who receive federal funds tied to population estimates. He suggests that a way to overcome this atmosphere is exploit the MAF and use the HUM ins a nationwide system of state demographic centers that participates in a meaningful partnership with the Census Bureau. He also suggests that state demographic centers, in turn, need an active and meaningful partnership with the local governments within their respective states. He identifies three major issues that need to be dealt with for this to work, confidentiality, costs, and technical issues inherent in the MAF/TIGER system, but observes that they are tractable problems. He therefore proposes that the MAF be more fully exploited by using the HUM as a universal means of population estimation for all areas of geography, administrative and statistical, and that state demographic centers be developed to a uniform level of capability. He suggests that this proposal be supported by state-federal matching funds and argues that this would lead not only to timely, accurate and cost-effective inter-censal population estimates, but also to greater equity in that there would be a uniformly higher level of demographic human capital in the country.

Swanson, D. 2004. “Advancing Methodological Knowledge within State and Local Demography: A Case Study.” *Population Research and Policy Review* 23 (4): 379-398. This paper examines a regression model developed in Nevada following the 2000 census that led to conflict over its use to estimate the population of Clark County, Nevada in 2002. The discussion reveals statistical and methodological shortcomings in this model that lead to an alternative model not subject to these shortcomings. This example illustrates how this type of analysis and discussion can lead to a wider understanding of methods on the part of practitioners through the corrective process of academic peer review. It also suggests that states in which estimates are used to allocate resources would be well-served by subjecting new methods being considered for use to academic peer review before they are adopted.

Swanson, D. 1989. “Confidence Intervals for Post-censal Population Estimates: A Case Study for Local Areas.” *Survey Methodology* 15: 271-280. The author shows how confidence intervals can developed for population estimates using the ratio-correlation method (and its variants), illustrates the proposed procedure using county data from Washington State and shows that it works well in an ex post facto test using 198o census data as a benchmark.

Swanson, D. 1986. “Evaluating Population Estimates and Short-Term Forecasts.” *Applied Demography* 2 (November): 5-6. The author provides some general guidelines for evaluating population estimates and short-term population forecasts and points out that some of the methods typically used in developing estimates can be used for developing short-term projections and that virtually all of the methods typically used in developing short-term forecasts can be used for estimates.

Swanson, D. 1981. “Allocation Accuracy in Population Estimates: An Overlooked Criterion With Fiscal Implications.” pp. 13-21 in *Small Area Population Estimates and Their Accuracy, Series GE-41, No. 7.* U. S. Bureau of the Census. Washington, DC: US Government Printing Office. The author argues that because population estimates are often used to distribute resources in a zero-sum fashion, an important component in

measuring the performance of estimation methods is missing, namely allocation error. He proposes using the Index of Dissimilarity to overcome this problem and illustrates its use with Washington State data.

Swanson, D. 1980. "Improving Accuracy in Multiple Regression Estimates of Population Using Principles from Causal Modeling." *Demography* 17: 413-428. This paper reports a mildly restricted procedure for using a theoretical causal ordering and principles from path analysis to provide a basis for modifying regression coefficients in order to improve the estimation accuracy of the ratio-correlation method of population estimation. The modification is intended to take into account temporal changes in the structure of variable relationships, a major element in determining the accuracy of post-censal estimates. The modification of coefficients is conservative in that it uses rank-ordering as a basis of change. Empirical results are reported for counties in Washington State that demonstrate the increased accuracy obtained using the proposed procedure.

Swanson, D. 1978. "An Evaluation of Ratio and Difference Regression Methods for Estimating Small, Highly Concentrated Populations: The Case of Ethnic Groups." *Review of Public Data Use* 6: 18-27. Using 1970 census data as a benchmark, this paper compares the accuracy of the ratio-correlation and difference-correlations methods for estimating the population by race in Washington State's 39 counties. The models are constructed using 1950 and 1960 data. The difference-correlation model shows much better characteristics (e.g., a higher r^2), but its accuracy is not shown in the test to be any higher than that of the ratio-correlation model. The paper notes that compared to difference-correlation method, the ratio-correlation method cannot deal well with sparse data where zero counts exist because a ratio with a zero in it is undefined while a difference with a zero leads to an actual value. Hence, the poor characteristics of the ratio-correlation model. The paper concludes that this is a case where model characteristics can be misleading since the ratio-correlation model performs as well as the difference-correlation model in the ex post facto test.

Swanson, D., and G. Hough. 2007. "An Evaluation of Persons Per Household (PPH) Data Generated By the American Community Survey: A Demographic Perspective." Paper presented at the Annual Meeting of the Southern Demographic Association. Birmingham, AL. This paper explores the usability of ACS data by examining "Persons Per Household (PPH), a variable of high interest to demographers and others preparing regular post-censal population estimates. The data used in this exploration are taken from 18 of the counties that formed the set of 1999 ACS test sites. The examination proceeds by comparing ACS PPH values to PPH values generated using a geometric model based on PPH change between the 1990 and 2000 census counts. The ACS PPH values represent what could be called the "statistical perspective" because variations in the values of specific variables over time and space are viewed largely by statisticians with an eye toward sample (and non-sample) error. The model-based PPH values represent a "demographic perspective" because PPH values are largely viewed by demographers as varying systematically, an orientation stemming from theory. The results suggest that the ACS PPH values lack sufficient temporal consistency to be used by demographers.

Swanson, D., and A. Al-Jiboury. 1988. "Inter-censal Net Migration Among the Three Major Regions of Iraq: 1957-1977." *Janasamkhya* 6 (December): 93-126. The authors generate total net migration estimates for the Northern, Central, and Southern regions in Iraq between 1957 and 1977. This was accomplished by applying the Reverse Survival Ratio Estimate Method (RSRE). Since the RSRE method has not been used in a nation such as Iraq, which lacks complete census and vital statistics information, it was tested using complete population data from the Mid-Atlantic Division of the United States. The test results indicated a reasonable level of accuracy in determining both the direction and the volume of the total net migration estimates over the 1960-70 period, given the authors' assumptions. The authors conclude that this method of estimation is useful for economic and social development planning in countries with incomplete census data.

Swanson, D., and J. McKibben. 2007. "New Directions in the Development of Population Estimates and Projections." Paper presented at the Satellite Conference on Small Area Statistics, 56th Conference of the International Statistical Institute, Pisa, Italy. The authors argue that a continuously updated Master Address File argue that the MAF can be uses to develop population estimates and projections if it is properly enhanced (EMAF). The authors describe a set of activities needed to develop EMAF and how EMAF data could be directly assessed for statistical uncertainty and its use as a basis for developing population projections containing a wide range of ascribed (e.g., age and sex) and achieved characteristics (e.g., educational attainment and employment). The authors point out that such a development would bring the US Census Bureau's small area population estimation programs more in line with its European counterparts, but that there are several important challenges that must be surmounted, including issues of public trust, confidentiality, and tradition.

Swanson, D., and L. Pol. 2008. "Applied Demography: Its Business and Public Sector Components." in *Demography* (Ed. Zeng Yi), *Encyclopedia of Life Support Systems (EOLSS*). Developed under the auspices of UNESCO, EOLSS Publishers, Oxford, UK (http://www.eolss.net, last accessed February, 2008). This paper contains a description of the methods and data used for population estimates and projections as well as a bibliography.

Swanson, D., and L. Pol. 2005. "Contemporary Developments in Applied Demography within the United States. *Journal of Applied Sociology* 21 (2): 26-56. In addition to covering contemporary developments in the field of applied demography and some ideas about its future, the paper describes the methods and data used in population estimation and projection.

Swanson, D., and J. Wicks (Eds). 1987. *Issues in Applied Demography: Proceedings of the 1986 National Conference.* Bowling Green, OH: PSRC Press. These are the proceedings of the 1986 Conference on Applied Demography. The focus of the conference was on ways in which demographic data and concepts can be used to produce better and more informed business and public policy decisions. Subject headings include applied demography and government policy; technical issues in business demography; relations between public sector and academic demographers; demography and management issues in business; and linking the demographic information needs of business, state and local governments, and universities.

Swanson, D., and D. Beck. 1994. "A New Short-term County Population Projection Method." *Journal of Economic and Social Measurement* 20: 25-50. This paper proposes a new method for short-term county population projections. It is based on a modification of the ratio-correlation method of population estimation. The modified ratio-correlation method can produce projections with a high potential for accuracy without requiring substantial data and intensive intellectual labor inputs. Tests of accuracy are examined for the modified ratio-correlation method and two currently available alternatives using data from Washington State. The authors find that the tests suggest that the new method performs well.

Swanson, D., and L. Tedrow. 1984. "Improving the Measurement of Temporal Change in Regression Models used for County Population Estimates." *Demography* 21: 361-372. The ratio-correlation method of population estimation is shown to contain an inconsistent temporal relationship between the model's empirical structure and its actual application. A simple transformation of the model's variables is provided that eliminates the inconsistency. Two tests of the relative accuracy of the original and transformed models show that the transformed model achieves accuracy levels equal to or higher than the original. In one test, all nine years show a higher degree of accuracy, of which four are statistically significant. Several possible reasons are given for the increased accuracy shown by the transformed model. The transformation, termed the "rate-correlation" model, is recommended as a logical starting point in the examination of coefficient stability and spatial autocorrelation as well as a method for estimating small populations.

Swanson, D., T. Burch, and L. Tedrow. 1996. "What is Applied Demography?" *Population Research and Policy Review* 15 (5-6): 403-418. The authors argue that applied demography is intrinsically distinct from basic demography because it exhibits the value-orientation and empirical characteristics of a decision-making science while the latter exhibits the value-orientation and empirical hallmarks of a basic science. Distinguishing characteristics of applied demography are based on the context in which it places precision and explanatory power relative to time and resources as well as the fact [that] its substantive problems are largely exogenously-defined, usually by customers. The substantive problems of basic demography, on the other hand, are largely endogenously-defined, usually by academic demographers. The authors examine this conceptualization of applied demography in terms of the methods and materials that fall within its purview and discuss some important consequences, including research agendas and training programs.

Swanson, D., J. Tayman, and C. Barr. 2000. "A Note on the Measurement of Accuracy for Sub-national Demographic Estimates." *Demography* 37: 193-201. Mean absolute percentage error (MAPE), the measure most often used for evaluating sub-national demographic estimates, is not always valid. The authors describe guidelines for determining when MAPE is valid. Applying them to case data, the authors find that MAPE understates accuracy because it is unduly influenced by outliers. To overcome this problem, the authors calculate a transformed MAPE (MAPE-T) using a modified Box-Cox method. Because MAPE-T is not in the same scale as the untransformed absolute percentage errors, the authors provide a procedure for calculating MAPE-R, a measure in the same scale as the original observations. The authors argue that MAPE-R is a more appropriate summary measure of average absolute percentage error when the guidelines indicate that MAPE is not valid.

Swanson, D., J. Tayman, and D. Beck. 1995. "On the Utility of Lagged Ratio-Correlation as a Short-term County Population Estimation Method: A Case Study of Washington State." *Journal of Economic and Social Measurement* 21:1-16. The utility of a new, short-term method of projecting future population at the county level is evaluated by the authors and compared with two alternative methods, exponential extrapolation and the cohort-component method, using data for Washington State for 1970, 1980, and 1990. The evaluation suggests that the lagged ratio-correlation method consistently has a high level of utility for all three time points. It achieves reductions in error that are comparable to those achieved by the Cohort-Component Method, yet with much less resource requirements. The exponential extrapolation method is found to have high and moderate utility in two of the three time points.

Swanson, D., G. Hough, J. Rodriguez, and C. Clemans. 1998. "K-12 Enrollment Forecasting: Merging Methods and Judgment. *ERS Spectrum* 16 (4): 24-31 pp. Arlington, Virginia. This article describes an enrollment forecasting process in which technical experts and local community stakeholders in Oregon worked together to produce data that were cost-efficient and yet accurate enough to serve as the basis for sound decisions. The large school district that employed this process gained valuable insights to guide its attendance zone and facilities planning decisions, and also created a group of key communicators to ensure community understanding and support. In an ex post facto evaluation, the authors find that the method produces forecasts sufficiently accurate for the school district to make the right decisions about its capital facility needs.

Swanson, D., Morrison, P., Sharkova, I., Tayman, J., and Popoff, C. 1999. *Evaluation of the Washoe County Population Estimation System (WCPEM), and Recommendations for Improving the County's Estimation Program*. Report submitted by S.A.I.C., Inc. to the Washoe County Department of Community Development, Reno, Nevada. The authors evaluated data and human capital resources available to Washoe County, Nevada for purposes of developing annual population estimates and made recommendations on how to use them to accomplish this goal, including the use of assessor parcel records.

Swanson, D., V. Kanhaiya; Y. Riad, B. Barry, and R. Prevost. "Impact of Census Error Adjustments on State Population Projections: the Case of Ohio. *Ohio Journal of Science* 89 (1) : 26-32 National undercount adjustment factors from the 1970 and 1980 U.S. censuses are used to prepare population projections for Ohio, which are in turn compared with unadjusted projections. The findings suggest that decisions concerning adjustment factors have varying effects on short-term, long-term, and strategic forecasting. These effects are particularly salient for selected age-groups and the impact on state government budget decisions typically associated with these age-groups. The authors recommend that

the effects of alternative adjustment possibilities be examined by state demographic centers and budget offices.

Sweet, J. 1984. “Components of Change in the Number of Households: 1970-1980.” *Demography* 21 (2): 129-140. This paper decomposes the growth in the number of households during the 1970s into components associated with changing age and marital status composition and changing age by marital status-specific propensities to form households. About one-third of the increase in the number of households was due to increased age by marital status propensity to form households, and two-thirds was due to shifts in the age by marital status distribution and population growth.

Tang, Z. 2008. “A New Approach to Measuring Migration and Local Population.” *Canadian Studies in Population* 35 (1): 27-48. The author proposes using IRS data to estimate migration at the sub-county level using a process that consists of two major steps. The author argues that this method can produce timely population estimates by race, sex, and age at the sub-county level in an efficient and economical manner. As well generating estimates at state and county levels using bottom-up method. A comparison between the author’s estimates and those produced by the US Bureau of the Census indicates that the historically problematic population underestimation for Massachusetts has been significantly improved through the application of this new methodology.

Taeuber, I. 1949. “Literature on Future Populations, 1943-1948. *Population Index* 15 91): 2-30. This bibliography begins with a statement by the author that it is intended as a supplement to bibliographies found in “*The Future Population of Europe and the Soviet Union: Population Projections, 1940 -1970*,” by Notestein et al. (1944). The author proceeds to give a general orientation to the field of “future estimation,” with specific examples, some discussion on methods (e.g., growth curves), prospects for obtaining data on future populations, carrying capacities, economic and social analyses, “short-time estimates, sub-national projection issues, and the basic research underway on the determinants of population change. The bibliography itself contains 224 entries, organized by thematic (e.g., migration) and geographic (e.g., Germany) areas.

Tarver, J. 1962. “Evaluation of Census Survival Rates in Estimating Inter-censal State Migration.” *Journal of the American Statistical Association* 57 (300): 841-862. Census survival rates have certain technical advantages over other residual methods, specifically the exclusive virtue of a built-in technique which corrects for net census undercount by age, including under-numeration and misstatements of age. Using three sets of estimated "closed" native populations and reported United States deaths as standards, this study evaluates three sets of national census survival rates. The study finds insignificant differences between adjusted and unadjusted survival rates and among the forward, reverse, and average formulas. Finally, it estimates net interstate migration for native whites, total whites, and nonwhites, by age and sex, for the 1940-50 decade, using census survival rates, and compares the results with the more correct estimates for all ages obtained with vital statistics data. By using the formula which most accurately duplicates "vital statistics" deaths over all ages, one obtains the most precise inter-censal state net migration estimates by age groups.

Tayman, J. 1999. "Population and Housing estimates for Census Blocks: The San Diego Experience." Paper presented the Estimates Methods Conference, U. S. Bureau of the Census, Suitland, MD. (available online, http://www.census.gov/population/www/coop/popconf/paper.html, last accessed June 2007). This paper reports on the integrated system developed by San Diego Association of Governments (SANDAG) for preparing annual population and housing estimates by census block. It also evaluates two sources of address-level information - utility hook-ups and assessors parcel file information - that SANDAG has used to develop these estimates, focusing on their accuracy and their strengths and weaknesses. The paper concludes with some general issues involved in estimating block-level population and housing, including maintaining an up-do-date GIS system, data management, temporal consistency, and controlling.

Tayman, J. 1996. "The Accuracy of Small-area Population Forecasts Based on a Spatial Interaction Land-use Modeling System." *Journal of the American Planning Association* 62 (1): 85-98. The author evaluates census tract forecasts based on a spatial interaction modeling system known as DRAM/EMPAL in a case study of San Diego County, California that focuses on performance measures that provide a broader look at forecast error. Mean absolute and algebraic percent errors are the most commonly used measures of precision and bias, but the author points out that measures based on the average can overstate the typical error because forecast error distributions are often positively skewed. Skewed error distributions, which occur with greater frequency in small areas, often result in different values for the median and the average of the central tendency.

Tayman, J. 1994. "Estimating Population, Housing and Employment for Micro-Geographic Areas." pp. 101-107 in K. V. Rao and J. W. Wicks (Eds.) *Studies in Applied Demography, 1994*. Bowling Green, OH: Population and Society Research Center: Bowling Green State University. This paper described procedures for developing current estimates of population, housing and place-of-work employment for micro-geographic areas in San Diego County, California. It demonstrates that with the proper technology and innovative use of nontraditional data sources, it is possible to estimate these activities for sub-census tract geographic areas. The procedures represent an interdisciplinary approach to the problem, blending the theory and methods from demography and geography. Residential meter and wage and salary employer address-level files allow a reasonably accurate determination of the micro-geographic location of employment and residential activities.

Tayman, J. 1991. "Population and Housing Estimates for Micro-Geographic Areas: A Blend of Demographic and GIS Techniques." Presented at the annual meeting of the Population Association of America. Washington, D.C. This paper describes the tools used to develop population and housing estimates for micro-geographic areas from electric meter residence files, census tract data, and its subdivisions Traffic Analysis Zones (TAZ), gridcells forming the Master Geographic Reference Area (MGRA), and the GBF/DIME file and a GIS procedure called ADMATCH, which maps and matches addresses to a geographic element. The biases occurring from mismatches or missing meters did not cause

a serious problem in estimating. Unique geographic areas can now be created using the procedure POLYGON-to-POLYGON Overlay where MGRA polygon boundaries are fit to the designated polygon boundaries. The procedure in ARC/INFO called NETWORK can generate estimates for an area 10 minutes from a given site. An example is given.

Tayman, J. 1986. "An Integrated System for Subcity Population Estimation." Presented at the annual meeting of the Population Association of America. San Francisco, CA. The author describes the integrated system implemented in the San Diego Region (San Diego County) for preparing census tract level household income estimates. He described the methods, data and procedures used in this system as well as the uses of the estimates generated by the system. He discusses future developments and plans.

Tayman, J. 1980. "Total and Occupied Housing Units as Useful Indicators in Regression Models for Population Estimation." Presented at the Spring FSCPE Meeting, Denver Colorado, April 8-9. Tayman constructs three ratio-correlation regression models based on 1960/1950 ratios for use in estimating the civilian population under 65 years for counties in Washington in 1970 (to compare results with1970 census data). He finds that by adding housing units to a model using voters, 1-8 enrollment, and covered employment the 1970 MAPE is reduced from 5.07% to 4.16% and that by using occupied housing units, the MAPE is reduced from 5.07% to 2.82%. Using the same general data and the ratio-correlation methodology, he then examines relative to 1970 censes data, the estimates from two more regression models, total housing units and occupied housing units. These two model allow him to estimate vacant units (total – occupied) and vacancy rates (vacant/total). In regard to the model for estimating total housing units, he finds that the MAPE is 5.61%, while for the model used to estimate occupied housing units, it is 4.15%. There are substantial differences by county in terms of estimated vacant units and vacancy rates, however, when he combines the results of the two regression models.

Tayman, J., and C. Rynerson. 1997. "An Integrated System for Estimating Subcounty Household Income Distributions. Presented at the annual meeting of the Population Association of America. Washington, DC. The authors described the integrated system implemented in the San Diego Region (San Diego County) for preparing census tract level household income estimates. The estimates were done first for the region, then for 41 groups of census tracts known as sub-regional areas (SRAs), then for the region's 347 census tracts. The modified lognormal curve was employed with a different equation using the parameters of income class, PARM (reflecting constant terms), and EXP (a nonlinear adjustment parameter). To estimate the household income distribution, the parameters of median income, PARM, and EXP were needed. Using 1990 census median incomes and income distributions as inputs, an iterative method was used for 9250 combinations of PARM and EXP to find the parameters that minimize the error between the observed and estimated income distributions. For the investigation of historical trends in the parameters, 1980 census data from eight income ranges were also assembled. Inputs to the model included the base year data, the estimated parameters, control totals from an independent estimate of the number of households for the region, SRAs and census tracts, and adjustment factors calculated from a comparison of the estimated and actual 1990 household income distributions. The integrated household income estimation system was fully implemented by means of finding solutions to the issues of forming households from individual tax returns, residual and adjustment factors, the applicability of the modified lognormal curve, and extrapolating parameters. The internal and external consistency of the estimates was also evaluated.

Tayman, J., and E. Schafer. 1985. “The Impact of Coefficient Drift and Measurement Error on the Accuracy of Ratio-Correlation Population Estimates.” *The Review of Regional Studies*. 15(2): 3-10. In this paper, the authors seek to determine the relative impact of 3 sources of error on the accuracy of population estimates produced by the ratio-correlation technique. These sources include coefficient instability, biased estimators and unreliable post-censal symptomatic indicators. The analysis focuses on a comparison of 1980 total population estimates of the counties of Washington developed from 6 specifications of a ratio correlation model. The analysis demonstrated that, at least in this particular situation, unstable regression coefficients contribute very little to overall estimate inaccuracy; the other 2 sources tended to produce estimates with greater inaccuracy. In addition it was found that the 2 sources of measurement error produced inaccuracies in estimates of about equal level. The findings question O'Hare's conclusion that estimate accuracy using the correlation method cannot be expected to improve unless ways of controlling or adjusting for coefficient drift can be found. From an applications perspective it would seem that greater benefit, i.e., less inaccuracy, can be realized by improving efforts in the area of symptomatic indicator measurement error rather than by modifying the basic ratio correlation model structure. Due to the fact that one is frequently required to produce population estimates from symptomatic data, the authors conclude that the potential impact of improving overall population estimate accuracy appears to be substantial.

Tayman, J., and D. Swanson. 1999. “On the validity of MAPE as a Measure of Population Forecast Accuracy.” Population Research and Policy Review 18 (4): 299 - 232. The authors note that mean absolute percent error (MAPE) is the summary measure most often used for evaluating the accuracy of population forecasts. While MAPE has many desirable criteria, they argue from both normative and relative standpoints that the widespread practice of exclusively using it for evaluating population forecasts should be changed. Normatively, they argue that MAPE does not meet the criterion of validity because as a summary measure it overstates the error found in a population forecast. They base this argument on logical grounds and support it empirically, using a sample of population forecasts for U.S. counties. From a relative standpoint, they examine two alternatives to MAPE, both sharing with it, the important conceptual feature of using most of the information about error. These alternatives are symmetrical MAPE (SMAPE) and a class of measures known as M-estimators. The empirical evaluation suggests M-estimators do not overstate forecast error as much as either MAPE or SMAPE and are, therefore, more valid measures of accuracy.

Tayman, J., and D. Swanson. 1996. “On the Utility of Population Forecasts.” *Demography* 33 (4): 523-528. The authors note that although the forecast evaluation literature is extensive, it is dominated by a focus on accuracy. They go beyond accuracy by examining the concept of forecast utility in an evaluation of a sample of 2,709 [U.S.] counties and census tracts. They find that forecasters provide `value-added' knowledge

for areas experiencing rapid change or areas with relatively large populations. For other areas, reduced value is more common than added value. The authors find that their results suggest that new forecasting strategies and methods such as composite modeling may substantially improve forecast utility.

Tayman, J., E. Schafer, and L. Carter. 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): 1-20. This paper describes a technique for making sub-county population forecasts and for generating confidence intervals around their forecast error. It also develops statistical equations for calculating point estimates and confidence intervals for areas with different population sizes. A non-linear, inverse relationship between population size and forecast accuracy was found and the authors demonstrate the ability to accurately predict average forecast error and confidence intervals based on this relationship. The data consist of 1980 and 1990 census observations and 1990 population forecasts for San Diego County, California.

Tayman, J., D. Swanson, and C. Barr. 1999. "In Search of the Ideal Measure of Accuracy for Sub-national Demographic Forecasts." *Population Research and Policy Review* 18 (5): 387-409. The authors examine nonlinear transformations of the forecast error distribution in hopes of finding a summary error measure that is not prone to an upward bias and uses most of the information about that error. MAPE, the current standard for measuring error, often overstates the error represented by most of the values because the distribution underlying the MAPE is right skewed and truncated at zero. Using a modification to the Box-Cox family of nonlinear transformations, the authors transform these skewed forecast error distributions into symmetrical distributions for a wide range of size and growth rate conditions. They verify this symmetry using graphical devices and statistical tests; examine the transformed errors to determine if re-expression to the scale of the untransformed errors is necessary; and develop and implement a procedure for the re-expression. The authors find that MAPE-R developed by this process is lower than the MAPE based on the untransformed errors and is more consistent with a robust estimator of location.

Thomsen, I., and A. Klievie-Holmy. 1998. "Combining Data from surveys and Administrative Record Systems: The Norwegian Experience." *International Statistical Review* 66 (2): 201-222. The authors note that at Statistics Norway administrative data have been extensively used in order to improve the quality of survey data. Various techniques have been used to reduce sampling variance and/or to reduce the effects of non-response. In this paper, the authors discuss some of the most commonly used methods, and based on empirical rather than theoretical evaluations, they provide their conclusions concerning their potentials and limitations.

United Nations.1969. *Methodology and Evaluation of Population Registers and Similar Systems. Studies in Methods, No. 15*. Sales No. E.69.XVII.15. New York City, NY: United Nations. This manual describes the methods and data commonly found in population registration systems, giving some history. It also describes suggested techniques for evaluating the efficacy of population registers.

United Nations. 1982. *Model Life Tables for Developing Countries*. Sales No. E.81.XIII.7. New York City, NY: United Nations. (Available online, www.un.org/esa/population/publications/Model_Life_Tables/Model_Life_Tables.htm, last accessed May, 2008). This report presents model life tables for developing countries. It gives sets of age-sex patterns of mortality which are based on reliably documented developing country data, thus supplementing the Coale-Demeny model life tables extracted mainly from historical European experience. Both the United Nations model life tables for developing countries and the Coale-Demeny model life tables are also available in the United Nations software package MORTPAK.

United Nations. 1983. *Manual X: Indirect Techniques for Demographic Estimation*. Population Studies, No. 81. Sales No. E.83.XIII.2. New York City, NY: United Nations. (Available online, www.un.org/esa/population/publications/Manual_X/Manual_X.htm, last accessed May, 2008). This manual is designed to assist demographers and population experts in conducting evaluations and exploitation of data sources, especially those that are incomplete or deficient. Manual X describes a wide range of techniques to make indirect estimates of demographic parameters. Each of the techniques presented is based on a mathematical model and explained in easy-to-follow examples.

