MEASURING UNCERTAINTY IN POPULATION FORECASTS MADE USING THE HAMILTON-PERRY METHOD:

David A. Swanson (Correspondence author)
Department of Sociology and
Center for Sustainable Suburban Development
University of California Riverside
Riverside, California 92521 USA
Tel: 1-951-827-4373 E-mail: **David.swanson@ucr.edu**

Jeff Tayman
Department of Economics
University of California San Diego
9500 Gilman Drive #0508
La Jolla, California 92093-0508 USA
Tel. 1-858-534-7040 Email: jtayman@san.rr.com

Abstract

Two basic approaches have been used to assess population forecast uncertainty: (1) a range of projections based on alternative scenarios; and (2) statistical prediction intervals. In terms of the latter, there are two complementary approaches: (1) model-based intervals; and (2) empirically-based intervals. We evaluate a model-based approach in this paper, but enhance it by using it the information in historical data, a feature found in the empirically-based approach. We describe and test in this paper a regression-based approach for developing 66% forecast intervals for age-group forecasts made using the Hamilton-Perry Method. We use a sample of four states (one from each of the four US Census Regions) with nine ex post facto tests, one for each census from 1930 to 2010, which yields 576 observations. The four states and the nine test points provide a wide range of characteristics in regard to population size, growth, and age-composition, factors that affect forecast accuracy. The tests reveal that the 66% intervals contain the census age-groups in 397 of the 576 observations (68.9 percent). We discuss the results, and make some observations regarding the limitations of our study. We conclude that the results are encouraging, however, and offer suggestions for further work.

Introduction[1]

Although they are widely used, population forecasts entail a tremendous amount of uncertainty, especially for long time horizons and for places with small or rapidly changing populations (Alho, 1984; Alho and Spencer, 1984, 1985, 1990, 1997, 2005; Lutz, Sanderson and Scherbov, 1999; Smith, Tayman and Swanson, 2001; Tayman, Smith, and Lin, 2007; Tayman, Smith and Rayer, 2011; Wilson, 2012). As such, virtually every forecast is wrong, making the task of an accurate forecast impossible, but the task is unavoidable (Keyfitz, 1987: 236). It is impossible in that the forecasted numbers turn out to be different from what actually occurs, but unavoidable in that forecasts must be done in the modern world. Swanson and Tayman (1995) describe this irony as the "rock" and the "hard place." As observed by Swanson and Tayman (1995), demographers have developed several strategies for dealing with the "irony" of forecasting. They include the use of the term "projection" rather than "forecast," (Keyfitz 1972, Pittenger 1978, Smith and Bayya 1992, Smith, Tayman, and Swanson, 2001), "normative" forecasting (Moen 1984), and providing measures of forecast uncertainty. One way to assess uncertainty is to produce several alternative projections or scenarios based on different sets of assumptions (Campbell, 1996; Cheeseman-Day, 1992; Spencer, 1989; Smith, Tayman, and Swanson, 2001; Thompson and Whelpton, 1933; U. S. Census Bureau, 1966). Another approach is to develop statistical forecast intervals based on historical data (Alho and Spencer, 2005; Stoto, 1983; Swanson and Beck, 1994). It is the latter that we explore in this paper.

Forecast intervals based on statistical theory and data on error distributions provide an explicit estimate of the probability that a given range will contain the future population. These intervals are sometimes called prediction intervals, probability intervals, confidence intervals, or confidence limits. We call them forecast intervals to distinguish them from traditional confidence intervals, which—strictly speaking—apply only to sample data.