United Nations. 1984. *Handbook of Household Surveys. Studies in Methods, Series F*, No. 31. Sales No. E.83.XVIII.13. New York City, NY: United Nations. This is a revised edition of the basic document in a technical series published by the United Nations to assist countries in planning, implementing, and utilizing the results from household surveys. It provides overall technical information and guidance of a relatively general nature to middle- and senior-level personnel who are producers or users of survey statistics. Included in the topics discussed are planning, strategy and technical design; selection of topics and their translation into survey instruments and preparation of the accompanying manuals, instructions and training activities; organization and implementation of field work; processing, compilation, tabulation and dissemination of data; evaluation of the data and the procedures used to collect and process them; and analyses and arrangements for active storage and retrieval of the survey results. Selected issues from survey experience in Africa, Asia, and Latin America are discussed.

United Nations. 1985. *Development and Design of Survey Questionnaires*. National Household Capability Programme. INT-84-014. New York City, NY: United Nations. This manual provides a broad review, with numerous illustrations, of questionnaires and study designs for use in the development and implementation of household surveys.

United Nations. 1993. *Sampling Errors in Household Surveys*. National Household Capability Programme. INT- 92-P80-15E. New York City, NY: United Nations. This manual provides a broad review, with numerous illustrations, of major sources of non-sampling errors in household surveys so as to enhance awareness among survey practitioners of factors which determine the quality of data produced through surveys.

The study also provides a practical guide to the measurement and control of non-sampling errors. There are a number of illustrations from a variety of experiences.

United Nations. 1998. *Principles and Recommendations for Population and Housing Censuses, Revision 1*.Sales No. E.98.XVII.8. New York City, NY: United Nations. The structure of the revised Principles and Recommendations closely follows that of the previous recommendations. Modifications are made, in part one, regarding operational aspects of censuses. Part two combines topics for population censuses and those for housing censuses. An entirely new part, part three, has been added to highlight the needs of users. It also contains a section focused on the need to consider the relationship between census topics and specific uses of census data. Formats for selected tabulations for each population and housing topic together with a brief statement for users are shown in annexes. References and an index are shown at the conclusion of the publication.

UNFPA. 1993. *Readings in Population Research Methodology, Volume 5: Population Models, Projections and Estimates*. New York City, NY: UNFPA. One of eight volumes, attempting to put together a complete reference source covering the full range of population research methodology, Volume 5 examines population models, projections, and estimates.

UNFPA. 1993. *Readings in Population Research Methodology, Volume 6: Advanced Basic Tools*. New York City, NY: UNFPA. One of eight volumes, attempting to put together a complete reference source covering the full range of population research methodology, Volume 6 is about advanced basic tools, including survival and event history methods, multi-state methods, contextual and multi-level methods, and techniques for qualitative analysis.

U. S. Census Bureau, Population Division. 2007. "County Housing Unit Estimates, April 1, 2000 to July 1, 2006." Washington, DC: Population Division, U. S. Census Bureau. (available online, http://www.census.gov/popest/housing/files/HU-EST2006_US.CSV, last accessed January, 2008). This is the data set that was developed using the data, methods, and procedures described in the following citation.

U. S. Census Bureau. 2006. Methodology for County Housing Unit Estimates for Vintage 2006." Washington, DC: Population Division, U.S. Census Bureau(available online, http://www.census.gov/popest/topics/methodology/2006_hu_meth.html, last accessed January 2008). This report describes the methods, data, and procedures to develop sub-county population estimates by a housing unit method that uses housing unit change to distribute county population to sub-county areas. It notes that the state and county housing unit estimates are aggregations of these housing unit estimates and that Housing unit estimates use building permits, mobile home shipments, and estimates of housing unit loss to update housing unit change since the last census. Census counts of housing units are geographically updated each year to reflect legal changes reported in the Boundary and Annexation Survey, Census corrections, and other administrative revisions. The 2006 data can be found via the link in the preceding citation.

U. S. Census Bureau. 2004. “The Census Bureau’s Master Address File (MAF): Census 2000 Address List Basics.” Washington, DC. Geography Division, U. S. Census Bureau (available online http://www.census.gov/geo/mod/maf_basics.pdf , last accessed January 2008). This report describes the content and procedures used to develop and maintain the MAF.

U. S. Census Bureau. 1990. “State Population and Household Estimates.” *Current Population Reports, P25, No. 1058*. Washington, DC: U. S. Government Printing Office. This report contains population and household estimates for U.S. States for July 1, 1989, revised household estimates for July 1, 1988, revised annual estimates for the resident and civilian populations for July 1, 1981-88, and components of population change for the 1980-89 period. The annual estimates of the resident population are further disaggregated by sex and age. The age detail includes 10-year and selected broad age groups.

U. S. Census Bureau. 1987. “State Population and Household Estimates, with Age, Sex, and Components of Change: 1981-86.” *Current Population Reports, P25, No. 1010*. Washington, DC: U. S. Government Printing Office. This report contains provisional estimates for States of the resident and civilian U.S. populations and of households for July 1, 1986, revised annual population and household estimates for July 1, 1981, through 1985, and components of population change for the 1980-86 period.

U. S. Census Bureau. 1986. “State Population and Household Estimates to 1985, With Age and Components of Change.” *Current Population Reports, P25, NO. 998*. Washington, DC: U. S. Government Printing Office. This report contains provisional estimates of the resident and civilian populations of U.S. States for July 1, 1985, revised population estimates for July 1, 1981, through 1984, and components of population change and revised household estimates for States for the 1980-85 period. Annual inter-censal estimates by age for 1971 to 1979 are also presented. This is the first time that estimates of population and households for States have been included in the same issue of Current Population Reports.

U. S. Census Bureau. 1985. “Estimates of Households for States: 1981 to 1984.” *Current Population Reports, P25. No. 974*. Washington, DC: U. S. Government Printing Office. This report presents estimates of the total number of households for U.S. States for July 1, 1981, 1982, and 1983, and provisional estimates for July 1, 1984. In addition to these estimates, census counts of households for 1970 and 1980 are included for comparative purposes.

U. S. Census Bureau. 1979. “Estimates of the Number of Households for States: 1977 and 1978.” *Current Population Reports, P25, No. 807*. Washington, DC: U. S. Government Printing Office. This report presents estimates of the total number of households for U.S. States for July 1, 1977 and 1978.

U. S. Department of Commerce. 2002. “A Better Strategy is needed for Managing the Nation’s Master Address File.” *Inspection Report No. OSE-12065*. Washington, D. C.: U.

S. Department of Commerce. This report found that Census did not establish procedures that would adequately ensure the MAF's accuracy and completeness, and thus used an error-prone database to conduct the decennial census in 2000. Among other problems, the report found that numerous addresses did not link to the correct location in the TIGER file, which hampered the Census Bureau's ability to accurately determine a housing unit's location. Some errors in the file occurred because address verification plans were flawed. In addition, the report finds that when the Census Bureau modified its software for eliminating duplicate addresses in an effort to improve its coverage of multiunit housing such as trailer parks and apartment buildings, the change allowed many suspected duplicate addresses to remain in the address file. The report also notes that maps often contained duplicate and missing streets, and were not always printed in a usable size and format.

U. S. GAO. 2006. *2010 Census: Census Bureau Needs to Take Prompt Actions to Resolve Long-standing and Emerging Address and Mapping Challenges*. GAO-06-272. Washington, D.C. U. S. Government Accountability Office (Note: Effective July 7, 2004, the GAO's legal name became the Government Accountability Office). This report found that the Census Bureau's address and map modernization efforts have progressed in some areas. The Bureau is researching how to correct addresses that were duplicated, missed, deleted, and incorrectly located on maps. However, some deadlines for completing research are not firm, while other deadlines that had been set continue to slip. Thus, whether research will be completed in enough time to allow the Bureau to develop new procedures to improve the 2010 address file is unknown. Also, the Bureau has not fully addressed emerging issues. For one such issue, the Bureau has acknowledged the compressed time frame for completing address canvassing--an operation where census workers walk every street in the country to verify addresses and maps--but has not reevaluated the associated schedule or staffing workloads. Also, the Bureau has allotted only 6 weeks to conduct address canvassing it completed in 18 weeks in 2000 and expanded the operation from urban areas in 2000 to the entire country in 2010. Whether the Bureau can collect and transmit address and mapping data using the MCD is unknown. The MCD, tested during 2006 address canvassing, was slow and locked up frequently. Bureau officials said the MCD's performance is an issue, but a new MCD to be developed through a contract awarded in March 2006 will be reliable. However, the MCD will not be tested until the 2008 Dress Rehearsal, and if problems emerge, little time will remain to develop, test, and incorporate refinements. If after the Dress Rehearsal the MCD is found unreliable, the Bureau could face the remote but daunting possibility of reverting to the costly paper-based census of 2000. Bureau officials do not believe a specific plan is needed to update the addresses and maps for areas affected by the hurricanes. Securing a count is difficult under normal conditions, and existing procedures may insufficient to update addresses and maps after the hurricanes' destruction--made even more difficult as streets, housing, and population will be in flux.

Verma, R.B.P. 1992. "Regression-Based Estimates of Total Population for Local Areas: Then and Now." pp. 167-178 in P. Krishnan, C. Tuan, and K. Mahadevan (Eds.) *Readings in Population Research: Policy, Methods and Perspectives.* Delhi, India: B.R. Publishing. This paper provides an overview of the methodological developments in the regression method for estimating local area population and described how the estimation procedure is applied to census divisions in Canada.

Verma, R., K. Basavarajappa, and R. Bender. 1984. "Estimation of Local Area Population: An International Comparison." pp. 324-329 in *1984 Proceedings of the Social Statistics Section,* Alexandria, VA: American Statistical Association. This paper discusses the current status of methodology on local area population estimation in five developed countries, Australia, Canada, England and Wales, New Zealand, and the U.S.A. It notes that the methods employed are compared with respect to the accuracy and the timeliness of estimates and the consistency of procedures and data sources used for different geographic levels.

Verma, R., K. Basavarajappa, and R. Bender. 1983. "The Regression Estimates of Population for sub-provincial Areas in Canada." pp. 512-517 in *1983 Proceedings of the Social Statistics Section,* Alexandria, VA: American Statistical Association. The authors describe two sets of post-censal population estimates that are published yearly by the government of Canada. The estimates are for census divisions and census metropolitan areas and appear 3 to 4 months and 12 to 15 months following the reference date. "The regression technique uses family allowance recipients as the main symptomatic indicator and where available, additional indicators to derive population change for the current year. The first set is obtained by adding this change to the second set for the previous year produced by the component method, with births and deaths from vital registers, and estimated migration from Revenue Canada taxation files. The two sets were found to be statistically similar, though the first set is more timely, and the second providing more details on the components of population change.

Voss, P., C. Palit, B. Kale, and H. Krebs. 1995. "Censal Ratio Methods." pp. 70-88 in N. W. Rives, W. Serow, A. Lee, H. Goldsmith, and P. Voss (Eds). *Basic Methods for Preparing Small-Area Population Estimate*s. Madison, WI: Applied Population Laboratory, University of Wisconsin. The authors provide a brief history of the development of censal ratio methods, describe its statistical properties, discuss its advantages and disadvantages, and suggest ways in which they can be improved. They also discuss its variations and provide the results of a 1980 test using data from Wisconsin.

Walashek, P., and D. Swanson. 2006. "The Roots of Conflict over U. S. Census Counts in the Late 20th Century and Prospects for the 21st Century." *Journal of Economic and Social Measurement* 31 (3-4): 185-205. Although not originally intended as such, the authors argue that the U. S. census has become a "Commons" in which private benefits are gained at the expense of public costs. The historical development of the census as a Commons first clearly emerged with the release of the 1970 census results, and since that time contentious litigation over census undercount error has become a standard part of the decennial census landscape. Political battles within the federal government have gone hand-in-hand with these litigation activities. They culminated with a Supreme Court decision on the legality of statistically adjusting census 2000 counts for estimated undercount error. As these battles raged, professional interest in providing methodological fixes for net census undercount error increased while public participation in the census generally declined. The authors examine the history behind these legal battles, the legislative acts, and judicial decisions that led to the 16th Amendment and the loss of the careful balance between public costs and private benefits crafted by the Founding Fathers in Article I of the Constitution. They identify the role that historical actions played in making the census into a Commons, thereby setting the stage for modern day census litigation and other forms of conflict. The authors observe that as a Commons, the census is facing a potential collapse that cannot be prevented by methodological developments and conclude by noting that a course of political action may be the best course for preventing such a collapse.

Wang, C. 1999. "Development of National Accounting of Address and Housing Inventory: The Baseline Information for Post-censal Population estimates." Paper presented the Estimates Methods Conference, U. S. Bureau of the Census, Suitland, MD. (available online, http://www.census.gov/population/www/coop/popconf/paper.html, last accessed June 2007). The paper discusses the use of the Master Address File for a number of purposes, including the development of data that can be used for population estimates made using the Housing Unit Method.

Warner, S. 1995. "Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias." *Journal of the American Statistical Association* 60: 63-69. For various reasons, the author notes that individuals in a sample survey may prefer not to confide to the interviewer the correct answers to certain questions. In such cases the individuals may elect not to reply at all or to reply with incorrect answers. The resulting evasive answer bias is ordinarily difficult to assess. In this paper, the author argues that such bias is potentially removable through allowing the interviewee to maintain privacy through the device of randomizing his response. A randomized response method for estimating a population proportion is presented as an example. Unbiased maximum likelihood estimates are obtained and their mean square errors of conventional estimates under various assumptions about the underlying population.

Webster, C. 1996. "Population and Dwelling Unit Estimates from Space." *Third World Planning Review* 18 (2): 155-76. This paper reports on attempts to measure the morphological patterns in an urban satellite scene and to use these for image interpretation. The interpretation task addressed is the estimation of residential dwelling units from the patterns discernible in high resolution satellite images of cities. The practical results include dwelling estimates that can be aggregated to any geographical unit of analysis, population estimates for cities and a dwelling density surface that can be categorized into any number of residential land-use classes.

Wetrogan, S. 2007. "Evaluating County Population Estimates Using a Housing Unit Based Approach." Paper presented at the annual meeting of the Southern Demographic Association. Birmingham, AL. The paper compares four alternative 1990-based housing unit estimates for 2000 to Census 2000 population counts and to a separate set of 1990-based county population estimates for 2000 prepared using the administrative records component of change method.

Williams, G. 1999. "The Measurement of Migraton Coverage Bias." Paper presented the Estimates Methods Conference, U. S. Bureau of the Census, Suitland, MD. (available online, http://www.census.gov/population/www/coop/popconf/paper.html, last accessed June 2007). The author observes that a key variable in the production of county and state estimates under the Census bureau's estimates methodology of the 1990's is IRS migration and notes that there is differential coverage of persons from tax returns, resulting either from no-filers or first time or last time filers, there is likely to be a cumulative bias in the measured amount of migration. He compares migration flows as measured by IRS migration and Alaska Permanent Fund and suggests ways to improve the measurement of migration.

Williams, B., J. Nichols, and M. Conroy. 2002. *Analysis and Management of Animal Populations: Modeling, Estimation, and Decision Making.* San Diego, CA. Academic Press. This book is an all-in-one reference for just about every quantitative technique that could be used in studying animal populations. Topics include sampling, experimental design, statistical analysis, mark-recapture and other techniques for the estimation of population size and vital rates, population modeling, and decision analysis. The authors have organized these subjects by dividing their book into four sections. The first section provides a framework for the book and includes discussions of the scientific method, the use of models in population ecology, hypothesis testing, sampling and experimental design. The second section deals with mathematical models applied to animal populations and contains, among other things, an overview of classical demographic models, a thorough introduction to stochastic processes, and a discussion of the applied uses of demographic models in management and conservation. The third section covers estimation of abundance and of population vital rates, including survival, movement, and recruitment. Estimation of community attributes, such as species diversity, is also covered. Finally, the fourth section of the book covers decision analysis with a focus on optimal management of populations in the presence of uncertainty. This section ends with a case study focused on management of game birds.

Wilson, T., and P. Rees. 2005. "Recent Developments in Population Projection Methodology: A Review." *Population, Space, and Place* 11 (5): 337-360. In this paper the authors survey five streams of research that have made important contributions to population projection methodology over the last decade. These are: (i) the evaluation of population forecasts; (ii) probabilistic methods; (iii) experiments in the projection of migration; (iv) projecting dimensions additional to age, sex and region; and (v) the use of scenarios for 'what if?' analyses and understanding population dynamics. Key developments in these areas are discussed, and a number of opportunities for further research are identified.

Wolter, K. 1990. "Capture-Recapture Estimation in the Presence of a Known Sex Ratio." *Biometrics* 46 (1): 157-162. The author presents new methods of estimating population size abased on capture-recapture data. The methods exploit knowledge of the sex ratio, males per female, and permit "estimability" even when both time of sampling and marking affect the “catchability” of an animal. An example is presented involving Microtus pennsylvanicus (meadow vole).

Wolter, K. and B. Causey. 1991. “Evaluation of Procedures for Improving Population Estimates for Small Areas.” *Journal of the American Statistical Association* 86 (414): 278-284. The authors provide and illustrate methods for evaluating across-the-board ratio estimation and synthetic estimation, two techniques that might be used for improving population estimates for small areas. The methods emphasize determination of a break-even accuracy of knowledge concerning externally obtained population totals, which marks the point at which improvement occurs.

Woodrow-Lafield, K. 2001 "Implications of Immigration for Apportionment." *Population Research and Policy Review* 20 (4): 267-289. The author observes that around the time of the US decennial censuses, a renewed interest emerges in the method for apportioning the US House of Representatives. Various methods may show slight variations in illustrative apportionments, with biases favoring less populous states, but the general pattern remains. Definition of certain groups as included in the apportionment counts and coverage levels for selected groups have been debated in the judicial system, legal journals, and government. Unauthorized residents, and, sometimes, lawful immigrants, are often singled out for exclusion. The legal issues are complex, and illustrating the effects of these groups' inclusion is problematic due to poor measures, nationally and geographically. Using approximate distributions, these analyses suggest this next apportionment might differ slightly under various scenarios such as ones excluding either recently entered unauthorized residents or all unauthorized residents. Allowing for net authorized immigration greater than official estimates for the 1990s might have some effect for large states.

Woodrow-Lafield, K. 1995. "An Analysis of Net Immigration in Census Coverage Evaluation." *Population Research and Policy Review* 14 (2): 173-204. The author notes that national surveys monitored growth in the foreign-born population for the 1980s, especially net undocumented migration's continuing role, but the 1990 census portrayed an even larger foreign-born population than these surveys. Under-coverage in 1990 could have been higher than initially presented because preliminary studies may have insufficiently accounted for decadal net immigration. Assumptions intended to maintain a high undocumented undercount performed poorly when census counts of foreign-born residents became known. Any point estimate for net undocumented migration, calculated as a residual, is likely to be biased by assumptions and data gaps for components of calculating net legal immigration, especially in the direction of underestimation. A reasonable statement is that at least 2.1–2.4 million undocumented residents were enumerated in the 1990 census. The number of un-enumerated undocumented residents may easily have ranged between 0.5 million and 3.0 million, and a narrower range of 1 million to 2 million is plausible. Despite the importance of undocumented migration measurement for census evaluation and policy purposes, differences among various undocumented estimates are more likely to stem from discrepancies in universe, reference dates, or individual judgment, rather than analytic refinement. Better measurement of the foreign born population or its census coverage would aid in setting upper limits on net undocumented migration.

Word, D., and M. Zitter. 1982. "Further Developments in Inter-censal Population Estimation Using Administrative Records." pp. 260-265 in *1982 Proceedings of the Social Statistics Section,* Alexandria, VA: American Statistical Association. This paper relates to the evaluation of U.S. intercensal population estimates for states and metropolitan areas ongoing at the Census Bureau. The focus of the paper extends the present use of Administrative Records (individual income tax records in this instance) for the purpose of preparing Black population estimates for both States and SMSAs.

Wunsch, G., And M. Termote. 1978. *Introduction to Demographic Analysis: Principles and Methods*. New York City, NY: Plenum Press. This book can serve as a textbook or references book. It provides a unified introduction to the principles and methods essential to conduct analyses of population trends and composition.

Zeng, Y., K. Land, Z. Wang, and D. Gu. 2006. "U. S. family household momentum and dynamics: an extension and application of the ProFamy Method." *Population Research and Policy Review* 25: 1–41. Unlike the headship-rate method, the ProFamy method uses demographic rates as input and can project detailed household types, sizes, and living arrangements for all members of the population. Tests of projections from 1990 to 2000 using ProFamy and based on observed U.S. demographic rates before 1991 show that discrepancies between ProFamy projections and census observations in 2000 are reasonably small, validating the new method.

Zachariah, K. 1962. "A Note on the Census Survival Ratio Method of Estimating Net Migration." *Journal of the American Statistical Association* 57 (297): 175-183. In the census survival ratio method, the author observes that one of the common assumptions is that the proportion which the enumerated population in any age-sex group bears to the actual population is the same at each census for each state as for the nation. The significance of this assumption is examined and it is established that such an assumption is not necessary. It is therefore replaced by a less stringent assumption, namely that the ratio of the degree of enumeration in a state (i.e. the proportion which the enumerated population in any age-sex group bears to the true population) to that of the nation is the same for the same cohort. In the light of the change in assumption, it is shown that much of the criticism of the method made by Daniel O. Price is untenable. Though his main conclusion (i.e. small relative differences in estimates of net migration should be interpreted with extreme caution) is correct, the errors in migration rates estimated by the census survival ratio method are probably much smaller than his analysis suggests.

Zaslavsky, A. 1993. "Combining Census, Dual-System, and Evaluation Study Data to Estimate Population Shares." *Journal of the American Statistical Association* 88 (423): 1092-105. This article addresses methods for combining the census, the Dual Sysatems Estimates (DSE), and bias estimates obtained from the evaluation programs to produce accurate estimates of population shares, as measured by weighted squared- or absolute-error loss functions applied to estimated population shares of domains. Several procedures are reviewed that choose between the census and the DSE using the bias evaluation data or that average the two with weights that are constant across domains. A multivariate hierarchical Bayes model is proposed for the joint distribution of the undercount rates and the biases of the DSE in the various domains. The specification of the model is sufficiently flexible to incorporate prior information on factors likely to be associated with undercount and bias. When combined with data on undercount and bias estimates, the model yields posterior distributions for the true population shares of each domain. The performance of the estimators was compared through an extensive series of simulations. The hierarchical Bayes procedures are shown to outperform the other estimators over a wide range of conditions and to be robust against misspecification of the models. The various composite estimators, applied to preliminary data from the 1990 Census and evaluation programs, yield similar results that are closer to the DSE than to the census. Analysis of a revised data set yields qualitatively similar estimates but shows that the revised post-stratification improves on the original one.

Zitter, M., and H. Shryock. 1964. "Accuracy of Methods of Preparing Post-censal Population Estimates for States and Local Areas." *Demography* 1(1): 227–241. A general assessment of methods that are and can be used for developing post-censal population estimates at the sub-national level in the United States. The authors discuss strengths and weaknesses of these methods.

II. The Housing Unit Method

Brown, W. 1999. " Use of Property Tax Records and Household Composition Matrices to Improve the Household Units Method for Small Area Population Estimates." Paper presented the Estimates Methods Conference, U. S. Bureau of the Census, Suitland, MD. (available online, http://www.census.gov/population/www/coop/popconf/paper.html, last accessed June 2007). This paper reports on a pilot effort in a single county to use property tax records in place of housing permits for building and demolition to improve count of housing units. In addition the property tax records include an x-y coordinate for spatially locating the housing unit structure. This enables the housing units to be geocoded to Census Blocks, as well as to non-Census geographic areas such as watersheds. Finally, the use of household composition matrices yields improved estimates of persons per household for small areas. The household composition matrices contain a refined measure of persons per household built from cells of persons per household cells by age of person and age of householder. The improved housing unit method is demonstrated for Census Block groups.

Byerly, E. 1990. "State and Local Agencies Preparing Population and Housing Estimates." *Current Population Reports, Series P-25, No. 1063*. Washington, DC: US Government Printing Office. This report presents the results of a mail survey of 1,049 State and local public agencies conducted from June to December 1989. The voluntary survey was intended to inventory official State and local governmental and other public agencies making estimates of population and/or housing units and the methods they use. An extensive bibliography of population estimation methods is included in the report.

Byerly, E. 1989. "State Population and Household Estimates, with Age, Sex, and Components of Change." *Current Population Reports, Series P-25, No. 1044*. Washington, DC: US Government Printing Office. This report contains provisional US state estimates of: (1) the resident and civilian populations and of households for July 1, 1987; (2) revised annual population and household estimates for July 1, 1981-1986; and (3) components of population change for the 1980-1987 period. Also shown are annual age estimates and median age of the resident population of states, 1981-1987, by sex for 10-year age groups and selected broad age groups.

Devine, J. and C. Coleman. 2003. "People Might Move but Housing Units Don't: An Evaluation of the State and County Housing Unit Estimates." *Population Division Working Paper Series No. 71*. Washington, DC: U. S. Census Bureau. This report presents an evaluation of estimates of total housing units for the Nation, states, and counties produced by the Population Division of the Census Bureau. The comparison of the April 1, 2000 estimates to the April 1, 2000 decennial census counts forms the basis for this report. The evaluation found that the 2000 state and county level housing unit estimates developed from building permit, mobile home shipment, and demolition data performed with a degree of accuracy similar to the state and county April 1, 2000 population estimates produced by the Population Division of the Census Bureau.

Findley, S. and H. Reinhardt. 1980. "Nonlinear Estimation of Household Size: The Minnesota Housing Unit Method." Paper presented at the annual meeting of the Population Association of America, Denver, CO. Data from the 1977 Minnesota household survey are used to determine housing categories and size of place distinctions to estimate household size (PPH) and vacancy parameters. Using evidence from California and Washington, the authors determine that PPH change during the 1970s had been non-linear.

Gibson, C. 1988. "Estimates of Households, for Counties: July 1, 1985." *Current Population Reports, Series P-23, No. 156*. Washington, DC: U.S. Bureau of the Census. The author describes the methodology used by the Census Bureau to estimate households at the county level. He discusses group quarters and the adult population in households. Tables provide estimates of the number of households and the average number of persons per household by county and county equivalent. An appendix contains corrections to the 1980 census counts of population and housing units for selected states and counties.

Gibson, C. 1986. "Post-censal estimates of households by size and type for states and of total households for counties." Paper presented at the Population Association of America Annual Meeting, San Francisco. This paper examines a demographic approach to sub-national household estimates which assumes that national trends in the average adult population per household and in the distribution of households by type and by size occur at the sub-national level. The assumption is tested for the 1970-1980 decade using US census data. Results show that, for states, the mean absolute error is lowest for the 20 years and up population (.69%), followed by the 18 years and up population (.76%), the 15 years and up population (1.05%), and all ages (1.61%). Counties show a lower mean error for the 18 and up group (1.47%), followed by the 20 and up (1.59%), the 15 and up (1.7%), and all ages (2.95%).

Gonzalez, M., and C. Hoza. 1978. "Small Area Estimation with Applications to Unemployment and Housing Estimates." *Journal of the American Statistical Association* 79: 7-15. The purpose of this study is to investigate methodologies for constructing inter-censal estimates of various characteristics of the population for small areas. The proposed methodology is illustrated mainly in the context of unemployment estimates, with one section utilizing dilapidated housing estimates. Alternative synthetic estimates of unemployment based on the 1970 Census 20-percent sample are investigated and their relative error is analyzed. The reliability of the synthetic estimates is discussed in the context of dilapidated housing estimates. Two types of regression models are studied, and the improvements obtained by excluding outliers from the regression are discussed.

Hepburn, G., T. Mayor, and J. Stafford. 1976. "Estimation of Market Area Population from Residential Electrical Utility Data." *Journal of Marketing Research* 13 (3): 230-236. An ability to project current population in small market trading areas is critical to many types of market research. The use of selected utility data to estimate current population in small geographic areas is evaluated.

Jarosz, B. 2008. “Using Assessor Parcel Data to Maintain Housing Unit Counts for Small Area Population Estimates.” pp. 89-102 in S. Murdock and D. Swanson (Eds.) *Applied Demography in the 21st Century*. Dordrecht, The Netherlands, Springer. This paper describes the reasons why the San Diego Association of Governments switched from census and permit data to county assessor records to build housing unit files for purposes of estimating population by the Housing Unit Method. It describes the advantages of using the county assessor data as well as the disadvantages and difficulties they entail. It concludes with a description of the uses to which the data will be put, notably in the micro model known by the acronym of PECAS.

Kimpel, T., and T. Lowe. 2007. “Estimating Household Size for use in Population Estimates.” *Population Estimates and Projections Research Brief No. 47*. Olympia, WA: Washington State Office of Financial Management. This Brief revisits a topic developed in Research Brief No. 10 where a regression procedure was developed using administrative data to update household size—a key variable used in local population estimates based on the Housing Unit Method. The authors find that generally the most accurate population estimates come from using several procedures and understanding the biases in each.

Lowe, T. 2000. “Occupied versus Rented: Census and Real Estate Occupancy Rates and their use in Population Estimates.” *Population Estimates and Projections Research Brief No. 9*. Olympia, WA: Washington State Office of Financial Management. This Brief identifies differences in the definition of “occupied” as used by the federal Bureau of the Census and local Real Estate Surveys. It also provides a simple procedure for adjusting the real estate survey data for use in the estimation process.

Lowe, T. 2000. “Developing Trends in Household Size for use in Population Estimates.” *Population Estimates and Projections Research Brief No. 10*. Olympia, WA: Washington State Office of Financial Management. This Brief illustrates a regression procedure using administrative data that can be used to update household size in the Housing Unit Method.

Lowe, T. 1988. “A Resurrection: The Potential of Postal Survey Data in Improving Housing Unit Population Estimates for Local Areas.” Paper presented at the annual meeting of the Population Association of America., New Orleans, LA. The accuracy of postal survey data in reporting residential housing unit occupancy estimates was compared to that of 1970 and 1980 U.S. census data for 26 Washington State cities. These postal surveys were conducted by the Federal Department of Housing and Urban Development in the 1970s within 2 months of collection of census data. Results are given in terms of absolute error of vacancy rates. Postal surveys almost always show lower vacancy rates than census data because they do not include unfinished or new units, or concealed unoccupied conversions in single family homes. Suburban single family housing generally has the highest occupancy rates. Postal data are much more accurate than census data when occupancy rates are variable, as in cities near military bases, and in multi-unit structures. This variability precludes the use of an adjustment factor in most cases. Adjusted postal occupancy rates were more accurate when city and postal boundaries corresponded and remained stable. The results of this study indicate the potential value of postal survey data when occupancy rates vary.

Lowe, T., and M. Mohrman. 2004. "The Pasco Project: An Opportunity to Identify Sources of Sample Bias." *Applied Demography* 17(1): 6-8. This paper describes the procedures, including GIS, that were used to develop a sample frame to be used in a re-assessment of the Housing Unit Method-based estimate of the city of Pasco for purposes of Washington State's annual population estimation program. In conclusion, it notes that an accurate list of housing units (e.g., the Master Address File) is a necessity for the American Community Survey.

Lowe, T., and M. Mohrman. 2003. "Using a Geographic Information System in a Population Estimate Program: The Pasco Case." *Population Estimates and Projections Research Brief No. 21*. Olympia, WA: Washington State Office of Financial Management. This paper describes the GIS procedures that were used to develop a sample frame to be used in a re-assessment of the Housing Unit Method-based estimate of the city of Pasco for purposes of Washington State's annual population estimation program.

Lowe, T., and M. Mohrman 2003. "Use of Postal Delivery Data in the Population Estimate Process." *Population Estimates and Projections Research Brief No. 18*. Olympia, WA: Washington State Office of Financial Management. This Brief extends the evaluation of postal vacancy/occupancy rates presented in Research Brief No.17. It examines the relationship between vacancy trends in postal data versus census data, describes how the postal rates were used in the 2003 estimation process using the Housing Unit Method, and identifies avenues for further evaluation.

Lowe, T., M. Mohrman, and D. Brunink. 2003. "Developing Postal Delivery Data for use in Population Estimates." *Population Estimates and Projections Research Brief No. 17*. Olympia, WA: Washington State Office of Financial Management. This case study underscores the problems inherent in developing address lists for census and sample purposes. The findings in this case study led the authors to caution that the ACS may have substantial coverage errors unless a great deal attention is paid to its frame, an extract of the Master Address File.

Lowe, T., L. Weisser, and B. Myers. 1984. "A Special Consideration in Improving Housing Unit Estimates: The Interaction Effect. Paper presented at the annual meeting of the Population Association of America, Minneapolis, MN. The authors observe that of some of the terms in the equation for the Housing Unit Method are themselves correlated and discuss how this can be dealt with in using the Housing Unit method to estimate population. They specifically point to the interaction of average household size and structure type.

Lowe, T., D. Pittenger, and J. Walker. 1977. "Making the Housing Unit Method Work: A Progress Report." Paper presented at the annual meeting of the Population Association of America, St. Louis, MO. This paper describes the procedures used by the Washington

State Office of Financial Management in its statewide program for using the Housing Unit Method to annually estimate the populations of Washington's cities and towns. It notes challenges and ways to overcome them.

McHenry, J. 1999. "A Local Alternative to Demographic Data Vendor Population and Household Estimates – Applying a Modified Housing Unit Approach to Individual Households." Paper presented the Estimates Methods Conference, U. S. Bureau of the Census, Suitland, MD. (available online, http://www.census.gov/population/www/coop/popconf/paper.html, last accessed June 2007). The author proposes a solution to problems encountered with small area data purchased from vendors that uses readily available local data combined with the housing method. Purchasers of small-area demographic data from national data vendors confront three interrelated problems. He tests his proposed solution using 1998 Miami-Dade County tax assessor records and complementary data and argues that the resulting population and household estimates compare very favorably to estimates sold by national demographic data vendors.

Myers, D. (Ed.). 1990. *Housing Demography: Linking Demographic Structure and Housing Markets.* Madison, WI: University of Wisconsin Press. This book consists of 13 papers selected to bridge the gap between the separate analysis of population data and housing data in the United States. The approach is interdisciplinary, involving demography, economics, geography, sociology, and urban planning. A major objective of the book is theory building. The papers are grouped under three main headings: linking housing characteristics with household composition, life course and cohort models of housing choice, and housing consumption among the elderly.

Myers, D., and A. Doyle. 1990. "Age-specific Population-per-Household Ratios: Linking Population Age Structure with Housing Characteristics." pp. 109-130 in D. Myers (ed.) *Housing Demography: Linking Demographic Structure and Housing Markets*. Madison, WI: University of Wisconsin Press. The authors develop a demographic framework for determining Persons per household and propose models for estimating these values with the age structure of the household population.

Perrone, S. 2008. "Address Coverage and Improvement and Evaluation Program – 2005 National Estimate of Coverage of the Master Address File." pp. 37-85 in S. Murdock and D. Swanson (Eds.) *Applied Demography in the 21st Century*. Dordrecht, The Netherlands, Springer. The author reports on the Master Address File (MAF) portion of the U.S. Census Bureau's 'Address Coverage Improvement and Evaluation Program.' She provides annual calculations, 2000-2005, of the national coverage of housing unit counts and updates the 2000 baseline accordingly. She compares the estimated true housing unit count to annual housing counts from the MAF. She concludes that maintaining an accurate list of addresses continues to be a challenge.

Pitkin, J. 1992. "A Comparison of Vendor Estimates of Population and Households With 1990 Census Counts in California." *Applied Demography* 7(1):5-8. This study compares the 1990 U.S. population estimates of four national demographic data vendors with the actual counts of the 1990 Census. It was performed for a consortium of three public utility companies in California. In addition to total population, this study evaluates the estimates of three other demographic variables: (1) number of households; (2) population of Spanish origin; and ((3) white population. Differences in accuracy are found among the vendors and between variables. On average, the estimates were most accurate for total population, considerably less so for race (white population) and households and by far the least accurate for Hispanic population.

Pittenger, D. 2004. "Use of Census Data and City Housing Reports in Creating Housing-Unit Demolition Rates." *Population Estimates and Projections Research Brief No. 28*. Olympia, WA: Washington State Office of Financial Management. The author assesses if census data can be used as a possible source for creating rates of housing unit losses in areas where losses (e.g., demolitions) are not reported. He finds that because of the ignorance of many census respondents' to accurately report "year structure built", that census-based housing cohort loss data are not reliable either for benchmarking housing stock by age or for creating loss rates by age of unit. He concludes that for the purposes of estimating losses for county of municipal sub-areas in Washington, it seems better to use or adapt rates based on city demolition reports.

Prevost, R. 1994. "State Housing Unit and Household Estimates: April 1, 1980, to July 1, 1993. *Current Population Reports, Series P-25, No.* 1123. Washington, D: US Government Printing Office. A description of methods, data, procedures, and results for the state level estimates of housing units done the period from 1980 to 1993.

Reese, A. J. 2006. "A Comparison of Housing Unit Estimates to the American Community Survey's Aggregated Master Address File." Paper Presented at the Annual meeting of the Southern Demographic Association, Durham, NC. This report presents a comparison of the estimates of housing units, produced by the Population Division of the U.S. Census Bureau, to the number of valid units in the Master Address File (MAF), at the national, state, and county levels of geography for 2002 through 2005. Geographic patterns based on differences between the two datasets were detected and discussed in an attempt to gain a deeper level of understanding of the differences between the two sources of data on the housing stock.

Roe, L., J. Carlson and D. Swanson. 1992. "A Variation of the Housing Unit Method for Estimating the Population of Small, Rural Areas: A Case Study of the Local Expert Procedure." *Survey Methodology* 18 (1): 155-163. The authors report on a random sampling study based on an adaptation of the Housing Unit Method and the local expert method to determine the socioeconomic features of three unincorporated rural communities near Yucca Mountain, Nevada. A sample of the study area was selected from a carefully screened fame comprised of electrical utility data. Meter readers from the local utility companies were the local experts and two of them worked together to authenticate the accuracy of recorded data which included number of person in the household as of July 15, 1990 and age and gender of each member. Data accuracy was tested and it was found that the 1990 US Census counts were within the relatively narrow

95% confidence intervals. The mean width was 7.2% of the estimated population, thus the authors conclude that the estimates were meaningful.

Smith, S. K. 1995. "Housing Unit Method." pp. 89-117 in N. W. Rives Jr., W. J. Serow, A. S. Lee, H. F. Goldsmith, & P. R. Voss (Eds.) *Basic Methods for Preparing Small-area Population Estimates*. Madison, WI: Applied Population Laboratory, University of Wisconsin. This chapter describes the general framework of the Housing Unit Method (HUM) and discusses the strengths and weaknesses of a n umber of the techniques and data sources that can be used with it. It provides examples and empirical evidence regarding the accuracy of different types of HUM estimates. It concludes with observations on the usefulness of the HUM in making local population estimates.

Smith, S. K. 1986. "A Review and Evaluation of the Housing Unit Method of Population Estimation." *Journal of the American Statistical Association* 81:287-96. The author notes that the housing unit (HU) method is used by public and private agencies throughout the United States to make local population estimates. He describes many of the different types of data and techniques that can be used in applying the HU method, and it discusses the strengths and weaknesses of each. Empirical evidence from four different states is provided, comparing the accuracy of HU population estimates with the accuracy of other commonly used estimation techniques. Several conclusions are drawn regarding the usefulness of the HU method for local population estimation.

Smith, S. K., and S. Cody. 2004. "An Evaluation of Population Estimates in Florida: April 1, 2000." *Population Research and Policy Review* 23(1): 1–24. The authors investigate the influence of differences in population size and growth rate on estimation errors; compare the accuracy of several alternative techniques for estimating each of the major components of the Housing Unit Method (HUM); compare the accuracy of 2000 estimates with that of estimates produced in 1980 and 1990; compare the accuracy of HUM population estimates with that of estimates derived from other estimation methods; consider the role of professional judgment and the use of averaging in the construction of population estimates; and explore the impact of controlling one set of estimates to another. Their results confirm a number of findings that have been reported before and provide empirical evidence on several issues that have received little attention in the literature. They conclude with several observations regarding future directions in population estimation research.

Smith, S. K., and M. Mandell. 1984. "A Comparison of Population Estimation Methods: Housing Unit versus Component II, Ratio-Correlation, and Administrative Records." *Journal of the American Statistical Association* 79: 282-289. The authors observe that the Housing Unit Method (HUM) is often characterized as inferior to other methods for estimating the population of states and local areas. They challenge this characterization and in this paper evaluate population estimates produced by the housing unit method and by three other commonly used methods: component II, ratio correlation, and administrative records. Basing their analysis on 1980 census data from 67 counties in Florida and testing for precision, bias, and the distribution of errors, the authors find that

their application of the HUM performs at least as well as the more highly acclaimed methods of local population estimation.

Smith, S. K., and S. Cody.1994. “Evaluating the Housing Unit Method: A Case Study of 1990 Population Estimates in Florida.” *Journal of the American Planning Association*. 60: 209-221. The authors observe that the housing unit method (HUM) is the most commonly used approach to making small-area population estimates in the US. They evaluate the accuracy and bias of HUM-based population estimates produced for counties and sub-county areas in Florida for April 1, 1990. They find that population size has a negative effect on estimation errors (disregarding sign) but no effect on bias; growth rates have a U-shaped effect on estimation errors (disregarding sign) and a negative effect on bias; electric customer data provide more accurate household estimates than building permit data; errors in household estimates contribute more to population estimation error than do errors in estimates of average household size or group quarters population; and the application of professional judgment improves the accuracy of purely mechanical techniques. The authors argue the HUM offers a number of advantages over other population estimation methods and provides planners and demographers with a powerful tool for small-area analysis.

Smith, S. K., and B. Lewis. 1983. "Some New Techniques for Applying the Housing Unit Method of Local Population Estimation: Further Evidence." *Demography* 20: 407-413. The authors note that the housing unit method (HUM) of population estimation is often characterized as being imprecise and having an upward bias. In an earlier paper, the authors argued that the method itself cannot be properly characterized by a particular level of precision or direction of bias. Only specific techniques of applying the new method can have such characteristics. In that paper, the authors present several new techniques for estimating the number of households and average number of persons per household (PPH). However, the testing of these new techniques was limited by the lack of census results against which the estimates could be compared. Complete census data on population, households, and PPH are available and can be used to test alternate estimation techniques. The tests reported in the earlier paper using 1980 census data for Florida's 67 counties are replicated and the authors argue that this provide further evidence that the new techniques produce more precise, less biased estimates than previously used techniques.

Smith, S. K., and B. Lewis. 1980. “Some New Techniques for Applying the Housing Unit Method of Local Population Estimation.” *Demography* 17 (3): 323-339. The authors argue that the housing unit method of population estimation is often characterized as being imprecise and having an upward bias. They believe that the method itself cannot properly be characterized by a particular level of precision or direction of bias. Only specific techniques of applying the method can have such characteristics. In this paper, the authors discuss several new techniques they have developed for estimating households and the average number of persons per household. The compare estimates produced by these techniques to estimates produced by several other techniques and find that special census results from Florida provide preliminary evidence that the new techniques produce more precise, less biased estimates than the other techniques.

Smith, S. K., J. Nogle, and S. Cody. 2002. “A Regression Approach to Estimating the Average Number of Persons per Household.” *Demography* 39(4): 697-712. The authors develop several regression models in which PPH estimates were based on symptomatic indicators of PPH change. They tested these estimates using county level data in four states and found them to be more precise and less biased than estimates based on more commonly used methods.

Starsinic, D., and M. Zitter. 1968. “Accuracy of the Housing Unit Method in Preparing Population Estimates for Cities.” *Demography* 5: 475-484. This paper reports the results of a test of the relative accuracy of the housing unit method in the estimation of the population of cities. Estimates were prepared for 47 cities in excess of 50,000 population in which special censuses were conducted during the years 1964-66. The test points up five features of the housing unit method for the estimation of the population of cities. (1) In general, the method yields estimates on the high side. When building-permit data were used as a basis for estimates, the deviations were positive in about 30 of the 47 cases. Of the cities in which the deviations were negative, about one-half had had substantial annexation after 1960. (2) The use of utility data instead of building-permit data generally reduces the size of errors, although here too there are substantially more positive than negative deviations. (3) Deviations are smaller when the average size of household is extrapolated than when the 1960 values are used. (4) When either building-permit or utility data were used, the average error in the estimate of the number of households was high. (5) The estimate of the number of households is a greater contributor to errors in estimates of population than is the estimate of the average size of household (for this test, extrapolated from 1950-60 values or assuming no change since 1960). Although the scope of the test was limited, the method made a relatively creditable showing, with average errors of 3.6 to 5.8 percent, excluding areas that are experiencing large annexations.

Stone, G. 2008. “Methods for Measuring the Population After a Disaster: Household Population Surveys in Post-Katrina New Orleans, October 2005 – October, 2006.” pp. 113-123 in S. Murdock and D. Swanson (Eds.). *Applied Demography in the 21st Century*. Dordrecht, The Netherlands, Springer. The author describes the use of three difference sampling designs to conduct household surveys in Post-Katrina New Orleans. He discusses their limitations, cost and time requirements and the usefulness of the estimates they produced.

Swanson, D., and G. Hough. 2007. An Evaluation of Persons Per Household (PPH) Data Generated By the American Community Survey: A Demographic Perspective.” Paper presented at the Annual Meeting of the Southern Demographic Association. Birmingham, AL. This paper explores the usability of ACS data by examining “Persons Per Household (PPH), a variable of high interest to demographers and others preparing regular post-censal population estimates. The data used in this exploration are taken from 18 of the counties that formed the set of 1999 ACS test sites. The examination proceeds by comparing ACS PPH values to PPH values generated using a geometric model based on PPH change between the 1990 and 2000 census counts. The ACS PPH values represent what could be called the "statistical perspective" because variations in the values of specific variables over time and space are viewed largely by statisticians with an eye toward sample (and non-sample) error. The model-based PPH values represent a "demographic perspective" because PPH values are largely viewed by demographers as varying systematically, an orientation stemming from theory. The results suggest that the ACS PPH values lack sufficient temporal consistency to be used by demographers.

Swanson, D., and J. McKibben. 2007. "New Directions in the Development of Population Estimates and Projections." Paper presented at the Satellite Conference on Small Area Statistics, 56th Conference of the International Statistical Institute, Pisa, Italy. The authors argue that a continuously updated Master Address File argue that the MAF can be uses to develop population estimates and projections if it is properly enhanced (EMAF). The authors describe a set of activities needed to develop EMAF and how EMAF data could be directly assessed for statistical uncertainty and its use as a basis for developing population projections containing a wide range of ascribed (e.g., age and sex) and achieved characteristics (e.g., educational attainment and employment). The authors point out that such a development would bring the US Census Bureau's small area population estimation programs more in line with its European counterparts, but that there are several important challenges that must be surmounted, including issues of public trust, confidentiality, and tradition.

Swanson, D. J. Carlson, L Roe, and C. Williams. 1995. "Estimating the Population of Rural Communities by Age and Gender: a Case Study of the Effectiveness of the Local Expert Procedure." *Small Town* 25 (6): 14-21. This article examines the accuracy of a survey-based technique called the Local Expert Procedure for estimating selected demographic characteristics of small, rural areas. The procedure employs local citizens to provide demographic information about households which were randomly selected from a residential sample extracted from utility records. The procedure is used to provide age and sex estimates for the population of three communities near Yucca Mountain, Nevada.

Swanson, D. A., B. Baker, and J. Van Patten. 1983. "Municipal Population Estimation: Practical and Conceptual Features of the Housing Unit Method." Paper presented at the annual meeting of the Population Association of America, Pittsburgh, PA. This paper argues that the housing unit method is a logical choice when municipalities can challenge estimates via a special census. In the case where a state demographic center has the capacity to assist municipalities in conducting special censuses and monitor the results, a special census will serve to update the components of the housing unit method used for estimation. The authors note that this feature is not available with other methods of estimation such as Component Method II, ratio-correlation, and Tax returns.

Swanson, D. A., D. Randall, and L. Weisser. 1977. "An Evaluation of the 1977 Round of Municipal Vacancy Estimates for 1 and 2 Unit Structures." *Staff Document No. 42*. Olympia, WA: Washington State Office of Financial Management. This paper assesses the methods (sampling by block) and procedures (field work) used to develop vacancy estimates and compares the results with the vacancy rates found in cities that did special censuses. The methods and procedures were found to perform well, but suggestions were made for refining the estimates using cluster analyses of municipalities and special census results to inform current estimates.

Swanson, D. A., Morrison, P., Sharkova, I., Tayman, J., and Popoff, C. 1999. *Evaluation of the Washoe County Population Estimation System (WCPEM), and Recommendations for Improving the County's Estimation Program*. Report submitted by S.A.I.C., Inc. to the Washoe County Department of Community Development, Reno, Nevada. In responding to RFP 2118-99, Science Applications International Corporation (SAIC) served as a demographic consultant to Washoe County to assist in the evaluation of its population estimates program and establishment of the County's "in-house" demographic capabilities, particularly to support the County's efforts in postcensal redistricting of county commission districts. This report is the deliverable product called for in the contract. Consonant with the County's recently-issued MIS Vision and Mission Statement, our recommendations on population estimation have a strategic aim: to make full and ongoing use of the County's data assets to enhance county-wide decision making. Our evaluation and recommendations are aligned specifically with the Mission Statement's mandate to "manage data as an asset and strategic resource of the County, ensuring its security and availability."

Tayman, J. 1994. "Estimating Population, Housing and Employment for Micro-Geographic Areas." pp. 101-107 in K. V. Rao and J. W. Wicks (Eds.) *Studies in Applied Demography, 1994*. Bowling Green, OH: Population and Society Research Center: Bowling Green State University. This paper described procedures for developing current estimates of population, housing and place-of-work employment for micro-geographic areas in San Diego County, California. It demonstrates that with the proper technology and innovative use of nontraditional data sources, it is possible to estimate these activities for sub-census tract geographic areas. The procedures represent an interdisciplinary approach to the problem, blending the theory and methods from demography and geography. Residential meter and wage and salary employer address-level files allow a reasonably accurate determination of the micro-geographic location of employment and residential activities.

Tayman, J. 1991. "Population and Housing Estimates for Micro-Geographic Areas: A Blend of Demographic and GIS Techniques." Presented at the annual meeting of the Population Association of America. Washington, D.C. This paper describes the tools used to develop population and housing estimates for micro-geographic areas from electric meter residence files, census tract data, and its subdivisions Traffic Analysis Zones (TAZ), gridcells forming the Master Geographic Reference Area (MGRA), and the GBF/DIME file and a GIS procedure called ADMATCH, which maps and matches addresses to a geographic element. The biases occurring from mismatches or missing meters did not cause a serious problem in estimating. Unique geographic areas can now be created using the procedure POLYGON-to-POLYGON Overlay where MGRA polygon boundaries are fit to the designated polygon boundaries. The procedure in ARC/INFO called NETWORK can generate estimates for an area 10 minutes from a given site. An example is given.

Tayman, J. 1986. "An Integrated System for Subcity Population Estimation." Presented at the annual meeting of the Population Association of America. San Francisco, CA. The author describes the integrated system implemented in the San Diego Region (San Diego County) for preparing census tract level household income estimates. He described the methods, data and procedures used in this system as well as the uses of the estimates generated by the system. He discusses future developments and plans.

Tayman, J., and C. Rynerson. 1997. "An Integrated System for Estimating Subcounty Household Income Distributions. Presented at the annual meeting of the Population Association of America. Washington, DC. The authors described the integrated system implemented in the San Diego Region (San Diego County) for preparing census tract level household income estimates. The estimates were done first for the region, then for 41 groups of census tracts known as sub-regional areas (SRAs), then for the region's 347 census tracts. The modified lognormal curve was employed with a different equation using the parameters of income class, PARM (reflecting constant terms), and EXP (a nonlinear adjustment parameter). To estimate the household income distribution, the parameters of median income, PARM, and EXP were needed. Using 1990 census median incomes and income distributions as inputs, an iterative method was used for 9250 combinations of PARM and EXP to find the parameters that minimize the error between the observed and estimated income distributions. For the investigation of historical trends in the parameters, 1980 census data from eight income ranges were also assembled. Inputs to the model included the base year data, the estimated parameters, control totals from an independent estimate of the number of households for the region, SRAs and census tracts, and adjustment factors calculated from a comparison of the estimated and actual 1990 household income distributions. The integrated household income estimation system was fully implemented by means of finding solutions to the issues of forming households from individual tax returns, residual and adjustment factors, the applicability of the modified lognormal curve, and extrapolating parameters. The internal and external consistency of the estimates was also evaluated.

U. S. Census Bureau, Population Division. 2007. "County Housing Unit Estimates, April 1, 2000 to July 1, 2006." Washington, DC: Population Division, U. S. Census Bureau. (available online, http://www.census.gov/popest/housing/files/HU-EST2006_US.CSV, last accessed January, 2008). This is the data set that was developed using the data, methods, and procedures described in the following citation.

U. S. Census Bureau. 2006. Methodology for County Housing Unit Estimates for Vintage 2006." Washington, DC: Population Division, U.S. Census Bureau(available online, http://www.census.gov/popest/topics/methodology/2006_hu_meth.html, last accessed January 2008). This report describes the methods, data, and procedures to develop sub-county population estimates by a housing unit method that uses housing unit change to distribute county population to sub-county areas. It notes that the state and county housing unit estimates are aggregations of these housing unit estimates and that Housing unit estimates use building permits, mobile home shipments, and estimates of housing unit loss to update housing unit change since the last census. Census counts of housing units are geographically updated each year to reflect legal changes reported in the

Boundary and Annexation Survey, Census corrections, and other administrative revisions. The 2006 data can be found via the link in the preceding citation.