Two types of forecast intervals have been used most frequently for population forecasts. One is based on the development of statistical models of population growth and the other is based on empirical analyses of errors from past population projections. Both rely on the assumption that historical or simulated error distributions can be used to predict future error distributions. To a large extent, the two approaches complement one another, but neither is fully satisfactory. On the one hand, model-based intervals exploit the theories and underlying inferential statistics, but fall short in utilizing the information available in historical data. On the other hand, empirically-based intervals utilize the information from historical data, but fall short in exploiting the theories underlying inferential statistics. We evaluate a model-based approach in this paper, but enhance it by using it the information in historical data

After discussing the model-based and empirically-based approaches, we discuss our model-based approach, which employs simple regression models applied to a forecasting method known as the Hamilton-Perry Method (Hamilton and Perry, 1962). This is followed by an empirical analysis and a discussion.

Model-Based Intervals

Model-based prediction intervals capitalize on the stochastic (or random) nature of population processes. They can be developed in a number of ways. Past applications have included maximum likelihood estimators of population growth rates (Cohen 1986), Monte Carlo simulations of fertility and migration rates (Pflaumer 1988), regression-based projection models (Swanson and Beck 1994), Bayesian projection models (Alkema et al. 2011), models based on the opinions of a group of experts (Lutz, Sanderson, and Scherbov 1999; San Diego County Water Authority 2002), and time series models covering mortality rates (Lee and Carter 1992), life expectancy (Torri and Vaupel 2012), fertility rates (Lee, 1993), net migration (De Beer 1993), and total population size (Alders, Keilman, and Cruijsen 2007; Hyndman and Booth 2008). Although much of the research on model-based intervals has focused on national or regional projections, recent research has extended the analysis to subnational projections as well (Cameron and Poot 2011; Tayman et al. 2007; Wilson and Bell 2004).

Time series models (especially ARIMA models) are the models most commonly used for developing prediction intervals for population projections. These models assume that the pattern (structure) of the data does not change over time, that errors are normally distributed with a mean of zero and a constant variance, and that errors are independent of each other (Makridakis et al., 1987). Time series models require a fairly long series of historical observations and can be difficult to apply, especially when attempting to combine prediction intervals for all three components of growth and developing intervals for various subgroups of the population.

Providing a detailed description of model-based prediction intervals is beyond the scope of this paper, but we can give several examples of the intervals produced by these models and compare them to the high and low projection series produced using the traditional approach. Lee and Tuljapurkar (1994) projected a population of 398 million for the United States in 2065, with a 95% prediction interval of 259-609 million. This range is considerably wider than the spread between the low and high projections produced by the Census Bureau at about the same time; those projections ranged from 276-507 million in 2050, with a medium projection of 383 million (Cheeseman-Day, 1992). The previous set of Census Bureau projections reported much lower numbers and a slightly smaller range, with a medium projection of 300 million and a range of 230-414 million for 2050 (Spencer 1989).

Pflaumer (1992) made two time series projections of the U.S. population, one based on population size and the other based on the natural logarithm of population size. The first model produced a medium projection of 402 million in 2050, with a 95% prediction interval of 277-527 million. These numbers are similar to the Census Bureau's projections from the same time. The second model produced a medium projection of 557 million, with a 95% prediction interval of 465-666 million. These numbers are much higher and provide a narrower range than the Census Bureau's projections.

McNown et al. (1995) made time series projections of the components of growth for the U.S. population, as well as total population size. For 2050, they projected a total population of 373 million, with a 95% prediction interval ranging from 243 million to 736 million. The total fertility rate was projected to be 2.46 in 2050, with a 95% prediction interval ranging from 0.91 to 5.53. Life expectancy at birth for males was projected to be 75.5, with a 95% prediction interval ranging from 68.5 to 82.8.

For fertility these intervals are much larger than those found in the Census Bureau projections, which assumed that the total fertility rate would range only from 1.83 to 2.52 in 2050 (Cheeseman-Day, 1992). For mortality the intervals are not much different than those reported by the Census Bureau, in which life expectancy at birth was projected to range between 75.3 and 87.6 in 2050.