U. S. Census Bureau. 2004. "The Census Bureau's Master Address File (MAF): Census 2000 Address List Basics." Washington, DC. Geography Division, U. S. Census Bureau (available online http://www.census.gov/geo/mod/maf_basics.pdf , last accessed January 2008). This report describes the content and procedures used to develop and maintain the MAF.

U. S. Census Bureau. 1990. "State Population and Household Estimates." *Current Population Reports, P25, No. 1058*. Washington, DC: U. S. Government Printing Office. This report contains population and household estimates for U.S. States for July 1, 1989, revised household estimates for July 1, 1988, revised annual estimates for the resident and civilian populations for July 1, 1981-88, and components of population change for the 1980-89 period. The annual estimates of the resident population are further disaggregated by sex and age. The age detail includes 10-year and selected broad age groups.

U. S. Census Bureau. 1987. "State Population and Household Estimates, with Age, Sex, and Components of Change: 1981-86." *Current Population Reports, P25, No. 1010*. Washington, DC: U. S. Government Printing Office. This report contains provisional estimates for States of the resident and civilian U.S. populations and of households for July 1, 1986, revised annual population and household estimates for July 1, 1981, through 1985, and components of population change for the 1980-86 period.

U. S. Census Bureau. 1986. "State Population and Household Estimates to 1985, With Age and Components of Change." *Current Population Reports, P25, NO. 998*. Washington, DC: U. S. Government Printing Office. This report contains provisional estimates of the resident and civilian populations of U.S. States for July 1, 1985, revised population estimates for July 1, 1981, through 1984, and components of population change and revised household estimates for States for the 1980-85 period. Annual inter-censal estimates by age for 1971 to 1979 are also presented. This is the first time that estimates of population and households for States have been included in the same issue of Current Population Reports.

U. S. Census Bureau. 1985. "Estimates of Households for States: 1981 to 1984." *Current Population Reports, P25. No. 974*. Washington, DC: U. S. Government Printing Office. This report presents estimates of the total number of households for U.S. States for July 1, 1981, 1982, and 1983, and provisional estimates for July 1, 1984. In addition to these estimates, census counts of households for 1970 and 1980 are included for comparative purposes.

U. S. Census Bureau. 1979. "Estimates of the Number of Households for States: 1977 and 1978." *Current Population Reports, P25, No. 807*. Washington, DC: U. S. Government Printing Office. This report presents estimates of the total number of households for U.S. States for July 1, 1977 and 1978.

U. S. Department of Commerce. 2002. "A Better Strategy is needed for Managing the Nation's Master Address File." *Inspection Report No. OSE-12065*. Washington, D. C.: U. S. Department of Commerce. This report found that Census did not establish procedures that would adequately ensure the MAF's accuracy and completeness, and thus used an error-prone database to conduct the decennial census in 2000. Among other problems, the report found that numerous addresses did not link to the correct location in the TIGER file, which hampered the Census Bureau's ability to accurately determine a housing unit's location. Some errors in the file occurred because address verification plans were flawed. In addition, the report finds that when the Census Bureau modified its software for eliminating duplicate addresses in an effort to improve its coverage of multiunit housing such as trailer parks and apartment buildings, the change allowed many suspected duplicate addresses to remain in the address file. The report also notes that maps often contained duplicate and missing streets, and were not always printed in a usable size and format.

U. S. GAO. 2006. *2010 Census: Census Bureau Needs to Take Prompt Actions to Resolve Long-standing and Emerging Address and Mapping Challenges*. GAO-06-272. Washington, D.C. U. S. Government Accountability Office (Note: Effective July 7, 2004, the GAO's legal name became the Government Accountability Office). This report found that the Census Bureau's address and map modernization efforts have progressed in some areas. The Bureau is researching how to correct addresses that were duplicated, missed, deleted, and incorrectly located on maps. However, some deadlines for completing research are not firm, while other deadlines that had been set continue to slip. Thus, whether research will be completed in enough time to allow the Bureau to develop new procedures to improve the 2010 address file is unknown. Also, the Bureau has not fully addressed emerging issues. For one such issue, the Bureau has acknowledged the compressed time frame for completing address canvassing--an operation where census workers walk every street in the country to verify addresses and maps--but has not reevaluated the associated schedule or staffing workloads. Also, the Bureau has allotted only 6 weeks to conduct address canvassing it completed in 18 weeks in 2000 and expanded the operation from urban areas in 2000 to the entire country in 2010. Whether the Bureau can collect and transmit address and mapping data using the MCD is unknown. The MCD, tested during 2006 address canvassing, was slow and locked up frequently. Bureau officials said the MCD's performance is an issue, but a new MCD to be developed through a contract awarded in March 2006 will be reliable. However, the MCD will not be tested until the 2008 Dress Rehearsal, and if problems emerge, little time will remain to develop, test, and incorporate refinements. If after the Dress Rehearsal the MCD is found unreliable, the Bureau could face the remote but daunting possibility of reverting to the costly paper-based census of 2000. Bureau officials do not believe a specific plan is needed to update the addresses and maps for areas affected by the hurricanes. Securing a count is difficult under normal conditions, and existing procedures may insufficient to update addresses and maps after the hurricanes' destruction--made even more difficult as streets, housing, and population will be in flux.

Wang, C. 1999. “Development of National Accounting of Address and Housing Inventory: The Baseline Information for Post-censal Population estimates.” Paper presented the Estimates Methods Conference, U. S. Bureau of the Census, Suitland, MD. (available online, http://www.census.gov/population/www/coop/popconf/paper.html, last accessed June 2007). The paper discusses the use of the Master Address File for a number of purposes, including the development of data that can be used for population estimates made using the Housing Unit Method.

Webster, C. 1996. “Population and Dwelling Unit Estimates from Space.” *Third World Planning Review* 18 (2): 155-76. This paper reports on attempts to measure the morphological patterns in an urban satellite scene and to use these for image interpretation. The interpretation task addressed is the estimation of residential dwelling units from the patterns discernible in high resolution satellite images of cities. The practical results include dwelling estimates that can be aggregated to any geographical unit of analysis, population estimates for cities and a dwelling density surface that can be categorized into any number of residential land-use classes.

Wetrogan, S. 2007. “Evaluating County Population Estimates Using a Housing Unit Based Approach.” Paper presented at the annual meeting of the Southern Demographic Association. Birmingham, AL. The paper compares four alternative 1990-based housing unit estimates for 2000 to Census 2000 population counts and to a separate set of 1990-based county population estimates for 2000 prepared using the administrative records component of change method.

Zeng, Y., K. Land, Z. Wang, and D. Gu. 2006. “U. S. family household momentum and dynamics: an extension and application of the ProFamy Method.” *Population Research and Policy Review* 25: 1–41. Unlike the headship-rate method, the ProFamy method uses demographic rates as input and can project detailed household types, sizes, and living arrangements for all members of the population. Tests of projections from 1990 to 2000 using ProFamy and based on observed U.S. demographic rates before 1991 show that discrepancies between ProFamy projections and census observations in 2000 are reasonably small, validating the new method.

III. Housing Unit Estimates

McDonald, J. and D. McMillen. 2000. "Residential Building Permits in Urban Counties: 1990–1997." *Journal of Housing Economics* 19 (3): 175-186. This paper presents simple empirical models of residential building permits for urban counties in the United States for the period 1990–1997. Building permits, as a percentage of the housing stock, are greater the larger are population growth, the proportion of units that are old (built before 1940), and the proportion of units that are new. A higher initial vacancy rate reduces building permits.

National Research Council. 2004. "Modernizing Geographic Resources." (Chapter 3) pp. 57-102 in *Re-engineering the 2010 Census: Risks and Challenges*. Washington, DC: National Academies Press. This chapter discusses the risk to Census 2010 if the "three legged" stool is not fully implemented, legs being the master address file (MAF) and TIGER. The chapter reviews MAF/TIGER enhancements and expresses concern about readiness for 2010.

Perrone, S. 2008. "Address Coverage and Improvement and Evaluation Program – 2005 National Estimate of Coverage of the Master Address File." pp. 37-85 in S. Murdock and D. Swanson (Eds.) *Applied Demography in the 21st Century*. Dordrecht, The Netherlands, Springer. The author reports on the Master Address File (MAF) portion of the U.S. Census Bureau's 'Address Coverage Improvement and Evaluation Program.' She provides annual calculations, 2000-2005, of the national coverage of housing unit counts and updates the 2000 baseline accordingly. She compares the estimated true housing unit count to annual housing counts from the MAF. She concludes that maintaining an accurate list of addresses continues to be a challenge.

Pittenger, D. 2004. "Use of Census Data and City Housing Reports in Creating Housing-Unit Demolition Rates." *Population Estimates and Projections Research Brief No. 28*. Olympia, WA: Washington State Office of Financial Management. The author assesses if census data can be used as a possible source for creating rates of housing unit losses in areas where losses (e.g., demolitions) are not reported. He finds that because of the ignorance of many census respondents' to accurately report "year structure built", that census-based housing cohort loss data are not reliable either for benchmarking housing stock by age or for creating loss rates by age of unit. He concludes that for the purposes of estimating losses for county of municipal sub-areas in Washington, it seems better to use or adapt rates based on city demolition reports.

Reese, A. J. 2006. "A Comparison of Housing Unit Estimates to the American Community Survey's Aggregated Master Address File." Paper Presented at the Annual meeting of the Southern Demographic Association, Durham, NC. This report presents a comparison of the estimates of housing units, produced by the Population Division of the U.S. Census Bureau, to the number of valid units in the Master Address File (MAF), at the national, state, and county levels of geography for 2002 through 2005. Geographic patterns based on differences between the two datasets were detected and discussed in an attempt to gain a deeper level of understanding of the differences between the two sources of data on the housing stock.

Swanson, D. A., J. McKibben. 2007. "New Directions in the Development of Population Estimates and Projections." Paper presented at the Satellite Conference on Small Area Statistics, 56th Conference of the International Statistical Institute, Pisa, Italy. Small area estimation programs of the US Census Bureau have taken a different direction than those found in the national statistical agencies of many European countries, where population and other registers are used to more effect. However, the advent of a continuously updated Master Area File (MAF) following the 2000 census represents an information resource that has not yet been fully tapped for purposes of developing timely, cost-effective, and precise population estimates and projections for even the smallest of geographical units (e.g., census blocks). The authors argue that the MAF can be enhanced (EMAF) for these purposes. In support of our argument, the authors describe a set of activities needed to develop EMAF, each of which is well within the current capabilities of the US Census Bureau and discuss various costs and benefits of each. They also describe how EMAF data could be directly assessed for statistical uncertainty and its use as a basis for developing population projections containing a wide range of ascribed (e.g., age and sex) and achieved characteristics (e.g., educational attainment and employment). Although such a development would bring the US Census Bureau's small area population estimation programs more in line with its European counterparts, there are several important challenges identified by the authors who describe that must be surmounted, including issues of public trust, confidentiality, and tradition.

Tayman, J. 1980. "Total and Occupied Housing Units as Useful Indicators in Regression Models for Population Estimation." Presented at the Spring FSCPE Meeting, Denver Colorado, April 8-9. Tayman constructs three ratio-correlation regression models based on 1960/1950 ratios for use in estimating the civilian population under 65 years for counties in Washington in 1970 (to compare results with1970 census data). He finds that by adding housing units to a model using voters, 1-8 enrollment, and covered employment the 1970 MAPE is reduced from 5.07% to 4.16% and that by using occupied housing units, the MAPE is reduced from 5.07% to 2.82%. Using the same general data and the ratio-correlation methodology, he then examines relative to 1970 censes data, the estimates from two more regression models, total housing units and occupied housing units. These two model allow him to estimate vacant units (total – occupied) and vacancy rates (vacant/total). In regard to the model for estimating total housing units, he finds that the MAPE is 5.61%, while for the model used to estimate occupied housing units, it is 4.15%. There are substantial differences by county in terms of estimated vacant units and vacancy rates, however, when he combines the results of the two regression models.

U. S. Census Bureau, Population Division. 2007. "County Housing Unit Estimates, April 1, 2000 to July 1, 2006." Washington, DC: Population Division, U. S. Census Bureau. (available online, http://www.census.gov/popest/housing/files/HU-EST2006_US.CSV, last accessed January, 2008). This is the data set that was developed using the data, methods, and procedures described in the following citation.

U. S. Census Bureau. 2006. Methodology for County Housing Unit Estimates for Vintage 2006." Washington, DC: Population Division, U.S. Census Bureau(available online, http://www.census.gov/popest/topics/methodology/2006_hu_meth.html, last accessed January 2008). This report describes the methods, data, and procedures to develop sub-county population estimates by a housing unit method that uses housing unit change to distribute county population to sub-county areas. It notes that the state and county housing unit estimates are aggregations of these housing unit estimates and that Housing unit estimates use building permits, mobile home shipments, and estimates of housing unit loss to update housing unit change since the last census. Census counts of housing units are geographically updated each year to reflect legal changes reported in the Boundary and Annexation Survey, Census corrections, and other administrative revisions. The 2006 data can be found via the link in the preceding citation.

U. S. Census Bureau. 2004. "The Census Bureau's Master Address File (MAF): Census 2000 Address List Basics." Washington, DC. Geography Division, U. S. Census Bureau (available online http://www.census.gov/geo/mod/maf_basics.pdf , last accessed January 2008). This report describes the content and procedures used to develop and maintain the MAF.

U. S. Department of Commerce. 2002. "A Better Strategy is needed for Managing the Nation's Master Address File." *Inspection Report No. OSE-12065*. Washington, D. C.: U. S. Department of Commerce. This report found that Census did not establish procedures that would adequately ensure the MAF's accuracy and completeness, and thus used an error-prone database to conduct the decennial census in 2000. Among other problems, the report found that numerous addresses did not link to the correct location in the TIGER file, which hampered the Census Bureau's ability to accurately determine a housing unit's location. Some errors in the file occurred because address verification plans were flawed. In addition, the report finds that when the Census Bureau modified its software for eliminating duplicate addresses in an effort to improve its coverage of multiunit housing such as trailer parks and apartment buildings, the change allowed many suspected duplicate addresses to remain in the address file. The report also notes that maps often contained duplicate and missing streets, and were not always printed in a usable size and format.

U. S. GAO. 2006. *2010 Census: Census Bureau Needs to Take Prompt Actions to Resolve Long-standing and Emerging Address and Mapping Challenges*. GAO-06-272. Washington, D.C. U. S. Government Accountability Office (Note: Effective July 7, 2004, the GAO's legal name became the Government Accountability Office). This report found that the Census Bureau's address and map modernization efforts have progressed in some areas. The Bureau is researching how to correct addresses that were duplicated, missed, deleted, and incorrectly located on maps. However, some deadlines for completing research are not firm, while other deadlines that had been set continue to slip. Thus, whether research will be completed in enough time to allow the Bureau to develop new procedures to improve the 2010 address file is unknown. Also, the Bureau has not fully addressed emerging issues. For one such issue, the Bureau has acknowledged the compressed time frame for completing address canvassing--an operation where census workers walk every street in the country to verify addresses and maps--but has not reevaluated the associated schedule or staffing workloads. Also, the Bureau has allotted only 6 weeks to conduct address canvassing it completed in 18 weeks in 2000 and expanded the operation from urban areas in 2000 to the entire country in 2010. Whether the Bureau can collect and transmit address and mapping data using the MCD is unknown. The MCD, tested during 2006 address canvassing, was slow and locked up frequently. Bureau officials said the MCD's performance is an issue, but a new MCD to be developed through a contract awarded in March 2006 will be reliable. However, the MCD will not be tested until the 2008 Dress Rehearsal, and if problems emerge, little time will remain to develop, test, and incorporate refinements. If after the Dress Rehearsal the MCD is found unreliable, the Bureau could face the remote but daunting possibility of reverting to the costly paper-based census of 2000. Bureau officials do not believe a specific plan is needed to update the addresses and maps for areas affected by the hurricanes. Securing a count is difficult under normal conditions, and existing procedures may insufficient to update addresses and maps after the hurricanes' destruction--made even more difficult as streets, housing, and population will be in flux.

Wang, C. 1999. "Development of National Accounting of Address and Housing Inventory: The Baseline Information for Post-censal Population estimates." Paper prepared for the Estimates Methods Conference, U.S. Bureau of the Census, Federal Office Building #3, Suitland, Maryland, June 8th. (available online at http://www.census.gov/population/www/coop/popconf/paper.html, last accessed 28 June 2007).

IV. Vacancy Rate Estimates

Brown, W. 1999. " Use of Property Tax Records and Household Composition Matrices to Improve the Household Units Method for Small Area Population Estimates." Paper presented at the Population Estimates Conference, U. S. Census Bureau. Suitland, MD. The Housing Unit Method for estimating small area population can be improved in three ways: improved counts of housing units; identification of geographic location of individual housing units; and refined estimates of persons per household. This paper reports on a pilot effort in a single county to use property tax records in place of housing permits for building and demolition to improve counts of housing units. In addition the property tax records include an x-y coordinate for spatially locating the housing unit structure. This enables the housing units to be geocoded to Census Blocks, as well as to non-Census geographic areas such as watersheds. Finally, the use of household composition matrices yields improved estimates of persons per household for small areas. The household composition matrices contain a refined measure of persons per household built from cells of persons per household cells by age of person and age of householder. The improved housing unit method is demonstrated for Census Block groups.

Edelstein, R. and D. Tang. 2007. "Dynamic Residential Housing Cycles Analysis." *The Journal of Real Estate Finance and Economics* 35 (3): 295-313. This paper develops and tests a theoretical model for residential housing market cyclical dynamics. The model employs an interactive supply and demand framework to engender housing price dynamics. Under our set of assumptions, the two equation system is econometrically identified: the first equation, housing demand, relates rent, property values, and capitalization rates with demand fundamentals. The second equation, housing supply, relates housing investment and property values with supply fundamentals. Using the model, the authors analyze empirically the cyclical dynamics for residential properties in Los Angeles, San Francisco, San Diego and Sacramento for the 1988–2003 time period. The theoretical and econometric design represents improvements and/or modifications of previous studies in at least four ways. First, many of the earlier commercial cyclical analyses have focused on office appraisal and have relied on sparse transactions data, which are likely to be less reliable than the copious amount of residential transactions data. Second, the cyclical volatility and timing of single-family housing is different than that of commercial real estate. Third, by examining different local MSA markets in California, our study distinguishes and isolates national-macro, regional and local market variable effects upon cycles. Finally, utilizing quarterly data (versus annual data) sharpens our ability to focus upon cyclical behavior. The authors argue that their empirical analyses suggest that fundamentals, such as employment growth and interest rates are key determinants of the residential real estate cycles. However, in general, local fundamentals tend to have greater cyclical impacts than those of national or regional fundamentals.

Gabriel, S. A., and F. E. Nothaft. 2001. "Rental Housing Markets, The Incidence and Duration of Vacancy, and the Natural Vacancy Rate." *Journal of Urban Economics* 49 (1): 121-49. The authors note that new intermetropolitan and time-series data from the BLS are used to derive and model the incidence and duration of rental vacancies and to assess their importance to the price adjustment mechanism for rental housing. Research findings indicate that duration varies with measures of MSA housing costs and housing stock heterogeneity, while incidence varies with measures of population mobility, public housing availability, and population growth. Results support a more general specification of rental price adjustment in which the rate of real rent change reflects deviations in observed vacancy incidence and duration from their equilibrium levels. They provide new estimates of equilibrium vacancy rates for a large set of metropolitan areas over the 1987–1996 period.

Gabriel, S. A., and F. E. Nothaft. 1988. "Rental Housing Markets and the Natural Vacancy Rate." *American Real Estate and Urban Economics Association Journal* 16(4): 419-429. This paper employs new census vacancy rate data to analyze the price-adjustment mechanism for rental housing. The study extends previous research on this topic, which provided conflicting evidence concerning the traditional theory of rental housing market adjustment. Cross-section and time-series data are pooled to estimate natural vacancy rates for sixteen United States cities for the 1981–85 period. The analysis further explores the determinants of variation in natural vacancy rates across those metropolitan areas.

Hendershott, P. H., B. D. MacGregor and R. Y. C. Tse. 2002. "Estimation of The Rental Adjustment Process." *Real Estate Economics* 30 (2):165-183. Rental adjustment equations have been estimated for a quarter century. In the U.S., models have used the deviation of the actual vacancy rate from the natural rate as the main explanatory variable, while in the UK, drivers of the demand for space have dominated the estimation. The recent papers of Hendershott (1996) and Hendershott, Lizieri and Matysiak (HLM, 1999) fall into the former category. They re-estimate these equations using alternative formulations but can do little to improve them overall. However, they identify econometric concerns with the specifications. The authors then derive a model incorporating both supply and demand factors within an Error Correction framework, and show how the U.S. and UK traditions are special cases of this more general formulation. They next estimate this equation using data from the City of London office market. Our initial specification of this more generalized model is greatly superior to the vacancy rate model. Finally, they estimate a two-equation variant with a separate vacancy rate equation; this model also performs much better than that of HLM. Importantly, our model passes standard modern econometric requirements for unit roots and co-integration. The authors find little evidence of special or temporal variation in natural vacancy rates.

Hsueh, L., H. Tseng, and C. Hsieh. 2007.. "Relationship Between the Housing Vacancy Rate, Housing Price, and the Moving Rate at the Township Level in Taiwan, in 1990 and 2000." *International Real Estate Review* 10 (1): 119-150. In this research, cross-sectional data for the township level obtained from the 1990 and 2000 Population and Housing Census are used to study the phenomenon of high housing vacancy rates in

Taiwan. Three simultaneous equations for housing price, vacancy rate, and moving rate are derived and estimated using 3SLS. The estimation results show that, in 1990, in a booming market situation, both expected housing price and current housing price had a strong, positive impact on the vacancy rate; however, the housing vacancy rate did not display a negative impact on housing price as expected. The results for 2000 show that housing price did not significantly affect the vacancy rate; however, the vacancy rate had a negative impact on housing price that was highly statistically significant. This result reflected the fact that housing market operation had swung to another extreme after the real estate bubble that started in the late 1980s and burst in the mid-1990s. The natural vacancy rate for each township can be obtained from the estimation results. The average rate for 2000 was 0.11 to 0.12, compared to an actual vacancy rate of 0.158, which implied that 75% of townships had an excess supply of housing. Only Taipei City, Kaohsiung City and townships in areas inhabited by Taiwan's indigenous peoples had, on average, a relatively low excess supply rate.

Khor, A., Y Ming, and L. Yuan. 2000. "The Natural Vacancy Rate of the Singapore Office Market." *Journal of Property Research* 17 (4): 329-338. The concept of a natural vacancy rate is relatively well established in the real estate literature. The natural vacancy rate is an equilibrium level of inventory of space, in the sense that both the matching process between landlord and tenant is facilitated, and that building owners hold an optimal buffer stock of inventory to meet future leasing contingencies. It is the current deviation from the natural vacancy rate (and not the absolute level of the current vacancy rate) which determines the degree to which the given market is in or out of equilibrium. When vacancy rates are above the natural vacancy rate, rents will fall and vacancies will drift upward toward equilibrium. The determination of the natural vacancy rate is therefore significant in that it can facilitate the monitoring of the market conditions since a vacancy rate below the natural vacancy rate signifies tight market. The converse is true if the vacancy rate is above the equilibrium level. The natural vacancy rate for the office space market in Singapore over the sample period 1979-1997 is found to be between 10% and 12% depending on the model used. These models have been selected based on their R2, D-W and other relevant statistics.

Lowe, T., et al. 2000. "Occupied versus Rented: Census and Real Estate Occupancy Rates and their use in Population Estimates." *Population Estimates and Projections Research Brief No. 9*. Olympia, WA: Washington State Office of Financial Management. This Brief identifies differences in the definition of "occupied" as used by the federal Bureau of the Census and local Real Estate Surveys. It also provides a simple procedure for adjusting the real estate survey data for use in the estimation process.

Lowe, T. 1988. "A Resurrection: The Potential of Postal Survey Data in Improving Housing Unit Population Estimates for Local Areas." Paper presented at the annual meeting of the Population Association of America., New Orleans, LA. The accuracy of postal survey data in reporting residential housing unit occupancy estimates was compared to that of 1970 and 1980 U.S. census data for 26 Washington State cities. These postal surveys were conducted by the Federal Department of Housing and Urban Development in the 1970s within 2 months of collection of census data. Results are given in terms of absolute error of vacancy rates. Postal surveys almost always show lower vacancy rates than census data because they do not include unfinished or new units, or concealed unoccupied conversions in single family homes. Suburban single family housing generally has the highest occupancy rates. Postal data are much more accurate than census data when occupancy rates are variable, as in cities near military bases, and in multi-unit structures. This variability precludes the use of an adjustment factor in most cases. Adjusted postal occupancy rates were more accurate when city and postal boundaries corresponded and remained stable. The results of this study indicate the potential value of postal survey data when occupancy rates vary.

Lowe, T., and M. Mohrman 2003. "Use of Postal Delivery Data in the Population Estimate Process." *Population Estimates and Projections Research Brief No. 18.* Olympia, WA: Washington State Office of Financial Management. This Brief extends the evaluation of postal vacancy/occupancy rates presented in Research Brief No.17. It examines the relationship between vacancy trends in postal data versus census data, describes how the postal rates were used in the 2003 estimation process using the Housing Unit Method, and identifies avenues for further evaluation.

Lowe, T., L. Weisser, and B. Myers. 1984. "A Special Consideration in Improving Housing Unit Estimates: The Interaction Effect. Paper presented at the annual meeting of the Population Association of America, Minneapolis, MN. The authors observe that of some of the terms in the equation for the Housing Unit Method are themselves correlated and discuss how this can be dealt with in using the Housing Unit method to estimate population. They specifically point to the interaction of average household size and structure type.

Lowe, T., M. Mohrman, and D. Brunink. 2003. "Developing Postal Delivery Data for use in Population Estimates." *Population Estimates and Projections Research Brief No. 17.* Olympia, WA: Washington State Office of Financial Management. This case study underscores the problems inherent in developing address lists for census and sample purposes. The findings in this case study led the authors to caution that the ACS may have substantial coverage errors unless a great deal attention is paid to its frame, an extract of the Master Address File.

Rosen, K., and L. Smith. 1983. "The Price-Adjustment Process for Rental Housing and the Natural Vacancy Rate." *The American Economic Review* 73 (4): 779-786. the authors confirm the view that rental price changes are significantly affected by excess supply or demand in the rental market for 17 U.S. cities. They use a pooled-time series, cross-sectional analysis to demonstrate that a close relationship exists between variations in rental price and deviations of actual vacancy rates from normal vacancy rates.