Swanson and Beck (1994) developed a regression-based model for making short-term county population projections in the state of Washington. They compared the 2/3 prediction intervals associated with this model to census counts of Washington's 39 counties in 1970, 1980, and 1990. They found the prediction intervals to contain the 1970 census count in 30 counties (77%), the 1980 census count in 24 counties (62%), and the 1990 census count in 31 counties (79%). These results suggest that Swanson and Beck's 2/3 prediction intervals provided a reasonably accurate view of forecast uncertainty.

Model-based prediction intervals are valid only to the extent that the assumptions underlying the models are valid. In spite of their objective appearance, they are strongly influenced by the analyst's judgment. The models themselves are often complex and require a substantial amount of base data. They are subject to errors in the base data, errors in specifying the model, errors in estimating the model's parameters, and future structural changes invalidating the model's parameter estimates (Lee 1992). In addition, it is the case that many alternative forecasting models can be specified, each providing different (perhaps dramatically different) prediction intervals (Cohen, 1986; Lee, 1974; Tayman, Smith and Lin, 2007).

In spite of these problems, model-based prediction intervals offer one important benefit: they provide explicit probability statements to accompany point forecasts. The intervals are often wide, exceeding the low and high projections produced by official statistical agencies. Given that many data users (and producers) tend to overestimate the accuracy of population projections, model-based prediction intervals provide an important reality check.

Empirically-Based Intervals

The second type of prediction interval is based on empirical analyses of errors from past projections (Keyfitz, 1981; Smith and Sincich, 1988; Stoto, 1983; Smith and Rayer, 2012; Tayman et al., 1998). Keyfitz (1981) took some 1,100 national projections made between 1939 and 1968 and, for each one, calculated the difference between the projected annual growth rate and the rate actually occurring over time. He found this difference to be largely independent of the length of horizon over which the projections were made. He calculated the RMSE for the entire sample to be approximately 0.4 percentage points and developed 2/3 prediction intervals by applying that error to the growth rates projected for each country. For example, if a country were projected to grow by 2% per year for the next 20 years, the probability would be approximately 2/3 that the actual growth rate would be somewhere between 1.6% and 2.4%.

Keyfitz (1981) refined his analysis by separating countries according to their population growth rates, finding a RMSE of 0.60 for rapidly growing countries, 0.48 for moderately growing countries, and 0.29 for slowly growing countries. He illustrated this refinement by applying the 0.29 RMSE to the U.S. growth rate of 0.79% per year projected by the Census Bureau, yielding annual growth rates of 0.50% and 1.08%. Applying those growth rates to the 1980 population of 260 million produced a range of 245-275 million in 2000. He concluded that the odds were about 2 to 1 that this range would contain the U.S. population in that year.

Stoto (1983) followed a similar approach, but analyzed projections containing more temporal and geographic diversity. Like Keyfitz (1981), he calculated forecast error as the difference between the projected annual growth rate and the rate actually realized over time. He differentiated between two components of error, one related to the launch year of the projection and the other to seemingly random events (the residual). For more developed countries he found the launch-year component to have a distribution that was stable over time and centered on zero, implying that the projections were unbiased. For less developed countries he found the variance of the launch-year component to be stable but that earlier sets of projections had a strong downward bias (although recent sets had little bias). The second component (the residual) was found to have a stable distribution but to have occasional outliers. For both components the variance was larger for less developed countries than more developed countries.

Stoto (1983) calculated the standard deviations for these two components of error and constructed prediction intervals in a manner similar to that used by Keyfitz (1981). He applied those intervals to projections of the U.S. population and estimated that there was about a 2/3 probability that an interval of 241-280 million would contain the actual population in 2000, and a 95% probability that an interval of 224-302 million would contain that population. He compared his results to projections produced by the Census Bureau, concluding that the Census Bureau's low and high series were very similar to a 2/3 prediction interval. Keyfitz (1981) had reached the same conclusion.

Smith and Sincich (1988) also used the distribution of past forecast errors to construct prediction intervals, but followed a different approach. They modified a technique developed by Williams and Goodman (1971), in which the predicted distribution of future forecast errors was based directly on the distribution of past forecast errors. An important characteristic of this technique is that it can accommodate any error distribution, including the asymmetric and truncated distributions typically found for absolute percent errors.