Smith, S. K. and M. House. 2007. "Temporary Migration: a Case Study of Florida." *Population Research and Policy Review* 26 (4): 437-454. Most migration statistics in the United States focus on changes in permanent residence, thereby missing temporary moves such as the daily commute to work, business trips, vacations, and seasonal migration. In this paper, the authors analyze temporary migration streams in Florida,

focusing on moves that include an extended stay. Using several types of survey data, they examine the characteristics of non-Floridians who spend part of the year in Florida and Floridians who spend part of the year elsewhere. The authors develop estimates of the number, timing, and duration of temporary moves and the origins, destinations, characteristics, and motivations of temporary migrants. This study presents the most comprehensive analysis yet of temporary migration in Florida and provides a model that can be used in other places. It also points to a serious shortcoming in the US statistical system, namely, the lack of information on temporary migration streams. They believe that the American Community Survey provides an opportunity to remedy this problem

Swanson, D. 1977."An Evaluation of the 1977 Round of Municipal Vacancy Estimates for 1 & 2 Unit Structures." *Staff Document No. 42*. Office of Financial Management, Population, Enrollment and Economic Studies Division. Olympia, WA: Washington State Office of Financial Management. The report compares the results of sample vacancy estimates for 45 Washington municipalities done using the windshield survey method with the vacancy rate values found in special headcount censuses done in the same municipalities. The author examines unadjusted, ratio-adjusted, and regression-adjusted VR values. In 29 of the 45 municipalities, the VA values found from windshield survey samples were less than the VR found in the census. In general, Swanson found that the regression method provided the most balanced estimates and restricted the range of VR errors to a smaller interval than either the ratio-adjusted method or the unadjusted VR estimates. One finding of interest is that there was a positive correlation between size of error and city population for all three VR estimation methods.

Tayman, J. 1980. "Total and Occupied Housing Units as Useful Indicators in Regression Models for Population Estimation." Presented at the Spring FSCPE Meeting, Denver Colorado, April 8-9. Tayman constructs three ratio-correlation regression models based on 1960/1950 ratios for use in estimating the civilian population under 65 years for counties in Washington in 1970 (to compare results with1970 census data). He finds that by adding housing units to a model using voters, 1-8 enrollment, and covered employment the 1970 MAPE is reduced from 5.07% to 4.16% and that by using occupied housing units, the MAPE is reduced from 5.07% to 2.82%. Using the same general data and the ratio-correlation methodology, he then examines relative to 1970 censes data, the estimates from two more regression models, total housing units and occupied housing units. These two model allow him to estimate vacant units (total – occupied) and vacancy rates (vacant/total). In regard to the model for estimating total housing units, he finds that the MAPE is 5.61%, while for the model used to estimate occupied housing units, it is 4.15%. There are substantial differences by county in terms of estimated vacant units and vacancy rates, however, when he combines the results of the two regression models.

Weidman, L., R. Creecy, D. Malec, and J. Tsay. 2008. "Exploration of the Use of Empirical Bayes Procedures for Estimating Changes in Occupancy Rate and Persons Per Household." *Research Report Series, Statistics #2008-07.* Statistical Research Division. Washington, DC: U.S. Bureau of the Census. The authors explore the use of an empirical Bayes approach as a means of updating PPH and VR values from the most recent Decennial Census in conjunction with ACS values. They examine national and state models using 1990 and 2000 long form data (to represent ACS values), while 1990 and 2000 short form data are used as the independent variables. The question they are aiming at ultimately answering is: Should statistically significant PPH and VR values from the ACS be used to estimate change over time in PPH and VR? Their work suggests several Research projects that need to be conducted before their question can be answered.

V. PPH Estimates

.

Acs, G., and S. Nelson. 2004. "Changes in living arrangements in the late 1990s: Do welfare policies matter?" *Journal of Policy Analysis and Management* 23(2): 273–290. Using data from the first two waves of the National Survey of America's Families (NSAF)1, the authors examine how living arrangements for families with children changed between 1997 and 1999. They find that the share of families composed of a single mother living independently declined; at the same time, the share of families composed of cohabiting couples with children rose. In addition, they report larger changes in living arrangements among the population subgroups most likely to be affected by welfare policies, lower-income and less-educated families, than among other subgroups, such as moderate-income families. The authors suggest that welfare policies may have contributed to the decline in single parenting and the rise in cohabitation between 1997 and 1999.

Akkerman, A. 1980. "On the Relationship between Household Composition and Population Age Distribution." *Population Studies* 34 (3): 525-534. Composition of households by age of head and by age of other household members has recently been presented in a convenient algebraic expression, the household composition matrix. It has been shown that this matrix operates as a linear transformation from the vector of household distribution by age of head to the vector of population age distribution. A further analysis will show that the first row of the matrix may be interpreted as representing a vector of average household fertility rates. If the linear relationship between household and population distributions is fully implemented, then a relationship between household fertility and the size of the youngest age group can be derived. Extension of this result will enable simultaneous projection of population and households, suitable for computer application to conventional five-year age groups.

Akkerman, A. 2000. "On the Leontief Structure of Household Populations." *Canadian Studies in Population* 27(1): 181-194. In a given population, the author considers the age distribution of all persons in households and the age distribution of household-heads or household-markers. He shows that the formal relationship holding between the two age distributions is equivalent to the input-output relationship in the Leontief model of the open economy. The notions of household composition and household accommodation which have emerged independently over the past two decades, are shown to be formally linked within this relationship.

Akkerman, A. 2004. "Age-specific Household Size as a Demographic Aspect of Regional Disparity: Czech Republic, 1991." *Canadian Studies in Population* 31(2): 237-260. The post-communist transition to market economy in Central Europe over the last decade of the twentieth century had a significant impact on the demographic profile of the former Soviet bloc countries. Largely due to government policy and market conditions related to housing, this observation is particularly true for the Czech Republic. The present study shows housing as a facet of regional demographic differences within the Czech Republic. The household composition matrix is applied here as a demographic gauge to the behavioral response of households to Czech housing markets and policy. The matrix provides here a glance at households' demographic behavior in the capital city of Prague and in the country's other regions, during the early transition period, based on observations from the 1991 census. A summary feature of household composition is the age-specific household size shown for the various regions of the Czech Republic to trace the reduced standard Gamma function. Anomalies detected in the trajectory of age-specific household size for Prague confirm the unique housing market conditions in the capital city, and point to a commensurate demographic response in Prague as opposed to the rest of the country.

Bell, M. and J. Cooper. 1990. "Household Forecasting: Replacing the Headship Rate Model." Paper presented at the Fifth National Conference, Australian Population Association, Melbourne, November. This is one of the earliest criticisms of the shortcomings of the classic headship-rate approach. It notes that the headship-rate method does not link easily to population processes and is generally inadequate for planning purposes.

Brown, W. 1999. "Use of Property Tax Records and Household Composition Matrices to Improve the Household Units Method for Small Area Population Estimates." Paper presented at the Population Estimates Conference, U. S. Census Bureau. Suitland, MD. The Housing Unit Method for estimating small area population can be improved in three ways: improved counts of housing units; identification of geographic location of individual housing units; and refined estimates of persons per household. This paper reports on a pilot effort in a single county to use property tax records in place of housing permits for building and demolition to improve counts of housing units. In addition the property tax records include an x-y coordinate for spatially locating the housing unit structure. This enables the housing units to be geocoded to Census Blocks, as well as to non-Census geographic areas such as watersheds. Finally, the use of household composition matrices yields improved estimates of persons per household for small areas. The household composition matrices contain a refined measure of persons per household built from cells of persons per household cells by age of person and age of householder. The improved housing unit method is demonstrated for Census Block groups.

Bumpass, L. 1990. "What's Happening to the Family? Interactions Between Demographic and Institutional Change." *Demography* 27(4): 483–498. In this article, which was the Presidential Address to the 1990 Annual Meeting of the Population Association of America, the author discusses recent changes affecting the family in the United States and implications for family structure in the future. He considers trends in marital disruption; cohabitation, family formation, and marriage; fertility; marital relationships; labor force needs; intergenerational relationships; and women's status.

Burch, T. 1967. "The Size and Structure of Families: A Comparative Analysis of Census Data." *American Sociological Review* 32 (3): 347-363. Recent hypotheses by Levy regarding the essential similarity of actual family structures in all societies are examined in the light of United Nations data on type and size of households, and in the light of more detailed census materials from selected nations, mostly in Latin America. Some support is found for his assertion that in no society has the very large residential family become the modal form. In passing, the measure average household size is shown to reflect mainly fertility, and to have little to do with extended family structure. Other parts of Levy's argument are seen to require further specification before they can be proven or disproven. Data are presented which suggest the need for qualifications in the common view linking urbanization with the "breakdown" of the extended residential family. Limitations of the data used underline the tentative character of the findings and point the way to needed future research.

Burch, T. 1970. “Some Demographic Determinants of Average Household Size: An Analytic Approach.” *Demography* 7: 61-69. Descriptions of non-nuclear family systems in terms of rules of residence imply large and complex households, yet such households are not encountered as modal or average for large populations. Demographic factors, in particular high mortality, have been suggested as possible explanations for the apparent discrepancy. The purpose of this paper is to investigate the influence of demographic variables (viz., mortality, fertility, age at marriage) on average household size under different family systems--nuclear, extended and stem. The approach used has been applied by Coale to stationary populations. It has here been modified to apply to stable populations. The results indicate that under all family systems, average household size is positively correlated with fertility, life expectancy, and average age at marriage. Households under nuclear and stem family systems never exceed 10 persons on average. By contrast, under extended family systems, when mortality is low and fertility is high, average household size reaches levels seldom if ever observed in reality, e.g., 25 persons per household. Large households under the extended family system also tend to be fairly complex, often containing 5 or more adults. A number of modifications in the model would make for greater fit between model and real family systems. A more fruitful approach would involve the simulation of household formation and development.

Burch, T. 1995. “Theories of Household Formation: Progress and Challenges.” pp. 85-108 in E. van Imhoff, A.C. Kuijsten, P. Hooimeijer, and L.J.C. van Wissen (eds.) *Household Demography and Household Modeling*. New York: Plenum Press.

Burch, T., S. Halli, A. Madan, K. Thomas, and L. Wai. 1987. “Measures of household composition and headship based on aggregate routine census data.” pp. 19-33 in J. Bongaarts, T. Burch and K. Wachter (eds.) Family Demography: Methods and Their Application. Oxford: Clarendon Press. The authors evaluate the use of routine aggregate census data in the study of changes and differences in household composition, with a focus on data concerning number of households and population by age, sex, and marital status. Several measures based on these data are examined, including two standardized headship indexes and a number of simple ratios such as average household size, adults per household, and married persons per household. "These ratios are examined in the context of a decomposition of the 'crude headship rate'. The use of these simple ratios as proxies for more complex measures...is explored. Finally, decomposition is used to study the sources of variation in average household size."

Coale, A. 1965. “Estimates of Average Size of Household.” pp. 64-69 in A. Coale, L. Fallers, M. Levy, D. Schneider, and S. Tompkins (Eds.) *Aspects of the Analysis of Family Structure*. Princeton, NJ: Princeton University Press. The author identifies and discusses the demographic determinants of household size.

Clark, W. A. V., and F. Dieleman. 1996. *Households and Housing: Choices and Outcomes in the Housing Market.* New Brunswick, NJ: Center for Urban Policy Research. The authors note that neighborhoods have become the focus of questions about how they affect the families that live within them and that a working assumption of some federal policies is that, with help, households can escape poverty neighborhoods and change their spatial context. How true is this, especially for low-income households, and does changing neighborhoods have measurable benefits? The study uses data from the Moving to Opportunity program, initiated by the Department of Housing and Urban Development, to test whether policy interventions by means of housing vouchers have aided moves away from low-poverty areas and into integrated residential settings. By examining the neighborhood demography of the initial and subsequent locations of the samples, it is possible to assess the success of the objectives of decreasing poverty and increasing integration. Although the program has shown some success in assisting households to live in lower-poverty neighborhoods, the findings here emphasize just how difficult it is to intervene in dynamic processes such as housing choice and mobility to create policy outcomes.

Courgeau, D, V. Nedellec, and P. Empereur-Bissonnet. 2000. “Duration of Residence in the Same Dwelling. A Test of Measurement Using Electricity Utility Company Records.” *Population* 12: 335-342. The authors use the records of the national French utility provider, EDF, to track how long people remain in the same housing units. They find that the average duration of a domestic account was 13. years during the period 1977-1999.

Ellen, I. and B. O’Flaherty. 2007. “Social Programs and Household Size: Evidence from New York City.” *Population Research and Policy Review* 26 (4): 387-409. What determines how many adults live in a house? How do people divide themselves up among households? Average household sizes vary substantially, both over time and in the cross-section. In this paper, the authors describe how a variety of government policies affect living arrangements, intentionally or not. Using data from a survey of households in New York City, they find that these incentives appear to have an impact. Specifically, households receiving these housing and income subsidies are smaller on average (measured by number of adults). The impacts appear to be considerably larger than those that would occur if the programs were lump-sum transfers. Small average household size can be extremely expensive in terms of physical and environmental resources, higher rents, and possibly homelessness. They encourage policymakers to pay greater heed to the provisions built into various social policies that favor smaller households.

Findley, S. and H. Reinhardt. 1980. "Nonlinear Estimation of Household Size: The Minnesota Housing Unit Method." Paper presented at the annual meeting of the

Population Association of America, Denver, CO. Data from the 1977 Minnesota household survey are used to determine housing categories and size of place distinctions to estimate household size (PPH) and vacancy parameters. Using evidence from California and Washington, the authors determine that PPH change during the 1970s had been non-linear.

Gibson, C. 1986. "Post-censal estimates of households by size and type for states and of total households for counties." Paper presented at the Population Association of America Annual Meeting, San Francisco. This paper examines a demographic approach to sub-national household estimates which assumes that national trends in the average adult population per household and in the distribution of households by type and by size occur at the sub-national level. The assumption is tested for the 1970-1980 decade using US census data. Results show that, for states, the mean absolute error is lowest for the 20 years and up population (.69%), followed by the 18 years and up population (.76%), the 15 years and up population (1.05%), and all ages (1.61%). Counties show a lower mean error for the 18 and up group (1.47%), followed by the 20 and up (1.59%), the 15 and up (1.7%), and all ages (2.95%).

Hobbs, F., and N. Stoops. 2002. "Demographic Trends in the 20th Century." *U.S. Census Bureau, Census 2000 Special Reports, Series CENSR-4.* Washington, DC: Government Printing Office. This publication highlights the broad trends that have served to redraw America's demographic portrait continually over the course of the 20th century. It shows changes during the 20^{th} century in housing units and households by size for the U.S., its regions, and its states.

Hogan, T. D. and D. N. Steinnes. 1996. "Arizona Sunbirds and Minnesota Sunbirds: Two Species of the Elderly Seasonal Migrant Genus." *Journal of Economic and Social Measurement* 22: 129-139. This paper provides a view on the duration of seasonal residence using survey data collected in the state of Minnesota.

Hogan, T. D. and D. N. Steinnes. 1994. "Towards an Understanding of Elderly Seasonal Migration Using Origin-Based Household Data." *Research on Aging* 16: 463-475. This paper provides a view on the duration of seasonal residence using survey data collected in the area of Phoenix, Arizona. Analyses of data from a statewide household survey of older Minnesotans provide estimates of seasonal migration activity that are substantially higher than previous census-based estimates. More important, the data allow for comparisons between the migrant and non-migrant populations of the same origin state (Minnesota). The findings show distinct differences between the two populations for many socioeconomic household characteristics.

Hogan, T.D. and D.N. Steinnes. 1993. "Elderly Migration to the Sunbelt: Seasonal Versus Permanent." *Journal of Applied Gerontology* 12: 246-260. A substantial body of research has focused on the interstate migration of elderly households to the Sunbelt. Most of this research has concentrated on permanent moves, but seasonal migration of elderly households to Sunbelt locations has become an increasingly important social phenomenon. Although some have suggested that such temporary migration serves as a

precursor of permanent locations, recent analyses have found that such seasonal migration constitutes an alternative elderly life-style. Using 1980 census data, this study empirically examines the similarities and differences in these two types of elderly migration flows to a Sunbelt state such as Arizona. The statistical results indicate that seasonal and permanent migration are correlated in different ways to the variables usually found to be determinants of elderly migration flows and suggest the two types of elderly migration are related but separate phenomena.

Keilman, N. 1988. "Dynamic Household Models." pp. 123-138 in: N. Keilman, A. Kuijsten, and A. Vossen. (eds.) *Modelling Household Formation and Dissolution* Oxford, England: Clarendon Press. This chapter reviews dynamic houshold models that have been constructed recently. The analysis is limited to purely demographic macromodels. It discusses: (1) the Canadian model developed by Akkerman; (2) Moller's ISP model for the Federal Republic of Germany; (3) the Swedish model constructed by Harsman, Snickars, Holmberg, and colleagues; (4) the model which resulted from research undertaken by Heida and Gordijn in the Netherlands; and (5) the model which is currently being constructed at the Netherlands Interuniversity Demographic Institute (NIDI). The comparison shows that it seems to be good strategy to model the behavior of individuals rather than that of households. The chapter also finds that the multidimensional approach holds considerable promise for modelling household behavior. Difficulties common to most models are: (1) the availability of data necessary to run the models; (2) inconsistencies that arise when individuals of different sex are modeled separately; and (3) the exponential increase in the number of states when a detailed household breakdown is considered.

Kimpel, T., and T. Lowe. 2007. “Estimating Household Size for use in Population Estimates.” *Population Estimates and Projections Research Brief No. 47*. Olympia, WA: Washington State Office of Financial Management. This Brief revisits a topic developed in Research Brief No. 10 where a regression procedure was developed using administrative data to update household size - a key variable used in local population estimates based on the Housing Unit Method. The authors find that generally the most accurate population estimates come from using several procedures and understanding the biases in each.

Korbin, F. 1976. “The Fall in Household Size and the Rise of the Primary Individual in the United States.” *Demography* 13 (1); 127-138. The long-term fall in household size in the United States is discussed within the framework of the aging of the population, continuing as the effects of fertility and mortality decline accumulate. Using distributions of households by size from U.S. census data 1790-1970 and a components of change analysis on primary individuals for 1950-1974, household changes are related to demographic change for the periods 1790-1900, 1900-1950, and 1950-1974. Fertility and mortality declines have unambiguous impact on household size until the increases in primary individuals begin. But these, too, have a theoretically interesting, if indirect relationship to population structure.

Lowe, T., et al. 2000. “Developing Trends in Household Size for use in Population Estimates.” *Population Estimates and Projections Research Brief No. 10*. Olympia, WA: Washington State Office of Financial Management. This Brief illustrates a regression procedure using administrative data that can be used to update household size in the Housing Unit Method.

Lowe, T., L. Weisser, and B. Myers. 1984. “A Special Consideration in Improving Housing Unit Estimates: The Interaction Effect. Paper presented at the annual meeting of the Population Association of America, Minneapolis, MN. The authors observe that of some of the terms in the equation for the Housing Unit Method are themselves correlated and discuss how this can be dealt with in using the Housing Unit method to estimate population. They specifically point to the interaction of average household size and structure type.

Mason, A. and R. Racelis. 1992. “A Comparison of Four Methods for Projecting Households.” *International Journal of Forecasting* 8 (3): 509-527. The purpose of this paper is to extend the headship rate method for projecting households to encompass both sexes. Four models are considered that explicitly incorporate the impact of changes in the number of men and women on the number and joint age distribution of husband-wife households. The models are applied to the Philippines using data from the 1988 National Demographic Survey to project households to 2010. The models are also evaluated by ‘backcasting’ and comparing the results with special tabulations from the 1970 and 1980 censuses and the 1975 National Demographic Survey.

Myers, D., and A. Doyle. 1990. “Age-specific Population-per-Household Ratios: Linking Population Age Structure with Housing Characteristics.” pp. 109-130 in D. Myers (ed.) *Housing Demography: Linking Demographic Structure and Housing Markets*. Madison, WI: University of Wisconsin Press. The authors develop a demographic framework for determining Persons per household and propose models for estimating these values with the age structure of the household population.

Roe, L., J. Carlson and D. Swanson. 1992. “A Variation of the Housing Unit Method for Estimating the Population of Small, Rural Areas: A Case Study of the Local Expert Procedure.” *Survey Methodology* 18 (1): 155-163. The authors report on a random sampling study based on an adaptation of the Housing Unit Method and the local expert method to determine the socioeconomic features of three unincorporated rural communities near Yucca Mountain, Nevada. A sample of the study area was selected from a carefully screened fame comprised of electrical utility data. Meter readers from the local utility companies were the local experts and two of them worked together to authenticate the accuracy of recorded data which included number of person in the household as of July 15, 1990 and age and gender of each member. Data accuracy was tested and it was found that the 1990 US Census counts were within the relatively narrow 95% confidence intervals. The mean width was 7.2% of the estimated population, thus the authors conclude that the estimates were meaningful.

Shapiro, G. and D. Kostanich. 1998. "High Response Error and Poor Coverage are Severely Hurting the Value of Household Survey Data." pp. 443-448 in 1998 *Proceedings of the Survey Research Methods Section, American Statistical Association*, Alexandria, VA. The authors discuss two major sources of non-sampling error: undercoverage of the population and response error. They find that both of these can be quite substantial and result in major biases for some survey estimates. They describe work that should be done on these errors and summarize how damaging these errors and that greater efforts are needed to reduce them.

Smith, S. K., J. Nogle, and S. Cody. 2002. "A Regression Approach to Estimating the Average Number of Persons per Household." *Demography* 39(4): 697-712. The authors develop several regression models in which PPH estimates were based on symptomatic indicators of PPH change. They tested these estimates using county level data in four states and found them to be more precise and less biased than estimates based on more commonly used methods.

Swanson, D., and G. Hough. 2007. An Evaluation of Persons Per Household (PPH) Data Generated By the American Community Survey: A Demographic Perspective." Paper presented at the Annual Meeting of the Southern Demographic Association. Birmingham, AL. This paper explores the usability of ACS data by examining "Persons Per Household (PPH), a variable of high interest to demographers and others preparing regular post-censal population estimates. The data used in this exploration are taken from 18 of the counties that formed the set of 1999 ACS test sites. The examination proceeds by comparing ACS PPH values to PPH values generated using a geometric model based on PPH change between the 1990 and 2000 census counts. The ACS PPH values represent what could be called the "statistical perspective" because variations in the values of specific variables over time and space are viewed largely by statisticians with an eye toward sample (and non-sample) error. The model-based PPH values represent a "demographic perspective" because PPH values are largely viewed by demographers as varying systematically, an orientation stemming from theory. The results suggest that the ACS PPH values lack sufficient temporal consistency to be used by demographers.

Swanson, D., and T. Lowe. 1979. "Comparative Analysis of Change in Average Household Size with Reference to IRS Data on Average Exemptions Per Return: Census Results from Selected Municipalities in Washington, 1970, 1977, and 1979." *Special Report No. 22*. Olympia, WA: Washington State Office of Financial Management. Using the results of the 1970 census and special census counts done in Washington State in 1977 and 1979, the authors evaluate the suitability of IRS data on the average number of exemptions of return as a basis for estimating Persons per household. They find deficiencies in the IRS data for this purpose, but suggest possible adjustments that could be tested in the future.

Swanson, D., and T. Lowe. 1979. "Recent Trends in Household Size for Rural, Predominantly White, Non-Hispanic Communities in Washington: Special Census Results from Three Towns in Washington." *Special Report No. 21*. Olympia, WA:

Washington State Office of Financial Management. The authors show the decline in average household size documented by special census counts during the 1970s.

Sweet, J. 1984. "Components of Change in the Number of Households: 1970-1980." *Demography* 21 (2): 129-140. There was an increase from 62.8 to 79.1 million households in the United States during the 1970s. The number of households increased much more rapidly than the population. This paper decomposes this growth in the number of households into components associated with changing age and marital status composition and changing age by marital status-specific propensities to form households. About one-third of the increase in the number of households was due to increased age by marital status propensity to form households, and two-thirds was due to shifts in the age by marital status distribution and population growth. The increased propensity to form households had its major impact at ages under 35, and primarily among never-married persons. The composition component had its primary impact at ages 25-44 as a result of the baby boom, and also because of the increased fractions never married and separated and divorced.

Van Imhoff, E., A.C. Kuijsten, P. Hooimeijer, and L.J.C. van Wissen (eds.). 1995. *Household Demography and Household Modeling*. New York: Plenum Press. In addition to the Introduction and Epilogue, this book consists of 13 chapters organized into three sections: (1) Trends and Theories, with three papers, by Well, Kuusten, and Burch, respectively; (2) Data and Analysis, with four papers, by Keilman, De Jong Gierveld, Ott, Courgeau, and Galler, respectively; and (3) Models, which has six papers by, respectively, Prinz et al., De Beer; Van Imhoff; Hooimeijer and Heida, and Nelissen. The papers by De Beer, Van Imhoff, Hooimeijer and Heida deal with projections of households, but the techniques they discuss can also be used for estimation.

Voss, P., and H. Krebs. 1979. "The use of Federal Revenue Data for Improving Current Estimates of Average Household Size for Minor Civil Divisions: An Evaluation." *Technical Series 70-5.* Applied Population Laboratory. Madison, WI: University of Wisconsin. The authors evaluate the suitability of IRS data on the average number of exemptions of return as a basis for estimating Persons per household for Minor Civil Divisions (townships) in Wisconsin. They find deficiencies in the IRS data for this purpose, but suggest possible adjustments that could be tested in the future.

Weidman, L., R. Creecy, D. Malec, and J. Tsay. 2008. "Exploration of the Use of Empirical Bayes Procedures for Estimating Changes in Occupancy Rate and Persons Per Household." *Research Report Series, Statistics #2008-07.* Statistical Research Division. Washington, DC: U.S. Bureau of the Census. The authors explore the use of an empirical Bayes approach as a means of updating PPH and VR values from the most recent Decennial Census in conjunction with ACS values. They examine national and state models using 1990 and 2000 long form data (to represent ACS values), while 1990 and 2000 short form data are used as the independent variables. The question they are aiming at ultimately answering is: Should statistically significant PPH and VR values from the ACS be used to estimate change over time in PPH and VR? Their work suggests several Research projects that need to be conducted before their question can be answered

Zeng, Y. K. Land, Z. Wang, and D. Gu. 2006. "U.S. family household momentum and dynamics: an extension and application of the ProFamy Method." *Population Research and Policy Review* 25: 1–41. The authors argue that the classic headship-rate method for demographic projections of households is not linked to demographic rates, projects a few household types without size, and does not deal with household members other than heads. By comparison, the ProFamy method uses demographic rates as input and projects more detailed household types, sizes, and living arrangements for all members of the population. Tests of projections from 1990 to 2000 using ProFamy and based on observed U.S. demographic rates before 1991 show that discrepancies between our projections and census observations in 2000 are reasonably small, validating the new method.

GLOSSARY

ADJUSTED RATE. (See STANDARDIZATION).

ADMINISTRATIVE RECORDS. Data collected by governmental (and sometimes private) organizations for taxation, registration, fee collection, and other administrative purposes that indirectly provide demographic information. These data are used by demographers for analyses, estimates, projections, and the evaluation of data specifically collected for demographic purposes (See also ADMINISTRATIVE RECORDS METHOD).

ADMINISTRATIVE RECORDS METHOD. In the United States, a member of the family of component methods for estimating population that relies on a past census, vital statistics data, and migration data derived from tax returns (See also POPULATION ESTIMATE).

AGGREGATION. The process of assembling individual elements into summary form for purposes of presentation or analysis. For example, to assemble census records for individuals in a given area into a summary for the area as a whole.

AGGREGATION BIAS. A type of distortion that can result by attributing relationships found among summaries to the individual elements from which the summaries were obtained.