Using population data for states from 1900 to 1980, Smith and Sincich used four simple extrapolation methods to make a series of projections covering 10- and 20-year horizons. They calculated absolute percent errors for each target year by comparing projections with census counts, focusing on the 90% intervals for each set of projections (i.e., the absolute percent error larger than exactly 90% of all absolute percent errors). They investigated two approaches to constructing 90% prediction intervals, one using the 90% interval from the previous set of projections and the other using the 90% interval from all other sets of projections. They found both approaches to provide relatively accurate prediction intervals. For most individual target years, 88%-94% of state forecast errors fell within the predicted 90% interval. Summing over all target years, 91% of all forecast errors fell within the predicted 90% interval. They concluded that stability in the distribution of absolute percent errors over time made it possible to construct useful prediction intervals for state projections.

Smith and Rayer (2012) used the Williams and Goodman approach to construct and test prediction intervals for county projections in Florida. Using forecast errors for target years 1985, 1990, and 1995, they constructed 2/3 prediction intervals for projections with launch years 1995, 2000, and 2005 and counted the number of counties in which the subsequent population counts or estimates fell within the predicted intervals. They found that 43 counties (64%) fell within the predicted range for 5-year horizons and 49 counties (73%) for both 10- and 15-year horizons. These numbers were fairly close to the 45 counties implied by the prediction intervals. Given the year-to-year volatility of Florida's population growth, this reflects a reasonably good forecasting performance.

Tayman et al. (1998) developed statistically-based prediction intervals for subcounty population forecasts in San Diego County. They started by projecting the population residing in grid cells, which are geographic areas of 2,000 ft. by 2,000 ft. defined for the most densely populated parts of the county. The projections had 1980 as a launch year and 1990 as a target year. Using repeated sampling techniques and randomly selected grid cells, they developed projections for a large number of areas varying in size from 500 to 50,000. Forecast errors were calculated by comparing the 1990 projections with 1990 census counts.

Rather than constructing prediction intervals for the population forecasts per se, Tayman and his colleagues developed predictions for the mean errors implied by those forecasts. Empirical prediction intervals for MAPEs and MALPEs were developed using an approach similar to that used by Williams and Goodman (1971) and Smith and Sincich (1988). For areas with 500 persons, they found a 95% prediction interval of 67.4%-80.3% for the MAPE. For areas with 50,000 or more, the interval was 9.7%-11.5%. For MALPE, the intervals were wider but centered closer to zero.

The Hamilton-Perry Method

Before describing the Hamilton-Perry method, it is useful to recall that any quantitative approach to forecasting is constrained to satisfy various mathematical identities (Land 1986). In regard to population forecasting, an approach should ideally satisfy demographic accounting identities, which is summarized in the identity known as the fundamental demographic equation:

$$P_t = P_0 + \text{Births} - \text{Deaths} + \text{Inmigrants} – \text{Outmigrants} \quad [1]$$

That is, the population at some time in the future, P_t, must be equal to the population at an earlier time, P_0, plus the births and in-migrants and less the deaths and out-migrants that occur between time =0 and time=t. The most commonly used approach to population forecasting, cohort-component method, satisfies the fundamental equation, but it is data-intensive (George et al. 2004, Smith Tayman and Swanson, 2001; Murdock and Ellis 1991, Pittenger 1976)

As we show at the end of this section, the Hamilton-Perry Method also satisfies the fundamental demographic equation. However, it has far less intensive input data requirements than does the cohort-component method (Hamilton and Perry 1962, Swanson Schlottmann and Schmidt, 2010, Swanson and Tedrow, 2012). Instead of mortality, fertility, migration, and total population data, which are required by the full-blown cohort-component method, the Hamilton-Perry method requires data only from the two most recent censuses (Smith, Tayman, and Swanson 2001: 153-158, Swanson Schlottmann and Schmidt 2010, Swanson and Tedrow 2012).

The Hamilton-Perry method moves a population by age (and sex) from time t to time t+k using cohort-change ratios (CCR) computed from data in the two most recent censuses. It consists of two steps. The first uses existing data to develop CCRs and the second applies the CCRs to the cohorts of the launch year population to move them into the future. As shown by Swanson, Schlottmann, and Schmidt (2010), the formula for the first step, the development of a CCR is:

$${}_nCCR_{x,i,t} = ({}_nP_{x,i,t}) / ({}_nP_{x-k,i,t-k}) \quad [2]$$

where

($_{n}P_{x,i,t}$) is the population aged x to x+n in area i at the most recent census (t),

($_{n}P_{x-k,i,t-k}$) is the population aged x-k to x-k+n in area i at the 2nd most recent census (t-k),

k is the number of years between the most recent census at time t for area i and the one preceding it for area i at time t-k.

The basic formula for the second step, moving the cohorts of a population into the future is:

$$_{n}P_{x+k,i,t+k} = (_{n}CCR_{x,i,t})*(_{n}P_{x,i,t}) \qquad [3]$$

where

$_{n}P_{x+k,i,t+k}$ is the population aged x+k to x+k+n in area i at time t+k

and both ($_{n}CCR_{x,i,t}$) and ($_{n}P_{x,i,t}$) are as defined in equation [2]

Given the nature of the CCRs, 10-14 is the youngest five-year age group for which projections can be made if there are 10 years between censuses. To project the population aged 0-4 and 5-9 one can use the Child Woman Ratio (CWR) or more generally an “Adult Child Ratio” (ACR). It does not require any data beyond what is available in the decennial census. For projecting the population aged 0-4, the ACR is defined as the population aged 0-4 divided by the population aged 20-34. For projecting the population aged 5-9, the ACR is defined as the population aged 5-9 divided by the population aged 25-39. Here are the ACR equations for projecting the population aged 0-4 and 5-9, respectively.

Population 0-4: ${}_5P_{0,t+k} = ({}_5P_{0,t} / {}_{15}P_{20,t}) * ({}_{15}P_{20,t+k})$ [4a]

Population 5-9: ${}_5P_{5,t+k} = ({}_5P_{5,t} / {}_{15}P_{25,t}) * ({}_{15}P_{25,t+k})$ [4b]

where

P is the population,

t is the year of the most recent census

and t+k is the projection year

Another way to project the youngest age groups is to take ratios of them at two points in time and apply that ratio to the launch year age group (t). In the first step, the ratios are as follows

Population 0-4: ${}_5R_{0,t} = ({}_5P_{0,t} / {}_5P_{0,t-k})$ [5a]

Population 5-9: ${}_5R_{5,t} = ({}_5P_{5,t} / {}_5P_{5,t-k})$ [5b]

In the second step, the projected population at t+k is found as follows.

Population 0-4: ${}_5P_{0,t+k} = ({}_5P_{0,t}) * {}_5R_{0,t}$ [6a]

$$\text{Population 5-9: } {}_{5}P_{5,t+k} = ({}_{5}P_{5,t}) * {}_{5}R_{5,t} \qquad [6b]$$

Projections of the oldest open-ended age group also differ slightly from the CCR projections for the age groups beyond age 10 up to the oldest open-ended age group. If, for example, the final closed age group is 70-74, with 75+ as the terminal open-ended age group, then calculations for the $CCR_{i,x+}$ require the summation of the three oldest age groups to get the population age 65+ at time t-k:

$${}_{\infty}CCR_{75,i,t} = {}_{\infty}P_{75,i,t} / {}_{\infty}P_{65,i,t-k} \qquad [7a]$$

The formula for estimating the population 85+ of area i for the year t+k is:

$${}_{\infty}P_{75+,i,t+k} = ({}_{\infty}CCR_{75,i,t})*({}_{\infty}P_{65,i,t}) \qquad [7b]$$

In terms of the Hamilton-Perry Method satisfying the fundamental demographic equation, we show that the former (equation [2]) can be re-stated using the latter (equation [1]) as follows.