ALLOCATION. The assignment of values to cases for which "item non-response" is found in a sample survey or census. Many allocation methods are available, including automated algorithms (See also IMPUTATION and SUBSTITUTION).

AMERICAN COMMUNITY SURVEY (ACS). In the United States, an on-going household survey conducted by the Census Bureau on a "rolling" geographic basis that is designed to provide demographic characteristics for counties, places, and other small areas. It may replace the long-form in the 2010 census (See also AMERICAN COMMUNITY SURVEY-DEFINED RESIDENT and CENSUS).

AMERICAN COMMUNITY SURVEY-DEFINED RESIDENT. The American Community Survey residency rule is more consistent with the Defacto approach to residency than with the De Jure approach used by the Decennial Census. It uses a two month residency rule, whereas the Census Residency Rule is based on self-reporting of one's usual place of residence (See also CENSUS DEFINED RESIDENT and RESIDENCY RULES)

ANNEXATION. In the United States, the legal act of adding territory to a governmental unit, usually an incorporated place, through the passage of an ordinance, court order, or other legal action.

AREA ANALYSIS. Measurements collected on a number of variables for each of many administrative/ statistical areas that are usually analyzed using multivariate techniques.

AT-RISK POPULATION. The persons to whom an event can potentially occur. In the form of the population at the middle of a given period, such as a year, it is used as an approximation of "Person-years lived"(See also EXPOSURE, PERSON-YEARS LIVED, and PROBABILITY).

AVERAGE FAMILY SIZE. The mean number of living children of an individual or couple.

BALANCING EQUATION. A term attributed to A. Jaffe that describes the basic population relation: Pt = P0 + I – O, where Pt equals a given population at time = 0 + t, P0= the given population at time =0, I = the number of persons entering the population through birth and immigration between time=0 and time = 0+t, and O = the number of persons exiting the population through death and emigration between time=0 and time=0+t (See also COHORT-COMPONENT METHOD, COMPONENT METHOD, ERROR OF CLOSURE, and RESIDUAL METHOD).

BASE PERIOD. In a population projection, this is the period between the initial year for which data are used to generate the projection and the last year, which is known as the launch year. (See also LAUNCH YEAR, PROJECTION HORIZON, and TARGET YEAR; and POPULATION PROJECTION).

BIAS. The deviation of an estimate or set of estimates from the correct value(s) in one direction (i.e., above or below the correct value(s)).

BLOCK. In the United States, the lowest level of geography for which census data are compiled. It is a typically a city block, but specifically is a small area bounded on all sides by identifiable features (e.g., roads, rivers, and city limits) that does not cross the boundaries of a given census tract. Each block is numbered uniquely within census tracts (See also BLOCK GROUP, BLOCK NUMBERING AREA, CENSUS TRACT, and CENSUS GEOGRAPHY).

BLOCK GROUP. In the United States, a cluster of blocks within a census tract that have the same first digit in their identifying numbers (See also BLOCK, BLOCK NUMBERING AREA, CENSUS TRACT and CENSUS GEOGRAPHY).

BLOCK NUMBERING AREA. In the United States, these were used in the1990 census as the framework for grouping and numbering blocks in counties that did not have census tracts and provided coverage only for the block-numbered portion of a county. Starting with the 2000 Decennial census all U.S. counties have census tracts. (See also BLOCK, BLOCK GROUP, CENSUS TRACT, and CENSUS GEOGRAPHY).

CAPTURE-RECAPTURE METHOD. A technique initially developed to estimate the size of a given wildlife population in which two samples are taken. In the first sample, individuals are uniquely marked, counted, and released. After a period of time to allow the marked individuals to redistribute themselves among the unmarked population, a second sample is taken. From the distribution of marked and unmarked individuals in the second sample, an estimate of population size can be obtained using certain assumptions (i.e., the population is closed and the two samples are independent) as follows: $N = (n_1 * n_2)/n_{12}$, where N is the number estimated for the total population, n_1 = the number marked in the first sample, n_2 is the number marked in the second sample and n_{12} is the number marked in both samples. This formula is an algebraic re-arrangement of the formula used to determine expected cell size in a 2x2 table for purposes of the Chi-squared test. In a human setting, two sets of administrative records usually take the place of the two samples (See also DUAL- SYSTEMS ESTIMATION and MATCHING).

CENSAL-RATIO METHOD. A set of population estimation techniques found within the "Change in Stock Method" family that uses crude rates (e.g., birth and death) as measured at the most recent census date(s) and post-censal administrative records. For example, a population estimate for 2002 can be obtained by dividing reported deaths for 2002 by the crude death rate measured in 2000 or by a crude death rate projected from 2000 to 2002. Often a series of Censal-Ratio estimates are averaged together. D. Swanson and R. Prevost showed in 1985 that the Ratio-Correlation Method is algebraically equivalent to a weighted average of censal-ratio estimates in which regression slope coefficients serve as weights (See also CHANGE IN STOCK METHOD, POPULATION ESTIMATE, RATIO-CORRELATION METHOD, and WEIGHTED AVERAGE).

CENSORED. A condition affecting time-ordered data because the time frame for which data are collected does not cover the entire time span over which an event of interest may occur (e.g., a pregnancy at future point beyond the time frame in which data were collected). "Left-Censored" is used to describe the period preceding the data collection time frame and "right-censored," the subsequent period.

CENSUS. The count of a given population (or other phenomena of interest) and record its characteristics, done at a specific point in time and usually at regular intervals by a governmental entity for the geographic area or subareas under its domain (See also AMERICAN COMMUNITY SURVEY, CENSUS COVERAGE, CENSUS DEFINED RESIDENT, POPULATION, POPULATION ESTIMATE, and SAMPLE).

CENSUS COVERAGE. An estimate of how complete a census was of a given population (See also COVERAGE ERROR, NET CENSUS UNDERCOUNT ERROR and TRUE POPULATION).

CENSUS COVERAGE ERROR. (See COVERAGE ERROR).

CENSUS COUNTY DIVISION. In the United States, a statistical subdivision of counties in states established cooperatively by the Census Bureau and local groups in which minor civil divisions (e.g., townships) are not suitable for presenting census data (See also CENSUS GEOGRAPHY).

CENSUS DEFINED RESIDENT. The concept of defining persons counted in a census in order to count each and every person once and only once. One of two counting bases is used: (1) De Jure, which attempts to locate persons at their usual residence; and (2) De facto, which counts people where they are found. The U.S. Decennial Census is based on the De Jure method. There is a difference between the definition of residency used in the Decennial Census and its related products and the definition of residency used in the American Community Survey. (See also AMERICAN COMMUNITY SURVEY-DEFINED RESIDENT, CENSUS, DE FACTO POPULATION, DE JURE POPULATION, DOMICLE, RESIDENCE, RESIDENCY RULES, and USUAL RESIDENCE).

CENSUS DESIGNATED PLACE (CDP). In the United States, a concentration of population enumerated during the decennial census in an area lacking legal boundaries, but recognized by the residents (and others) as a distinctive area with a name. A CDP is defined cooperatively by local officials and the Census Bureau. CDPs have been used since the 1980 census; from 1940 to 1970, they were called Unincorporated Places. (See also CENSUS GEOGRAPHY).

CENSUS ERROR (See COVERAGE ERROR).

CENSUS GEOGRAPHY. In the United States, this refers to the hierarchical system of geographic areas that is used in conjunction with each decennial census. It consists of two major components: (1) areas defined by political or administrative boundaries (e.g., states, counties, townships, and cities.); and (2) areas defined by "statistical" boundaries (e.g., block, census designated place, census tract). The areas so defined are used for analytical, political, and administrative purposes. Any country conducting a census uses some type of census geography. (See also BLOCK, CENSUS COUNTY DIVISION, CENSUS DESIGNATED PLACE, CENSUS TRACT, CITY, COUNTY, METROPOLITAN AREA).

CENSUS TRACT. In the United States, this is the lowest level of "statistical geography" found in the decennial census designed to be homogenous with respect to population and economic characteristics (note that blocks and block groups, while at a lower level, are not designed with respect to population or economic homogeneity). Once established it is designed to be consistent in its boundaries for a long period of time. Starting with the 2000 census, all areas in the United States are tracted. (See also BLOCK, BLOCK GROUP, BLOCK NUMBERING AREA; and CENSUS GEOGRAPHY).

CENTRAL CITY. Within the U.S. Census Bureau's geography system, the core area in a metropolitan area. However, in other contexts, it is usually viewed as the concentrated inner area of a city consisting of business districts and urban housing.

CENTRAL RATE. An event-exposure ratio where the numerator is the number of events in a given period to the estimated number of people exposed to the event as of the mid-point of the period.

CHANGE IN STOCK METHOD. A family of techniques for estimating population that is based on the measuring the total change in population since the last census rather than the components of change. Examples include the censal-ratio method, housing unit method, and the ratio-correlation method (See also COMPONENT METHOD, CENSAL-RATIO METHOD, HOUSING UNIT METHOD, and POPULATION ESTIMATE).

CHILD-WOMAN RATIO. A measure formed by dividing the number of children (aged 0-4) by the number of women of child-bearing age (aged either 15-49 or 15-44). (See also EFFECTIVE FERTILITY).

CITY. In the United States, a type of incorporated place (See also CENSUS GEOGRAPHY).

CIVILIAN NON-INSTITUTIONAL POPULATION. In the United States, persons 16 years and over who are not inmates of institutions and who are not on active duty in the armed forces (See also CURRENT POPULATION SURVEY).

CIVILIAN POPULATION. Persons who are not members of the armed forces. This may, however, include persons who are dependents of members if the armed forces.

CLOSED POPULATION. A population for which in and out migration is minimal, if at all. For example, the population of the world as a whole is "closed," whereas the population of New York City is not.

COHORT. A group of people who experience the same demographic event during a particular period of time such as their year of marriage, birth, or death. Cohorts are typically defined on the basis of a initiating signal event (e.g., birth), but they also can be defined on the basis of a terminating signal event (e.g., death). (See also COHORT ANALYSIS, COHORT EFFECT, COHORT MEASURE, and PERIOD).

COHORT ANALYSIS. An analysis that traces the demographic history of a cohort as it progresses through time (See also AGE-PERIOD-COHORT EFFECT, COHORT, LEXIS DIAGRAM, and PERIOD ANALYSIS).

COHORT CHANGE RATIO (see HAMILTON-PERRY METHOD).

COHORT-COMPONENT METHOD. A projection technique that takes into account the components of population change, births, deaths, and migration, and a population's age and sex composition, (See also BALANCING EQUATION, COMPONENT METHOD, and POPULATION PROJECTION).

COMPONENT METHOD. In general, this refers to any technique for estimating population that incorporates births, deaths, and migration. Also known as a "Flow Method" (See also BALANCING EQUATION, CHANGE IN STOCK METHOD, COMPONENT METHOD I, COMPONENT METHOD II, and POPULATION ESTIMATE).

COMPONENT METHOD I. A component method of estimating population that uses the relationship between local and national school enrollment data to estimate the net migration component. (See also COMPONENT METHOD, COMPONENT METHOD II, and POPULATION ESTIMATE).

COMPONENT METHOD II. A component method of estimating population that uses the relationship between expected (survived) and actual local school enrollment data to estimate the net migration component. (See also COMPONENT METHOD, COMPONENT METHOD I, and POPULATION ESTIMATE).

COMPONENTS OF CHANGE. There are four basic components of population change: births, deaths, in-migration, and out-migration. The excess of births over deaths results in natural increase, while the excess of deaths over births results in natural decrease. The difference between in- and out-migration is net migration. In an analysis of special characteristics or groups, the number of components is broadened to include relevant additional factors (e.g., aging, marriages, divorces, annexations, and retirements), depending on the group (See also BALANCING EQUATION).

COMPOSITE METHOD. A technique for estimating total population that is based upon independent estimates of age or age-sex groups that are summed to obtain the total population (See also POPULATION ESTIMATE).

CONSOLIDATED METROPOLITAN STATISTICAL AREA. (See METROPOLITAN AREA).

CONTROLLING. The act of adjusting a distribution to an independently derived total value (See also CONTROLS).

CONTROLS. Independently derived estimates of a "total value" to which distributions are adjusted for purposes of improving accuracy, reducing variance and bias, or maintaining consistency. Controls can be univariate (one-dimensional) or multivariate (n-dimensional). Many methods may be used, including those that take account of whether the distributions have only positive values or both positive and negative values. (See also CONTROLLING, ITERATIVE PROPORTIONAL FITTING and PLUS-MINUS METHOD).

COVERAGE ERROR. In principle, this refers to the difference between the "true population" and the number reported in a set of data such as a census, survey, or set of administrative records. In practice, it is the difference between an estimate of the true number and the number reported in a set of data such as a census, survey, or set of administrative records (See also CENSUS, NET CENSUS UNDERCOUNT ERROR, TOTAL ERROR, and TRUE POPULATION).

COUNTY. In the United States, a type of governmental unit that is the primary administrative subdivision of every state except Alaska and Louisiana (See also CENSUS GEOGRAPHY).

COUNTY EQUIVALENT. In the United States, a geographic entity that is not legally recognized as a county but referred to by the Census Bureau as the equivalent of a county for purposes of data presentation. Boroughs and certain statistically defined areas are county equivalents in Alaska and parishes are county equivalents in Louisiana (See also COUNTY and CENSUS GEOGRAPHY).

CROSS SECTIONAL ANALYSIS. Studies that focus on phenomena that occur during a precise time interval (such as a calendar year) among several cohorts (See also COHORT ANALYSIS and PERIOD ANALYSIS).

CURRENT POPULATION SURVEY(CPS). In the United States, a sample survey conducted monthly by the Census Bureau designed to represent the civilian non-institutional population that obtains a wide range of socio-economic-demographic data (See also CIVILIAN NON-INSTITUTIONAL POPULLATION).

CURVE. A mathematical function, usually continuous and otherwise "well-behaved" that can be used as a model for a demographic process such as the change in the size of a population over time. Examples include the Exponential, Geometric, Gompertz, Linear, Logistic, and Polynomial.

CURVE-FITTING. The process of finding a mathematical function that serves as a model for a given demographic process.

DATA LINKAGE. (see MATCHING).

DECREMENT. The exit of an individual or set of individuals from a "population" of interest, where the population is often defined by a model. In the case of a model such as the standard life table, such an exit would be due to death (See also INCREMENT and INCREMENT-DECREMENT LIFE TABLE).

DE FACTO POPULATION. A census concept that defines an enumerated person on the basis of his or her actual location at the time of the census (See also AMERICAN COMMUNITY SURVEY- DEFINED RESIDENT, CENSUS DEFINED RESIDENT, DE JURE POPULATION, and RESIDENCY RULE)

DE JURE POPULATION. A census concept that defines an enumerated person on a basis other than his or her actual location at the time of the census. The most common basis is the person's usual place of residence at the time of a census. (See also AMERICAN COMMUNITY SURVEY-DEFINED RESIDENT, CENSUS DEFINED RESIDENT, DE FACTO POPULATION, and RESIDENCY RULE).

DEMOGRAPHIC ACCOUNTING. The process of analyzing the change in a population using "stocks" (e.g., conditions such as the number of people in a given age-sex group) and "flows" (e.g., events such as births and deaths by age and sex) to show how the flows affect stocks over time. Ideally the stocks and flows should be measured without error and form mutually exclusive and exhaustive categories.

DEMOGRAPHIC ANALYSIS. Generally, this refers to the methods of examination, assessment, and interpretation of the components and processes of population change, especially births, deaths, and migration. In the United States, it also refers to a specific method of estimating net census undercount using the components and process of population change.

DEMOGRAPHICS. A popular term for demography also used to represent demographic data and the application of demographic data, methods, and perspectives to activities undertaken by non-profit organizations, businesses, and governments (See also DEMOGRAPHY).

DEMOGRAPHY. The study of population, typically focused on five aspects: (1) size; (2) geographic distribution; (3) composition; (4) the components of change (births, deaths, migration); and (5) the determinants and consequences of population change. This term is usually used to refer to human populations, but it also is used to refer to non-human, particularly wildlife, populations. (See also DEMOGRAPHICS, FAMILY DEMOGRAPHY, HOUSEHOLD DEMOGRAPHY, ORGANIZATIONAL DEMOGRAPHY, and POPULATION).

DENSITY. The number of people per unit area (e.g., persons per square kilometer).

DIFFERENCE-CORRELATION METHOD. (See RATIO-CORRELATION METHOD).

DIRECT ESTIMATION. The measurement of demographic phenomena using data that directly represent the phenomena of interest.(See also INDIRECT ESTIMATION).

DIRECT STANDARDIZATION. The adjustment of a summary rate (e.g., the crude death rate) for a population in question found by computing a weighted average of group-specific rates (e.g., age specific death rates) for the population in question, where the weights consist of specific groups (e.g., the proportion in each age group) found in a "standard" population. This procedure is designed to produce a summary rate that controls for the effects of population composition (e.g., age) and is usually used for purposes of comparison with directly standardized rates for other populations computed using the same standard population. To standardize a crude death rate by the direct method, multiply the age-specific death rates for the population in question by the age-specific proportions in a standard population and sum the products. (See also INDIRECT STANDARDIZATION, STANDARD POPULATION, and STANDARDIZATION).

DIURNAL FLUCTUATION. For a given area, the change in its De Facto population over the course of a day (i.e., a 24 hour period) (See also DE FACTO POPULATION).

DOMESTIC MIGRATION. The movement of people within a given country across political or administrative boundaries. People leaving an area are out-migrants and those entering an area are in-migrants. It is a synonym for internal migration. (See also FOREIGN MIGRATION and MIGRATION)

DOMICILE. A person's fixed, permanent, and principal home for legal purposes (See also HOUSEHOLD, HOUSING UNIT, RESIDENCE and USUAL RESIDENCE).

DUAL RESIDENCE. The state of having two usual places of residence over a given period of time, which must be resolved in a De Jure census through the use of a set of procedures designed to count persons once and only once.

DUAL-SYSTEMS ESTIMATION. Estimation of the true number of events or persons by matching the individual records in two data collections systems (See also CAPTURE-RECAPTURE METHOD and MATCHING).

EMIGRANT. A resident of a given country who departs to take up residence in another country (See also DOMESTIC MIGRATION, FOREIGN MIGRATION, and MIGRATION).

EMIGRATION. (see FOREIGN MIGRATION).

EMIGRATION RATE. An out-migration rate for a country as a whole (See also FOREIGN MIGRATION and OUT-MIGRATION RATE).

ENUMERATION. The act of counting the members of a population in a census.

ENUMERATION DISTRICT. The area assigned to an enumerator during a census or survey of a given area.

ERGODICITY. A process whereby a closed population subject to fixed or nearly fixed fertility and mortality schedules eventually acquires a constant or nearly constant age composition that is independent of its starting age composition (See also STABLE POPULATION).

ERROR OF CLOSURE. The difference between the change in population implied by census counts at two different dates and the change implied by an estimate not dependent on both census counts. This also can refer to a term added to the demographic balancing equation to account for errors in the components of change that cause them not to exactly match the change in measured independently for the population to which they apply. (See also BALANCING EQUATION and RESIDUAL METHOD).

ESTIMATE. (See POPULATION ESTIMATE).

EXPOSURE. The condition of a population being at risk of having an event occur to it during a specified period (See also AT-RISK POPULATION).

EXTENDED FAMILY. A group within a specified degree of consanguinity or marriage who tend to collaborate in support activity for one another.

EXTINCT GENERATIONS. A technique introduced by P. Vincent in the early 1950s that is designed to estimate the number of extremely old persons in a population at a given date by cumulating deaths (to include, as needed, reported, estimated, and projected deaths) to given cohorts to the point where all members of the given cohorts have expired.

EXTRAPOLATION. The process of determining (estimating or projecting) values that go beyond the last known data point in a series (e.g., the most recent census or estimate). It is typically accomplished by using a mathematical formula, a graphic procedure, or a combination of the two. (See also INTERPOLATION).

FAMILY. Those members of a household who are related through blood, adoption, or marriage (See also HOUSEHOLD and NON-FAMILY HOUSEHOLD).

FAMILY DEMOGRAPHY. The study of the size, distribution, and composition of families, along with their components of change, and the determinants and consequences of family change (See also DEMOGRAPHY and HOUSEHOLD DEMOGRAPHY).

FAMILY LIFE CYCLE. An approach to the study of the family that considers its evolution through various stages (e.g., marriage, birth of first child, birth of last child, divorce, widowhood).

FAMILY SIZE . Number of living children of an individual or couple.

FERTILITY. The reproductive performance of a woman, man, couple, or group. Also a general term for the incidence of births in a population or group. One of the components of population change (See also COMPONENTS OF CHANGE and FECUNDITY).

FIPS CODE. In the United States, one of a series of codes issued by the National Institute of Standards and Technology for the identification of geographic entities. FIPS stands for "Federal Information Processing Standards"

FLOW METHOD. (See COMPONENT METHOD).

FORECAST. (See POPULATION FORECAST).

FOREIGN MIGRATION. The movement of residents across national boundaries. Residents of a given country leaving it to take up permanent residence in another country are called emigrants and residents of a given country entering another country in order to take up permanent residence are called immigrants (See also DOMESTIC MIGRATION and MIGRATION).

FORWARD SURVIVAL RATE. A type of rate that expresses survival of a population group from a younger age to an older age. Where a survival rate is not further labeled, forward survival is to be assumed (See also REVERSE SURVIVAL RATE and SURVIVAL).

FORWARD-REVERSE SURVIVAL METHOD. A technique used in both estimating intercensal populations and net migration between two censuses in which an "average" is taken between the results of using forward and reverse survival rates to age and "young" a given population, respectively, over the period between the two censuses (See also FORWARD SURVIVAL RATE, REVERSE SURVIVAL RATE, and SURVIVAL).

GENERAL RATE. A rate that relates a demographic event to a set of people in a given population generally thought to be exposed to the event of interest, but one for which no distinction is made regarding different exposure levels to the event. A GENERAL RATE is distinguished from a CRUDE RATE because of the former's attempt to limit the population at risk to those actually exposed to the event in question, typically on the basis of age. Examples include the General Activity Rate, General Divorce Rate, General Enrollment Rate, and the General Fertility Rate. (See also AGE-SPECIFIC RATE, CRUDE RATE and RATE)

GEOCODING. The assignment of geographic or spatial information to data, such as coordinates. It is the most fundamental operation in the development of a "GIS" - Geographic Information System (See also GEOGRAPHIC INFORMATION SYSTEM).

GEOGRAPHIC INFORMATION SYSTEM (GIS). A chain of operations involving the collection, storage, manipulation, and display of data referenced by geographic or spatial coordinates (e.g., coded by latitude and longitude).

GIS. (See GEOGRAPHIC INFORMATION SYSTEM).

GRADUATION. (See SMOOTHING).

GRAVITY MODEL. A model (borrowed from classical physics) based on the hypothesis that movement (migration, commuting, retail purchasing, etc.) between two areas is directly related to the population size of each area and inversely related to the distance between the two areas.

GROSS MIGRATION. The sum of in-migration and out-migration for a given area (See also MIGRATION, and NET MIGRATION).

GROUP QUARTERS. In the United States, a term used by the Census Bureau for places in which people reside that are not considered as "housing units." Such places include prisons, long-term care hospitals, military barracks, and school and college dormitories. (See also HOUSING UNIT and HOUSEHOLD POPULATION).

GROWTH RATE. Often used as a general expression to describe the rate of change in a given population, even one that is declining (See also RATE and RATE OF CHANGE).

HAMILTON-PERRY METHOD. A technique developed by H. Hamilton and J. Perry used in population projections that refers to a type of survival rate calculated for a cohort from two censuses. It includes not only the effects of mortality, but also the effects of net migration and relative census enumeration error (See also SURVIVAL RATE).

HAZARD FUNCTION. One of three algebraically related functions used in survival analysis, the other two being the "Death Density Function" and the "Survivorship Function." The hazard function is found by dividing the death density function by the survivorship function (See also DEATH DENSITY FUNCTION, HAZARD RATE, and SURVIVORSHIP FUNCTION).

HAZARD RATE. The probability that an event occurs within a given time interval, no matter how small the interval, given that the event has not occurred to the subject of interest prior to the start of the interval. Typically, the event of interest is a "decrement" such as death (See also DECREMENT, FORCE OF MORTALITY, and HAZARD FUNCTION).

HEAD OF HOUSEHOLD. A "marker" for a household, its type and structure. It is usually defined as the principal wage-earner or provider for a multi-person household, or, alternatively, is a person in whose name the housing unit is rented or owned. Persons living alone also are designated as heads of households. In principle, the number of households is equal to the number of household heads (See also HEADSHIP RATE and HOUSEHOLD).

HEADSHIP RATE. Usually defined as the proportion of the (household) population who are "heads" of households. (i.e., divide the number of households by the household population), often by age. It is often used in conjunction with population projections to obtain household projections (See also HEAD OF HOUSEHOLD, HOUSEHOLD and POPULATION PROJECTION).

HETEROGENEITY. The presence of variation among the members of a population with respect to a given characteristic of interest (See also HOMOGENEITY).

HOMELESS PERSON. Member of a population without a home or an official address usually found in shelters, on the streets, in vacant lots or vacant buildings not intended for residence (See also DOMICLE, GROUP QUARTERS, HOUSING UNIT, and HOUSEHOLD).

HOMOGENEITY. Lack of variation among the members of a population with respect to a given characteristic of interest (See also HETEROGENEITY).

HORIZON. (See PROJECTION HORIZON).

HOT DECK IMPUTATION. (See IMPUTATION).

HOUSEHOLD. Either a single person or a group of people making provision for food and other essentials of living, occupying the whole, part of, or more than one housing unit or other provision for shelter. (See also DOMICLE, FAMILY, GROUP QUARTERS, HEAD OF HOUSEHOLD, HOMELESS PERSON, HOUSEHOLD POPULATION, HOUSING UNIT, RESIDENCY RULE, and VACANT HOUSING UNIT).

HOUSEHOLD DEMOGRAPHY. The study of the size, distribution, and composition of households, along with their components of change and the determinants and consequences of household change. Sometimes called HOUSING DEMOGRAPHY (See also DEMOGRAPHY, FAMILY DEMOGRAPHY, and HOUSEHOLD).

HOUSEHOLD POPULATION. Members of a population living in households, (as opposed to those who are homeless or living in group quarters - e.g., prisons, long-term care hospitals, military barracks, and school and college dormitories) (See also GROUP QUARTERS, HOMELESS PERSONS, HOUSEHOLD, HOUSING UNIT, and RESIDENCY RULE).

HOUSEHOLDER (See HEAD OF HOUSEHOLD)

HOUSING DEMOGRAPHY. (See HOUSEHOLD DEMOGRAPHY).

HOUSING UNIT. Generally a shelter intended for "separate use" by its occupants, such that there is independent access to the outside and the shelter is not a group quarters. A housing unit may be occupied or vacant. (See also DOMICLE, FAMILY, GROUP QUARTERS, HOMELESS PERSONS, and HOUSEHOLD).

HOUSING UNIT METHOD. A population estimation technique found within the "Change in Stock Method" family that uses current housing unit counts, vacancy estimates, and estimates of the number of persons per household to estimate the total household population, to which can be added an estimate of the group quarters population to obtain an estimate of the total population (See also CHANGE IN STOCK METHOD, HOUSEHOLD, HOUSING UNIT, GROUP QUARTERS, and OCCUPANCY RATE POPULATION ESTIMATE, and VACANCY RATE).

IMMIGRANT. Residents of a given country entering another country in order to take up permanent residence (See also DOMESTIC MIGRATION, FOREIGN MIGRATION, and MIGRATION).

IMMIGRATION. (see FOREIGN MIGRATION).

IMPUTATION. In a sample survey or census, a general term used to describe the assignment of values to cases for which one or more variables have missing values due to "non-response." Four common methods are: (1) deductive imputation, which is based on other information available from the case in question; (2) hot-deck imputation, which is based on information from "closest-matching" cases; (3) mean-value imputation, which uses means of variables as the source of assignment; and (4) regression-based imputation, in which models are constructed using cases with no missing values and a dependent variable is the one whose missing values will be imputed and the independent variables are those that yield acceptable regression equations (See also ALLOCATION and SUBSTITUTION).

INCIDENCE RATE. The frequency with which an event, such as a new case of illness, occurs in a population at risk to the event over a given period of time

INCREMENT. The entry of an individual or set of individuals into a population of interest, where the population of interest is often defined by a model. In the case of a model of nuptiality, such an entry would be marriage (See also DECREMENT and INCREMENT-DECREMENT LIFE TABLE).

INCREMENT-DECREMENT LIFE TABLE. A life table in which there are both entries and exits to the population of interest. It is often used in reference to multiple increments and multiple decrements. That is, when there is more than one way to enter and exit a population of interest (e.g., enter via marriage and in-migration and exit via divorce, death, and out-migration). In such a life table it is potentially possible to exit and re-enter the population of interest (See also DECREMENT, INCREMENT, LIFE TABLE, and MULTI-STATE LIFE TABLE).

INDIRECT ESTIMATION. The measurement of demographic phenomena using data that do not directly represent the phenomena of interest.(See also DIRECT ESTIMATION).

INDIRECT STANDARDIZATION. The adjustment of a summary rate (e.g., the crude death rate) for a population in question found in part by computing a weighted average of group-specific rates (e.g., age-specific death rates) of a "reference" population, where the weights are the specific groups (e.g., proportion in each age group) of the population in question. This procedure is designed to produce a summary rate that controls for the effects of population composition (e.g., age) and is usually used for purposes of comparison with indirectly standardized rates for other populations computed using the same reference population. To standardize a crude (death) rate by the indirect method, first multiply the age-specific-(death) rates in the reference population by the population in the corresponding age groups of the population in question and sum the products to get the "expected" total (deaths) for the population in question. Then divide the expected total (deaths) into the total reported (deaths) for the population in question and multiply this ratio by the crude (death) rate of the reference population (See also DIRECT STANDARDIZATION and STANDARDIZATION).

INFLATION-DEFLATION METHOD. A technique that compensates for census coverage error by adjusting the demographic composition of the population of interest, but not its total number. It is sometimes used in conjunction with the cohort-component method of population projection, with the population in the launch year subject to "inflation" and the subsequent projection(s) subject to a compensating "deflation." It also is employed in the preparation of the official estimates of the population of the United States by age, sex, race, and ethnicity (Hispanic and non-Hispanic) (See also COVERAGE ERROR, COHORT-COMPONENT METHOD, LAUNCH YEAR, and POPULATION ESTIMATE).

IN-MIGRANT. A person who takes up residence within a "migration-defined" receiving area (the destination) after leaving a residence at a location outside of the receiving area (the origin), but one within the same country. For most countries, the destination and origin must be in different areas as defined by a political, administrative, or statistically-defined boundary. In the U.S., the destination must be in a different county than the origin for a person to be classified as an in-migrant by the Census Bureau (See also DESTINATION, IMMIGRANT, IN-MIGRATION RATE, MIGRANT, MIGRATION, MOVER, NET MIGRATION, NON-MIGRANT, ORIGIN, and OUT-MIGRANT).

IN-MIGRATION. (See IN-MIGRANT).

IN-MIGRATION RATE. The ratio of the number of in-migrants to a receiving area (the destination) over a given period to any one of a number of measures of the population of the receiving area, including the population at the end of the period, the population at the beginning of the period, and so on. Sometimes the denominator is formed by using an approximation of the population at risk of migrating, e.g., the national population outside of the destination (See also DESTINATION, IN-MIGRANT, MIGRATION, NET MIGRATION RATE, and OUT-MIGRATION RATE).

INTERNAL MIGRATION. (See DOMESTIC MIGRATION).

ITEM NON-RESPONSE. (See NON-RESPONSE).

INTERCENSAL. The period between two successive censuses.

INTERNATIONAL MIGRATION. The movement across an international boundary for the purpose of establishing a new permanent residence (See also DOMESTIC MIGRATION).

INTERPOLATION. The calculation of intermediate values for a given series of numbers. It is typically accomplished by using a mathematical formula, a graphic procedure, or a combination of the two. It typically imparts or even imposes a regularity to data and can, therefore, be used for smoothing, whether or not the imposed regularity is realistic (See also EXTRAPOLATION and SMOOTHING).

INTRINSIC RATE. A rate that would eventually be reached if a given population were subject to fixed mortality and fertility schedules, such that it became a "stable population" in the formal demographic sense. Intrinsic rates include the intrinsic birth rate, intrinsic death rate, and the intrinsic rate of increase. (See also BIRTH RATE, DEATH RATE, and RATE OF NATURAL INCREASE).

ITERATIVE PROPORTIONAL FITTING. A method for adjusting a multi-way distribution to a set of independently derived total values that approximates a least-squares approach. (See also CONTROLLING, CONTROLS and PLUS-MINUS METHOD).

JUMP-OFF YEAR. (See LAUNCH YEAR).

KARUP-KING METHOD. A technique used to interpolate between given points or to subdivide groups. It is based on a polynomial osculatory formula (See also INTERPOLATION).

KINSHIP NETWORK. Family support system that operates both within and outside of a household.

LAUNCH YEAR. The year in which a population projection is launched, typically the year of the most recent census. Sometimes referred to as the "Jump-Off" year, it is the starting point of the projection horizon (See also BASE PERIOD, PROJECTION HORIZON, TARGET YEAR; and POPULATION PROJECTION).

LEFT-CENSORED. (See CENSORED).

LESLIE MATRIX. An approach to population projection developed by P.H. Leslie in the late 1940s. It represents the calculations for cohort-component projections of the age distribution of the population in terms of a square matrix incorporating age-specific birth rates and survival rates and a vector containing the initial age composition of the population.

LEXIS DIAGRAM. A graphic technique developed apparently independently by several people, but largely attributed to Wilhelm Lexis (hence, the name "Lexis Diagram"), that is designed to reveal the relationship between age, time and population change, with particular applications to cohort analysis, life table construction, and population estimation (See also COHORT ANALYSIS and LIFE TABLE).

LIFE CYCLE. A sequence of significant periods through which an individual, group (e.g., family, household), or culture passes over time.

LIFE EXPECTANCY. The average number of years of life remaining to a person or group of persons who reached a given age, as calculated from a life table. (See also LIFE TABLE and SURVIVAL RATE).

LIFE TABLE. A statistical model comprised of a combination of age-specific mortality rates for a given population. A period life table (Also known as a cross-sectional life table) is constructed using mortality and age data from a single point in time; a generation life table (also known as a cohort life table) is based on the mortality of an actual birth cohort followed over time (to its extinction). A complete life table contains mortality information for single years of age, while an abridged table contains information by age group (See also INCREMENT-DECREMENT LIFE TABLE, PERIOD LIFE TABLE, LIFE EXPECTANCY and SURVIVAL RATE).

LIFE TABLE FUNCTIONS. The fundamental elements of a life table, to include the number surviving to a given age, the number of deaths to those surviving to a given birthday before they reach a subsequent birthday, the probability of dying before reaching a subsequent birthday for those who survived to a given birthday, the number alive between two birthdays, and the years of life remaining for those who survive to a given birthday (including birth). Life table functions can be interpreted in two ways: (1) as a depiction of the lifetime mortality experience of a cohort of newborns; and (2) as a stationary population that would result from a fixed mortality schedule and a constant number of annual births equal to the constant number of annual deaths resulting from the fixed mortality schedule (See also LIFE TABLE).

LIFE TABLE SURVIVAL RATE. (see SURVIVAL RATE)

LIFETIME MIGRATION. Migration that has occurred between birth and a given point in which a census or survey is conducted.

LOGISTIC CURVE. A mathematical model that depicts an "S-Shaped" curve indicative of three stages of population change: (1) an initial period of slow growth; (2) a subsequent period of rapid growth; and (3) a final period in which growth slows and comes to a halt (See also DEMOGRAPHIC TRANSITION).

LOGIT: A mathematical transformation, often used in event history analysis. For a number between 0 and 1, its logit is usually defined as the natural logarithm of the number divided by one minus the same number: $\text{logit}(n) = \ln[n/(1-n)]$, where $0 < n < 1$. Sometimes it is defined as $\text{logit}(n) = (.5)\{\ln[n/(1-n)]\}$. (See also EVENT HISTORY ANALYSIS, ODDS RATIO, and PROBIT).

LOGIT TRANSFORM. (See LOGIT).

LOGIT LIFE TABLE SYSTEM. A system of model life tables initially developed by William Brass that relies on forming logits of the proportion of deaths to those who survived to a given birthday before reaching a subsequent birthday (See also LOGIT).

LONG FORM. In the United States, the decennial census form given on a sample basis (approximately 1 in 6 households) that is designed to collect a wide range of population and housing data. The data collected go well beyond the basic information collected in the short form, which is given to the remaining households. Note, however, that the questions on the short form are contained in the long form (See also SHORT FORM).

LONGEVITY. (See LIFE SPAN).

LORENZ CURVE. Named after M. Lorenz, who introduced it in 1905, it is used to measure the distributional equality of two variables (e.g., the distribution of income across a population; the geographic distribution of the Hispanic population relative to the distribution of the non-Hispanic population over the same geography). To plot the curve, units in both variables are aggregated and the cumulative proportion of one variable is plotted against the corresponding cumulative proportion of the other, e.g., x percent of the people have y percent of all income. (See also GINI INDEX, INDEX OF DISSIMILARITY, and SEGREGATION INDEX).

MAJOR CIVIL DIVISION. A "primary" subnational political area established by law or a related process. (See also CENSUS GEOGRAPHY and MINOR CIVIL DIVISION).

MALTHUSIAN GROWTH. The hypothesis that unless negative checks (i.e., T. Malthus's idea of "moral restraint") are introduced, a population increases geometrically until some type of positive check is imposed (i.e., famine, war, pestilence).

MARKOV PROCESS (also known as a MARKOV CHAIN). A systems model named after A. A. Markov that is specified by transition probabilities between the different states of the system, where the transition probabilities are dependent solely on the present distribution of the population in these states (See also MULTI-STATE LIFE TABLE).

MARRIAGE. The social institution involving legal and/or religious sanction whereby men and women are joined together for the purpose of founding a family unit. In some countries, marriage includes couples joined for purposes other than founding a family unit (See also DIVORCE, SINGLE, and WIDOWED).

MASTER ADDRESS FILE (MAF). In the United States, the set of records maintained by the Census Bureau for purposes of conducting the decennial census. It is intended to represent the geographic location of every housing unit.

MATCHED GROUPS. A group constructed on a case-by-case basis through matching of sets of records according to a limited number of characteristics.

MATCHING (of Records). Assembly of data in a common format from different sources but pertaining to the same unit of observation, (e.g., a person, household, or an event such as death). Also known as Record Matching and Data Linkage (See also CAPTURE-RECAPTURE and DUAL-SYSTEMS ESTIMATION).

MEAN POPULATION AGE. The average age of all members of a population.

MEDIAN POPULATION AGE. The age at which a population is divided into two equally-sized groups.

METROPOLITAN AREA. In the United States, this refers to a family of specific census geographies intended to represent a large population nucleus and aggregations thereof. Specific types of include “Primary Metropolitan Statistical Area” and “Standard Consolidated Statistical Area.” (See also CENSUS GEOGRAPHY, PRIMARY METROPOLITAN STATISTICAL AREA, and STANARD CONSOLIDATED STATISTICAL AREA).

MIGRANT. A person who makes a relatively permanent change of residence from one country, or region within a country (an origin), to another (the destination) during a specified (migration) period. For most countries, the change must be across a political, administrative, or statistically-defined boundary for a person to be classified as a migrant. In the U.S., the origin and destination must be in different counties for a person to be classified as a migrant (See also DESTINATION, EMIGRANT, IMMIGRANT, IN-MIGRANT, MIGRATION, MOVER, NON-MIGRANT, ORIGIN, and OUT-MIGRANT).

MIGRATION. A general term for the incidence of movement by individuals, groups or populations seeking to make relatively permanent changes of residence. One of the components of population change (See also ASYLEE, COMMUTING, COMPONENTS OF CHANGE, DESTINATION, DOMESTIC MIGRATION, EMIGRANT, FOREIGN-BORN, GROSS MIGRATION, IMMIGRANT, IN-MIGRANT, INTERNALLY DISPLACED PERSONS, INTERNATIONAL MIGRATION, MIGRANT, MOBILITY, MOVER, NATIVE, NET MIGRATION, NON-MIGRANT, ORIGIN, OUT-MIGRANT, and REFUGEE).

MIGRATION HISTORY. Information obtained in a census or a sample survey that provides lifetime migration data (See also LIFETIME MIGRATION).

MIGRATION PREFERENCE INDEX. As defined by R. Bachi in 1957, the ratio of the actual to the expected number of migrants to a given area, where the expected number is directly proportional to both the population at the origin and the destination.

MIGRATION STATUS. (See MOBILITY STATUS).

MIGRATION STREAM. A group of migrants with a common origin and destination over a given period. (See also COUNTERSTREAM).

MILITARY POPULATION. Persons who are members of the armed forces.

MILITARY DEPENDENT POPULATION. Persons who are dependents of members of the armed forces.

MINOR CIVIL DIVISION. A “secondary” subnational political area established by law or a related process (See also CENSUS GEOGRAPHY and MAJOR CIVIL DIVISION).

MOBILITY, GEOGRAPHIC. Any move resulting in a change of residence (See also DOMESTIC MIGRATION and MIGRATION).

MOBILITY RATE. The ratio of the number of movers over a given time period to the population at risk of moving over the same period (See also IN-MIGRATION RATE, MIGRATION, and OUT-MIGRATION RATE).

MOBILITY STATUS. A Classification of people based on their residential locations at the beginning and end of a given time period.

MODEL. A generalized representation of a demographic process, set of demographic relationships, pattern of mortality, fertility, migration, or marriage, or method of population estimation or projection.

MODEL LIFE TABLE. A life table based on the generalization of empirical relationships derived from a group of observed life tables.

MOMENTUM OF POPULATION GROWTH. The tendency of a population to increase for as many as 70 years after reaching replacement level fertility. It may be measured as the projected percent increase in the population between the current level and the level when it is projected to become stationary in the absence of migration (See also STATIONARY POPULATION).

MORTALITY. A general term for the incidence of deaths in a population or group. One of the components of population change (See also COMPONENTS OF CHANGE).

MOVER. A person who reports in a census or survey that he or she lived at a different address at an earlier date (e.g., five years before the census or survey). In the U.S., a mover is classified by the Census Bureau as a person who changed residence, but within the same county (See also MIGRATION).

MULTIPLE DECREMENT TABLE. (See MULTI-STATE LIFE TABLE).

MULTIPLE INCREMENT TABLE. (See MULTI-STATE LIFE TABLE).

MULTIPLE INCREMENT-DECREMENT TABLE. (See MULTI-STATE LIFE TABLE).

MULTIPLE INCREMENT-DECREMENT LIFE TABLE. (See MULTI-STATE LIFE TABLE).

MULTI-REGIONAL ANALYSIS. An analysis of multi-regional systems in which spatial and demographic factors are linked.

MULTI-STATE LIFE TABLE. An extension of the standard life table in which multiple transitions between states are possible and the transitions are expressed in terms of transition probabilities between states (See also DECREMENT, INCREMENT, INCREMENT-DECREMENT LIFE TABLE, and MARKOV PROCESS).

NATIVE Persons born in a particular country or region as distinguished from foreign-born.

NATURAL DECREASE. (See NATURAL INCREASE).

NATURAL FERTILITY. The level of fertility in a population in which deliberate control of childbearing (e.g., contraception, abstinence) is not practiced.

NATURAL INCREASE. The excess of births over deaths in a population is defined as natural increase; an excess of deaths over births is defined as natural decrease.

NATURAL VACANCY RATE. A concept that represents the average vacancy rate for a rental property market that would result if supply and demand were in balance, often a judgmental benchmark against which observed vacancy rates in a given market can be compared (See also VACANCY RATE).

NET CENSUS UNDERCOUNT ERROR. The estimated level of coverage error in a census computed by algebraically adding estimated overcounts and estimated undercounts for population groups (e.g., age-sex-race) and summing them. (See also COVERAGE ERROR, NON-RANDOM ERROR and TRUE POPULATION)

NET MIGRATION. The difference between the number of in-migrants and the number of out-migrants for a given area (e.g., a county) over a given period of time: Net = In - Out (See also GROSS MIGRATION, IN-MIGRANT, MIGRATION, NET MIGRATION RATE, and OUT-MIGRANT).

NET MIGRATION RATE. The ratio of net migration for a given area (e.g., a county) over a given period to any one of a number of measures of the population of the area, including the population at the end of the period, the population at the beginning of the period, and so on. Sometimes the denominator is formed by using a population outside of the area (e.g., the national population outside of the county) (See also IN-MIGRATION RATE, MIGRATION, NET MIGRATION, and OUT-MIGRATION RATE).

NET NUMBER OF MIGRANTS (See NET MIGRATION).

NON-FAMILY HOUSEHOLD. One or more persons living together in a housing unit that are not related by blood or marriage (See also FAMILY and HOUSEHOLD)

NON-METROPOLITAN POPULATION. The number of people living outside large urban settlements. In the U.S., this represents the population outside Metropolitan Statistical Areas (See also CENSUS GEOGRAPHY).

NON-MIGRANT. In a census or survey, an individual who resided in an area both at the beginning and end of the designated migration period. Alternatively, an individual who has neither migrated into nor migrated out of his or her area of residence. (See also IN-MIGRANT, MIGRATION, MOVER, NET MIGRATION and OUT-MIGRANT).

NON-RANDOM ERROR. All errors not due to the effects of random sample selection (i.e., random error). It can occur both in a sample survey and in a population census. Examples include non-response, incorrect answers by a valid respondent, answers given by a non-valid respondent, as well as coding and other processing errors. Statistical inference can only be used to estimate random error, not non-random error (See also NET CENSUS UNDERCOUNT ERROR, NON-RESPONSE, POPULATION, RANDOM ERROR, SAMPLE, and TOTAL ERROR).

NON-RESPONSE. Missing data on a form used in a survey or census due to a number of reasons, including the refusal of a respondent to answer, the inability to locate a potential respondent, the inability of a respondent (or informant) to answer questions, or the omission of answers due to a clerical or some other form of error. Total non-response refers to a case (i.e., an observation) in which all variables have missing values and item non-response refers to a case in which fewer than all variables have one or more missing values. Imputation is often used to estimate values for cases in which they are missing (See also IMPUTATION, NON-RANDOM ERROR, NON-RESPONDENT)

NON-RESPONSE ERROR. (See NON-RESPONSE).

NON-RESPONDENT. In a sample survey or census, a respondent who refuses to be interviewed, or is otherwise unable to take part (See also NON-RESPONSE)

OCCUPIED HOUSING UNIT. A Housing unit that is not vacant. In principle, a household. (See also, HOUSEHOLD, OCCUPANCY RATE, RESIDENCY RULE, and VACANT HOUSING UNIT).

OCCUPANCY RATE. The ratio of occupied housing units to the total number of housing units, usually multiplied by 100 so that this ratio is expressed as a percent. (See also HOUSING UNIT, OCCUPIED HOUSING UNIT, and VACANCY RATE).

OCCURRENCE-EXPOSURE RATIO. The ratio of the number of events occurring during a given period to the population at risk during the same period. The population at risk may be measured in different ways.

ODDS RATIO. As defined for a dichotomous variable, the ratio of the proportion of the population having a characteristic of interest to the proportion not having the characteristic. For example, the percent of the population to the percent not in poverty The logarithm of the odds ratio is termed a logit. (See also LOGIT).

OPEN-ENDED INTERVAL. A class interval in a distribution of grouped data that is not bounded on one end. For example, in a distribution of data on income, the highest income class may be given as $100,000 or more; in a life table the last age interval may be given as 85 years and over. In a longitudinal analysis, the period between the most recent occurrence of an event of interest (e.g., a live birth) and a subsequent time point . For example, in a survey of birth histories, the period between the second birth and the survey would constitute an open-ended interval for a woman reporting two births, whereas the periods between her first and second birth would be a closed interval.

OPEN INTERVAL. (See OPEN-ENDED INTERVAL).

OPEN LIVE-BIRTH INTERVAL. The time elapsed since the most recent birt, typically measured as an average for a group of women; an index that directly reflects the effect of increased spacing between births.

ORIGIN. The place of residence that a migrant left at the start of a given (migration) period (See also DESTINATION, MIGRANT, and MIGRATION).

ORPHANHOOD METHODS. A set of survey and census-based techniques for estimating adult mortality in a population that lacks reliable mortality data. It is based on identifying the proportion of respondents with living mothers and fathers. (See also SIBLING METHODS).

OSCULATORY INTERPOLATION. An interpolation method that involves combining higher-order polynomial formulas into one equation, designed to provide a smooth junction between two adjacent groups of data (e.g., age group 5-9 and age group 10-14). (See also INTERPOLATION).

OUT-MIGRANT. A person who leaves his or her residence in a "migration-defined" sending area (the origin) to take up residence at a location outside of the sending area (the destination), but one within the same country. For most countries, the origin and destination must be in different areas as defined by a political, administrative, or statistically-defined boundary In the U.S., the origin must be in a different county than the destination for a person to be classified as an out-migrant by the Census Bureau (See also DESTINATION, EMIGRANT, IN-MIGRANT, MIGRANT, MIGRATION, MOVER, NET MIGRATION, NON-MIGRANT, ORIGIN, and OUT-MIGRATION RATE)

OUT-MIGRATION.(See INTERNAL MIGRATION).

OUT-MIGRATION RATE. The ratio of the number of out-migrants from a sending area (the origin) over a given period to some measure of the population of the sending area, including the population at the beginning of the period, the population at the end of the period, and so on. (See also ORIGIN, OUT-MIGRANT, IN-MIGRATION RATE, MIGRATION, and NET MIGRATION RATE)

OVERCOUNT. In a census, this can be due to counting some people more than once, counting people in a census who are not members of the population in question, or a combination of both (See also NET CENSUS UNDERCOUNT ERROR and UNDERCOUNT).

OWN-CHILD METHOD. A census or survey-based method for measuring fertility that uses counts of children living with their mothers.

PARTIAL MIGRATION RATE. The number of in-migrants from a particular origin to a given destination relative to the population of either the origin or destination.

PARTICIPATION RATE. The proportion of a population or segment of a population with a certain characteristic, usually social or economic, e.g., the proportion aged 10-14 who are enrolled in school.

PARITY SPECIFIC BIRTH RATE. Live births to women of specific parities.

PERCENT. (See PROPORTION).

PERIOD ANALYSIS. The analysis of demographic data observed during a brief period of time (usually one year), such as death registrations, or a single date, such as census data on marital status, and sometimes at several points of time. The data are typically comprised of more than one cohort (See also AGE-PERIOD-COHORT EFFECT, COHORT ANALYSIS, PERIOD EFFECT, and PERIOD MEASURE).

PERIOD EFFECT. An analytical perspective that attempts to determine the effect of a period event (e.g., a war, famine, or natural disaster) on a variable of interest (See also AGE EFFECT, AGE-PERIOD-COHORT EFFECT, COHORT EFFECT, and PERIOD ANALYSIS).

PERIOD MEASURE. A summary measure of data collected during a brief period of time (usually one year) that typically represent more than one cohort (See also COHORT MEASURE and PERIOD ANALYSIS).

PERIOD LIFE TABLE. A life table based on mortality data collected at a given point in time (a year) or a short period (two or three years) for a given population (See also GENERATION LIFE TABLE and LIFE TABLE).

PERSON-YEARS LIVED. The total number of years (and fractions thereof) lived by a given population or population segment during a given period of time. It is approximated by computing the product of: (1) the number of persons in the population or population segment; and (2) the amount of time in years (and fractions thereof) lived by these same persons during the time in question (See also AT-RISK POPULATION and LIFE TABLE).

PERSONS PER HOUSEHOLD. The ratio of the total number of persons living in households to the total number of households, often referred to as the average number of persons per household. (See also HOUSEHOLD and HOUSING UNIT METHOD).

PLACE OF RESIDENCE. (See USUAL RESIDENCE)

PLUS-MINUS METHOD. A "controlling" technique that attempts to compensate for both increasing and decreasing subsets of a population of interest by using two separate adjustment factors. For example, one might use the plus-minus method in adjusting post censal population estimates of census tracts to an estimate of the county containing the tracts, if some tracts show growth since the last census and others show decline (See also CONTROLLING, CONTROLS, and ITERATIVE PROPORTIONAL FITTING).

POPULATION. In the demographic sense, the "inhabitants" of a given area at a given time, where inhabitants could be defined either on the De Facto or De Jure basis (but not a mixture of both). Note that the concept of "area" can be generalized beyond the geographical sense to include, for example, formal organizations. In the statistical sense, the term "population" refers to the entire set of persons (or phenomenon) of interest in a particular study, as compared to a sample, which refers to a subset of the whole (See also CENSUS, DE FACTO POPULATION, DE JURE POPULATION, DEMOGRAPHY, SAMPLE, and SPECIAL POPULATION)

POPULATION AT RISK. (See AT-RISK POPULATION).

POPULATION COMPOSITION. The classification of members of a population by one or more characteristics such as age, sex, race, and ethnicity. It can be presented in either absolute or relative numbers. "Population distribution" and "population structure" are often used as synonyms (See also POPULATION DISTRIBUTION).

POPULATION DECREASE. Reduction in the number of inhabitants in an area.

POPULATION DENSITY. Number of persons per unit of land area.

POPULATION DISTRIBUTION. Usually used to refer to the location of a population over space at a given time, but sometimes used as a synonym for population composition (See also POPULATION COMPOSITION)

POPULATION DYNAMICS. Changes in population size and structure due to fertility, mortality, and migration, or the analysis of population size and structure in these terms.

POPULATION ESTIMATE. An approximation of a current or past population of a given area at a given time, or its distribution and composition, in the absence of a complete enumeration, ideally done in accordance with one of two standards for defining a population, De Facto or De Jure. In the United States, most population estimates are done using data that represent the set of Census defined residents, although some may be done that represent the De Facto population and modifications of it (See also ADMINISTRATIVE RECORDS METHOD, CENSAL-RATIO METHOD, CENSUS, CENSUS DEFINED RESIDENT, CHANGE IN STOCK METHOD, COMPONENT METHOD, COMPOSITE METHOD, DE FACTO POPULATION, DE JURE POPULATION, HOUSING UNIT METHOD, POPULATION PROJECTION, RATIO-CORRELATION METHOD, RATIO ESTIMATION, RESIDENCY RULES, SYNTHETIC METHOD, and VITAL RATES METHOD).

POPULATION FORECAST. An approximation of the future size of the population for a given area, often including its composition and distribution. A forecast usually is one of a set of projections selected as the most likely representation of the future (See also POPULATION ESTIMATE and POPULATION PROJECTION).

POPULATION MOMENTUM. (See MOMENTUM OF POPULATION GROWTH).

POPULATION PROJECTION. The numerical outcome of a particular set of implicit and explicit assumptions regarding future values of the components of population change for a given area in combination with an algorithm. Strictly speaking, it is a conditional statement about the size of a future population (often along with its composition and distribution), ideally made in accordance with one of the two standards used in defining a population, De Facto or De Jure (See also BASE PERIOD, CENSUS, CENSUS DEFINED RESIDENT, COHORT-COMPONENT METHOD, DE FACTO POPULATION, DE JURE POPULATION, LAUNCH YEAR, POPULATION FORECAST, POPULATION ESTIMATE, PROJECTION HORIZON, and TARGET YEAR).

POPULATION PYRAMID. A graphic device that shows the age-sex composition of a given population and possibly other characteristics as well. It is in pyramidal form, ranging from an equilateral triangle to a near-rectangle, with the shape determined by the effects of the components of population change on a prior age-sex composition.

POPULATION REGISTER. An administrative record system used by many countries (e.g., China, Finland, Japan, and Germany) that requires residents to register their place of residence, usually at a local police station. By itself, such a system provides limited demographic information (e.g., total population), but where it can be matched to other administrative record systems (e.g., tax, social and health care services), the result is often a system that provides a wide range of longitudinal and cross-sectional demographic information.

POPULATION REPLACEMENT. In general, the process of renewal by which a population replaces losses from deaths by means of births. In stable population theory, it refers to the extent to which women in the population are being replaced over the course of a generation or a year. It is measured by the net reproduction rate for a generation and by the intrinsic rate of increase on an annual basis. These measures allow for both the level of fertility and the level of mortality through the childbearing ages. Exact replacement requires a net reproduction rate of 1.00.

POPULATION SIZE. The number of persons inhabiting a given area at a given time. (See also CENSUS and POPULATION).

POPULATION STATISTICS. These are generally comprised of vital statistics, migration statistics, and census and survey data, but they vary by country in that not all countries collect all types of data. They also may include administrative record data, including population register data,

POPULATION STRUCTURE. (See POPULATION COMPOSITION).

PREVALENCE. The number of persons who have a given characteristic (e.g., disease, contraceptive use, impairment, labor force participation) in a given population at a designated time or who had the characteristic at any time during a designated period, such as a year (See also PREVALENCE RATE).

PREVALENCE RATE. The proportion of persons in a population who have a particular disease or attribute at a specified time (point prevalence) or at any time during a designated period, such as a year (period prevalence). (See also PREVALENCE).

PRIMARY METROPOLITAN STATISTICAL AREA. In the United States, a census-based piece of geography defined by the Office of Management and Budget that is comprised of a central city and county and adjoining counties linked to the central city by social and economic interactions that meet prescribed standards. (See also CENSUS GEOGRAPHY, METROPOLITAN AREA, and STANDARD CONSOLIDATED AREA).

PROBABILITY. A ratio in which the numerator consists of those in a population experiencing an event of interest (e.g., death) over a specified period of time, while the denominator consists of the at-risk population. (See also AT-RISK POPULATION, PROPORTION, RATE, and RATIO).

PROBIT. A mathematical transformation, often used in event history analysis, for "linearizing" the cumulative normal distribution of a variable of interest. The probit unit is $y = 5 + Z(p)$, where p = the prevalence of response at each dose level and $Z(p)$ = the corresponding value of the standard cumulative normal distribution (See also EVENT HISTORY ANALYSIS and LOGIT).

PROJECTION. (See POPULATION PROJECTION).

PROJECTION HORIZON. In a population projection, the period between the launch year and the target year (See also Base Period, Launch Year, and Target Year; and POPULATION PROJECTION).

PROPORTION. A ratio used to describe the status of a population with respect to some characteristic (e.g., married), where the numerator is part of the denominator. When multiplied by 100, a proportion is known as a “percent.” (See also PROBABILITY, RATE, and RATIO).

PROPORTIONAL HAZARDS. The examination of “exogenous” explanatory variables in the analysis of survival time and related data (See also PROPORTIONAL HAZARDS MODEL and SURVIVAL)

PROPORTIONAL HAZARDS MODEL. A regression-based approach to the analysis of survival or duration data designed to examine survival time, failure time, or other duration data in terms of the effect of exogenous explanatory variables (See also PROPORTIONAL HAZARDS)

PUBLIC USE MICRODATA SAMPLE (PUMS). In the United States and elsewhere this usually refers to a hierarchically-structured data set that contains individual, family, and household information in a given record and for which confidentiality is maintained by deleting identifying information. It is typically obtained by sampling from census records.

QUASI-STABLE POPULATION. A population not affected by migration with constant fertility and gradually changing mortality. It also is used to refer specifically to a formerly stable population in which fertility remains constant but mortality is gradually changing (See also STABLE POPULATION).

RADIX (OF A LIFE TABLE). A hypothetical cohort of newborns used as the starting point of a life table, typically 100,000.

RAKING. (See CONTROLLING).

RANDOM ERROR. The difference between a statistic of interest (e.g., mean age) found in a sample unaffected by non-random error and its corresponding parameter (e.g., mean age) found in the population from which the sample was drawn. Random error can only occur in a sample, never in a population. It is often referred to as sample error or sampling error (See also NON-RANDOM ERROR, POPULATION, SAMPLE, and TOTAL ERROR).

RATE. Technically, this type of ratio is the same as a probability. However, the term is often applied to the type of ratio known as a proportion, as in the case of “vacancy rate,” which is the ratio of unoccupied housing units to all housing units. It is also applied to other types of ratios in which the denominators are not precisely the “at-risk populations," as is the case of the crude birth rate (See also AT-RISK POPULATION, PROBABILITY, PROPORTION, and RATIO).

RATE-CORRELATION METHOD. (See RATIO-CORRELATION METHOD).

RATE OF CHANGE . The change of population during a given period express as a rate. The rate may relate to the entire period, in which case the denominator is usually the initial population. Alternatively, it may be an average annual rate, in which case the rate may assume annual compounding, continuous compounding, or some other function (See also POPULATION CHANGE)

RATE OF NATURAL INCREASE. The result of subtracting the crude death rate from the crude birth rate. For a population closed to migration it provides the rate of increase (or the rate of decrease if the crude death rate exceeds the crude birth rate) (See also CRUDE BIRTH RATE, CRUDE DEATH RATE, and INTRINSIC RATE).

RATIO. A single number that expresses the relative size of two other numbers - i.e., a quotient, which is the result of dividing one number by another. (See also PROBABILITY, PROPORTION, and RATE).

RATIO-CORRELATION METHOD. A regression-based subnational population estimation technique found within the "Change in Stock Method" family. Introduced by R. Schmitt and A. Crosetti in the early 1950s: (1) the dependent variable consists of the ratio formed by dividing the most recent population proportion for a set of sub-areas (e.g., proportion of a state population in each of its counties at the most recent census) by the population proportion for the same subareas at an earlier time (i.e., the previous census); and (2) the independent variables consist of corresponding ratios of proportions for symptomatic indicators of population (e.g., school enrollment, automobile registrations, births, deaths) available from administrative records. Variations of the Ratio-Correlation Method include the Difference-Correlation Method introduced by R. Schmitt and J. Gier in 1966 and the Rate-Correlation Method introduced by D. Swanson and L. Tedrow in the 1984 (See also CENSAL-RATIO METHOD, CHANGE IN STOCK METHOD, POPULATION ESTIMATE, and WEIGHTED AVERAGE).

RATIO ESTIMATION. A set of techniques used to estimate population based on ratios across geographic areas, variables, or both. (See also POPULATION ESTIMATE).

RECORD LINKAGE. (See MATCHING).

RECORD MATCHING. (See MATCHING).

REFERENCE POPULATION. (See STANRARD POPULATION).

RESIDENCE. The place where a person lives. Defined differently in different censuses, but often interpreted as "usual residence," which is the case in the U.S. decennial census based on the De Jure method (See also AMERICAN COMMUNITY SURVEY,

AMERICAN COMMUNITY SURVEY-DEFINED RESIDENT, CENSUS, CENSUS-DEFINED RESIDENT, DE FACTO POPULATION , DE JURE POPULATION, DOMICLE, HOMELESS PERSON, HOUSEHOLD, GROUP QUARTERS, RESIDENCY RULES, and USUAL RESIDENCE).

RESIDENCY RULES. A set of criteria used directly in an enumeration such that everybody is counted once, and only once. The criteria are used indirectly for purposes of estimation. In principle, the Decennial Census is based on the De Jure concept, which translates as a person's place of usual residency, which adds "in the right place" to the objective that everybody should be counted once and only once. The American Community Survey is based on a modification of the De Facto concept that accepts a person as a resident if that person has lived in a given location for two months. As such, it tends to ignore the "in the right place" feature found in the Decennial Census (See also AMERICAN COMMUNITY SURVEY- DEFINED RESIDENT, CENSUS DEFINED RESIDENT, DE FACTO POPULATION, DE JURE POPULATION, HOUSEHOLD, and POPULATION).

RESIDENTIAL MOBILITY. A change of residence, either in the same city or town, or between cities, states, countries, or communities.

RESIDUAL METHOD. A technique that estimates intercensal net migration for a given area by subtracting from the most recent census count, the algebraic sum of intercensal births and deaths added to the population counted at the preceding census. Resulting estimates are confounded by differences in net census undercount error (See also BALANCING EQUATION, ERROR OF CLOSURE, and NET MIGRATION).

RETURN MIGRATION. A move back to point of origin, whether domestic or foreign (See also MIGRATION).

REVERSE RECORD CHECK. A technique used to estimate census coverage error that attempts to match a sample drawn from a reliable source of records independent of the census with data collected in the census. For example, a reverse record check may attempt to match a sample of births over a 10-year period with children under 10 in the census, or a sample of enrollees under Medicare with the elderly population in the census (See also CENSUS and COVERAGE ERROR).

REVERSE STREAM. (See COUNTERSTREAM)

REVERSE SURVIVAL METHOD. Any method of estimating population involving backward "survival" of a population to an earlier date (See also SURVIVAL RATE).

RIGHT-CENSORED. (See CENSORED).

RISK ASSESSMENT. The qualitative or quantitative estimation of the likelihood of adverse effects attributable to exposure to specified health hazards or medical procedures or treatments, such as contraceptives.

RURAL POPULATION. Usually defined as the residual population after the urban population has been identified (See also URBAN POPULATION).

RURAL-URBAN MIGRATION. The migration from rural to urban areas, both internal and international.

SAMPLE. A subset of a population (in the statistical sense) for which data are typically collected in a "survey," which is a way of providing respondents with questions to be answered (e.g., through personal interviews, telephone interviews, mail-out/mail-back questionnaires). Samples may also be selected from administrative and other records such that interviews are not needed because data are taken directly from the records themselves (e.g., from Medicare files). Samples may be defined in a number of ways, but if statistical inference is to be used, a sample's elements should have a known probability of selection, or at least a reasonable approximation thereof, so that "random error" can be estimated (See also CENSUS, NON-RANDOM ERROR, POPULATION, and RANDOM ERROR).

SAMPLE ERROR. (See RANDOM ERROR).

SCHOOL-AGE POPULATION. Children if school age, usually defined by the ages for which school attendance is compulsory, which varies from country to country and sometimes with a given country.

SEASONAL ADJUSTMENT. A statistical modification to a data series to reduce the effect of seasonal variation (See also SEASONAL VARIATION).

SEASONAL VARIATION. Seasonal differences in the occurrence of data collected over time and reported at least quarterly (See also SEASONAL ADJUSTMENT)

SEGREGATION. There are many different interpretations of the term "segregation" and at least five dimensions of segregation have been identified: (1)centralization; (2) concentration; (3) clustering; (4) evenness; and (5) exposure. In demography, segregation usually refers to the spatial separation or isolation of a race, ethnic, or socioeconomic group by residence (See also SEGREGATION INDEX)

SEGREGATION INDEX. There are at least 20 indices that have been developed to measure the different dimensions of segregation, many of which are algebraically related to one another. Those typically used in demography are concerned with residential segregation, with the most common being the Index of Dissimilarity (See also GINI INDEX, INDEX OF DISSIMILARITY, LORENZ CURVE, and SEGREGATION).

SELF-ENUMERATION. A method of conducting a census or sample survey in which respondents fill out questionnaire themselves, usually in connection with a mail-out/mail-back design for distributing and retrieving the questionnaires.

SEPARATION FACTORS. The proportions used to assign deaths at each age in each calendar year to birth cohorts in connection with the calculation of probabilities of dying, especially for constructing life tables. Special separation factors are applied to infant deaths because of the tendency for deaths to be concentrated in the earliest days, weeks, and months of infancy (See also LIFE TABLE FUNCTIONS).

SETTLEMENT AND RESETTLEMENT. The relocation of refugees and other displaced persons in a new place. (See also ASYLEES, INTERNALLY DISPLACED PERSONS, and REFUGEES).

SEX COMPOSITION. (See POPULATION COMPOSITION).

SEX DISTRIBUTION. (See SEX COMPOSITION)

SEX RATIO. The ratio of males to the number of females in a population, usually computed for age groups and expressed per 100 females.

SEX STRUCTURE. (See POPULATION COMPOSITION).

SHORT FORM. In the United States, the decennial census form asking a limited range of basic population and housing questions and distributed to about five-sixths of the households, with the so-called "long form" being distributed to the remaining households. Note, however, that the questions on the short form are contained in the long form, so in effect all households receive the short form (See also LONG FORM).

SIBLING METHOD. A set of survey and census-based techniques for measuring mortality in a population that lacks otherwise reliable mortality data. It is based on asking respondents for dates of birth and ages at death (if applicable) of brothers and sisters with living mothers and fathers. (See also ORPHANHOOD METHODS).

SINGLE. A general term for a person not currently married. It could be applied to a person who has never been married or a person who is divorced or widowed and not yet re-married (See also MARRIAGE).

SMALL AREA. The subdivisions of the primary political subdivisions of a country. In the United States, counties and their subdivisions are usually considered small areas, although some limit the term to subcounty areas such as census tracts, block groups, and blocks and the areas that can be aggregated from them (See also CENSUS GEOGRAPHY).

SMOOTHING. The adjustment of data to eliminate or reduce irregularities and other anomalies assumed to result from measurement and other errors. A common application of smoothing procedures is in connection with single-year-of-age data that appear to be affected by age heaping (See also AGE-HEAPING and INTERPOLATION).

SPECIAL POPULATION. Population groups identified separately for purposes of a census and or sample survey because of their distinctive living arrangements, such as college students, prison inmates, residents of nursing homes, and military personnel and their dependents. Special populations usually are characterized by components of change very different from the broader populations in which they are found, sometimes because of laws or regulations governing them. (See also COMPONENTS OF CHANGE and POPULATION).

STABLE POPULATION. A population with an unchanging relative age composition and a constant rate of change in its total size, resulting from conditions of constant fertility and mortality rates over an extended period, about 70 years (See also QUASI-STABLE POPULATION, POPULATION COMPOSITION, STABLE POPULATION METHOD, and STATIONARY POPULATION).

STABLE POPULATION METHOD. The use of a "reference" stable population and its parameters approximating the conditions of an observed population to evaluate and estimate the composition and the fertility and mortality levels of an observed population of interest (See also STABLE POPULATION).

STANDARD CONSOLIDATED AREA. In the United States, a combination of Primary Metropolitan Statistical Areas, with a total population of at least 1,000,000, established by the Office of Management and Budget. (See also CENSUS GEOGRAPHY, METROPOLITAN AREA, and PRIMARY METROPOLITAN STATISTICAL AREA).

STANDARD LIFE TABLE. A life table against which values from another life table are compared, or from which a life table for a population of interest is constructed. The term also is used to refer to the conventional life table, representing the diminution of a cohort of births through age-specific death rates of a particular year or short group of years without additional decrements or any increments (See also LIFE TABLE).

STANDARD METROPOLITAN STATISTICAL AREA. (See PRIMARY METROPOLITAN AREA).

STANDARD POPULATION. A "reference" population used for purposes of analyzing a population of interest. Also, specifically, a population whose age distribution is employed in the calculation of standardized rates by the direct method. (See also DIRECT STANDARDIZATION and STANDARDIZATION)

STANDARDIZATION. The adjustment of a summary rate (e.g., the crude death rate) to remove the effects of population composition (e.g., age), usually done to compare rates across populations with different compositions. There are two general types of standardization, direct and indirect. The type selected is dependent on the data available for the population(s) of interest (See also DIRECT STANDARDIZATION, INDIRECT STANDARDIZATION, POPULATION COMPOSITION, STANDARD POPULATION, and STANDARDIZED RATE).

STANDARDIZED RATE. A rate that has been subjected to standardization. (See also STANDARDIZATION).

STATIONARY POPULATION. A stable population in which the rate of increase is zero and the total size and both the absolute and relative age composition are constant (See also MOMENTUM OF POPULATION GROWTH and STABLE POPULATION).

STATIONARITY. The condition where a population is stationary. (See also STATIONARY POPULATION)

SUBSTITUTION. In a sample survey or census, the process of assigning values for a case in which there is "total non-response." Many substitution methods are available, including automated algorithms (See also ALLOCATION and IMPUTATION).

SUBURBAN. A popular term referring to the residential area surrounding a central city. Such an area may follow the transportation lines and be dependent on the central city both economically and culturally but, increasingly, such areas are becoming the equivalent of central cities to suburbs of their own. (See also URBAN FRINGE)

SUBURBANIZATION. The spatial diffusion of population growth affecting areas adjoining a city.

SURVEY. (See SAMPLE).

SURVIVAL. Primarily a condition where an individual or group remains alive after a specified interval, and secondarily a condition where an individual or group maintains membership in the group of interest, such as a school enrollment cohort, marriage cohort, or the non-poor population. (See also SURVIVAL RATE).

SURVIVAL CURVE. A graph depicting a survivorship function (See also SURVIVORSHIP FUNCTION)

SURVIVAL RATE. A rate expressing the probability of survival of a population group, usually an age group, from one date to another and from one age to another. A survival rate can be based on life tables or two censuses. When based on two censuses, the rate includes not only the effects of mortality, but also the effects of net migration and relative census enumeration error. (See also FORWARD SURVIVAL RATE, HAMILTON-PERRY METHOD, LIFE TABLE, SURVIVAL, and SURVIVORSHIP FUNCTION).

SURVIVORSHIP. (See SURVIVAL).

SURVIVORSHIP FUNCTION. The probability that an individual survives to at least time = t before an event of interest (e.g., death) occurs. It is one of three algebraically related functions used in survival analysis, the other two being the "death density function" and the "hazard function." The survivorship function is found by dividing the death density function by the hazard function (See also DEATH DENSITY FUNCTION, HAZARD FUNCTION, and SURVIVAL RATE).

SURVIVORSHIP RATIO. (See SURVIVAL RATE).

SYNTHETIC METHOD. A member of the family of ratio estimation methods that is used to estimate characteristics of a population in a subarea (e.g., a county) by reweighting ratios (e.g., prevalence rates or incidence rates) obtained from survey or other data available at a higher level of geography (e.g., a state) that includes the subarea in question. (See also POPULATION ESTIMATE, RATIO ESTIMATION and WEIGHTED AVERAGE).

TARGET YEAR. In a population projection, the final year for which a population is projected, the end point of the projection horizon (See also BASE PERIOD, LAUNCH YEAR, and PROJECTION HORIZON; and POPULATION PROJECTION).

TEMPORARY MIGRATION. A type of migration, both internal and international, in which the duration of stay is temporary. Data for temporary migration are not normally included in the official data on internal or international migration and are usually obtained from a special sample survey.

TIGER. (See TOPOLOGICALLY INTEGRATED GEOGRAPHIC ENCODING AND REFERENCING SYSTEM).

TOPOLOGICALLY INTEGRATED GEOGRAPHIC ENCODING AND REFERENCING SYSTEM. (TIGER). A digital database of geographic features (e.g., roads, rivers, political boundaries, census statistical boundaries, etc.) covering the entire United States. It was developed by the U.S. Census Bureau to facilitate computerized mapping and areal data analysis. (See also GEOGRAPHIC INFORMATION SYSTEM).

TOTAL ERROR. In a sample, the theoretical sum of random error and non-random error, which in practice can at best only be roughly approximated because of the difficulty of estimating non-random error. Also known as Total Sample Error. In a census, total error is comprised solely of non-random error (see also NON-RANDOM ERROR, RANDOM ERROR, and TRUE POPULATION).

TREND EXTRAPOLATION: (See EXTRAPOLATION).

TRUE POPULATION. In theory, the population that would be counted if there were no errors in a census. In practice, it is a value representing the theoretical actual number for the population at a given date, which cannot be precisely measured, but which can be roughly approximated by adjusting a census for net census undercount error (See also CENSUS and NET CENSUS UNDERCOUNT ERROR)

TRUNCATION BIAS. Distortion of results due to the systematic omission from an analysis of values that fall below or above a given range.

TURNOVER. A term sometimes employed to refer to the sum of the components of change during a period, i.e., births plus deaths plus immigrants/in-migrants plus emigrants/out-migrants.

UNDERCOUNT. In a census, the omission of valid members of the population in question (See also NET CENSUS UNDERCOUNT ERROR and OVERCOUNT).

UNDER-ENUMERATION. (See UNDERCOUNT).

UNDER-REGISTRATION. The omission of persons or events from a registration system or other administrative record system.

UNINCORPORATED PLACE. (See CENSUS DESIGNATED PLACE)

URBAN FRINGE. The densely settled area surrounding the core city of an urbanized area. Sometimes population referred to as the suburban area (See also SUBURBAN).

URBAN POPULATION. Usually defined as a large population in a densely-packed area that meets criteria derived from geographic, social, and economic factors, which, in turn, may vary by country (See also RURAL POPULATION).

URBANIZATION. Growth in the proportion of persons living in urban areas; the process whereby a society changes from a rural to an urban way of life.

USUAL RESIDENCE. The place where one usually eats and sleeps, a concept associated with a De Jure census (See also CENSUS, CENSUS-DEFINED RESIDENT, DE JURE, DOMICILE, and RESIDENCE).

VACANCY RATE. The ratio of vacant housing units to the total number of housing units, usually multiplied by 100 so that this ratio is expressed as a percent. (See also HOUSING UNIT, NATURAL VACANCY RATE, and OCCUPANCY RATE)

VACANT HOUSING UNIT. A vacant housing unit is one that is not occupied, but intended for occupancy. For vacant units, the criteria of separateness and direct access are applied to the intended occupants whenever possible (See also HOUSING UNIT, HOUSEHOLD, and VACANCY RATE)

VITAL EVENTS. Births, deaths, fetal losses, abortions, marriages, annulments, divorces--any of the events relating to mortality, fertility, marriage, and divorce recorded in registration systems (See also VITAL STATISTICS).

VITAL RATES METHOD. A censal-ratio method of population estimation introduced by D. Bogue in the 1950s that uses crude birth and crude death rates (See also CENSAL-RATIO METHOD and POPULATION ESTIMATE)).

VITAL RECORDS. (See VITAL STATISTICS).

VITAL STATISTICS. Data on births, deaths, fetal losses, abortions, marriages, and divorces usually compiled through registration systems or other administrative record systems (See also VITAL EVENTS).

WEIGHTED AVERAGE. Usually an arithmetic mean of an array of specific rates or ratios, with variable weights applied to them representing the relative distribution of the populations on which the rates or ratios are based. More generally, a summary measure of a set of numbers (absolute numbers or ratios), computed as the cumulative product of the numbers and a set of weights representing their relative importance in the population. An unweighted average is one in which each number in the set has the same weight (e.g., 1 or 1/n, where n is the total set of numbers) (See also CENSAL RATIO METHOD and SYNTHETIC METHOD).

ZIP CODE. Administrative areas set up by the U.S. Postal Service as postal delivery areas and used for marketing and related purposes in the United States. They have fluid boundaries that do not correspond to any established political area or statistical area of the decennial census but may approximate some small areas defined by the census (See also CENSUS GEOGRAPHY).

Endnotes to the Annotated Bibliography and Glossary

*The author thanks Theresa Lowe, Stan Smith and Jeff Tayman for reviewing Appendix B, which incorporates entries from many sources, including article abstracts, Popline, Population Index, and bibliographies compiled by:

(1) Lucky Tedrow and David Swanson (1987), which appears in "Appendix C, Bibliography," pp. 41-52 in E. Byerly in "State and Local Agencies Preparing Population Estimates." *Current Population Reports, P-25, No. 1063*. U. S. Bureau of the Census, September, 1990;

(2) Anne Lee (1995), which is found on pp. 148- 161 in Rives, N.W., W.J. Serow, A. Lee, H. Goldsmith and P. Voss. 1995. *Basic Methods for Preparing Small-Area Population Estimates*. Madison, WI: Applied Population Laboratory, University of Wisconsin;

(3) the Population Association of America (2003), *A Bibliography of Population Estimates* (available online, http://www.popassoc.org/files/public/Pop_Estimates.pdf, last accessed, March, 2008);

(4) Tommy Wright and Joyce Farmer (2000), "A Bibliography of Selected Statistical methods and Development Related to Census 2000," *Statistical Research Report Series, No. No. RR 2000/02*, Washington, DC: U.S. Census Bureau (available online, http://www.census.gov/srd/papers/pdf/rr2000-02.pdf, last accessed May, 2008); and

(5) Henry S. Shryock (1959). "Annotated List of Estimates of Net Migration, 1940 to 1950, by the Residual Method." *Population Index* 25 (1): 16-24.

The section of the annotated bibliography on general population estimates is largely based on the bibliography compiled by L. Tedrow and D. Swanson, which appears in "Appendix C, Bibliography," pp. 41-52 in E. Byerly. "State and Local Agencies Preparing Population Estimates." *Current Population Reports, P-25, No. 1063*. U.S. Bureau of the Census (September, 1990)

Some of the entries in the section on the Housing Unit Method go beyond it, but are included because they contain methods that could be used to develop data for the Housing Unit Method.

Some of the entries in this section on the determination of housing units include methods for estimating households

Some of the entries in the section on vacancy rates deal indirectly with vacancy rate estimation (e.g., they are aimed at estimating the total number of households and housing units).

Some of the entries in the section on the estimation of persons per household deal indirectly with estimating the number of persons per household (e.g., they are aimed at estimating the household population and households)

The Glossary is largely based on the glossary by D. Swanson and G. Stephen found on pp. 751-778 in J. Siegel and D. Swanson (Eds.) *The Methods and Materials of Demography* (2004). New York, NY: Elsevier Academic Press. It excludes many of the entries in the original and includes some entries not found there.