Since $P_{i,t+k} = P_{i,t} + B_i – D_i + I_i - O_i$

where

$P_{i,t}$ = Population of area i at time t (the launch date)

$P_{i,t+k}$ = Population of area i at time t+k (the estimate date)

B_i = Births in area i between time t and t+k

D_i = Deaths in area i between time t and t+k

I_i = In-migrants in area i between time t and t+k

O_i = Out-migrants in area i between time t and t+k

Then $_{n}CCR_{x,i,t} = (_{n}P_{x-k,i,t-k} + B_i – D_i + I_i - O_i)/(_{n}P_{x-k,i,t-k})$ [8]

And we since can express equation [3] also in terms of equation [1]

$_{n}P_{x+k,i,t+k} = (_{n}P_{x-k,i,t-k} + B_i – D_i + I_i - O_i)/(_{n}P_{x-k,i,t-k}))*(_{n}P_{x,i,t})$ [9]

and where x+k >= 10 then

$_{n}CCR_{x,i,t} =(_{n}P_{x-k,i,t-k} – D_i + I_i - O_i)/(_{n}P_{x-k,i,t-k})$

and since $N_i = I_i - O_i$, we have, where $x+k \geq 10$,

$_{n}CCR_{x,i,t} =(_{n}P_{x-k,i,t-k} – D_i + N_i)/(_{n}P_{x-k,i,t-k})$ [10]

Equations [8], [9] and [10] show that the Hamilton-Perry Method is not only consistent with the fundamental demographic equation, but also closely related to the cohort-component method. The Hamilton-Perry Method simply expresses the individual components of change - births, deaths, and migration - in terms of Cohort Change Ratios. As such, it satisfies the fundamental demographic equation. As we will see in the following section, this way of expressing the components of population change can be exploited. An important reason for a demographic forecasting method to be consistent with the fundamental demographic equation is to minimize the potential errors associated with hidden heterogeneity (Vaupel and Yaushin 1985).

Developing Forecast Intervals for the Hamilton-Perry Method: The Regression Approach to Estimating CCRs

As should be clear from the preceding discussion, the Hamilton-Perry Method is deterministic. This is not surprising given its consistency with the fundamental demographic equation, which by its nature is an accounting method. However, we also know that population forecasting is subject to uncertainty since we do not precisely know the future components making up the fundamental equation. So, the question is how to introduce an element of statistical uncertainty into a method that is inherently deterministic. One answer to this question is found by employing a simple regression method to estimate CCRs and then applying the regression-estimated CCRs to the launch-year age groups to obtain forecasts of these age groups.

Recall from equation [2] that ${}_nCCR_{x,i,t} = ({}_nP_{x,i,t}) / ({}_nP_{x-k,i,t-k})$. From this, we can define the CCR for the preceding census period as ${}_nCCR_{x,i,t-k} = ({}_nP_{x,i,t-k}) / ({}_nP_{x-k,i,t-2k})$. We can then construct a regression model with ${}_nCCR_{x,i,t}$ as the dependent variable and ${}_nCCR_{x,i,t-k}$ as the independent variable. We note that for age groups 0-4, 5-9, and the terminal open-ended age group that the dependent and independent observations follow the equations provided earlier. Given this adjustment, we can generally describe the estimated CCRs at time t as follows.

$$ {}_{n}\hat{CCR}_{x,i,t} = a + b^{*}({}_{n}CCR_{x,i,t-k}) \quad [11] $$

We can then apply the regression-estimated CCR for time t to forecast the CCR at time t+k and utilize the measure of statistical uncertainty (the standard error of estimate) for the model along with the sample size and other characteristics of the data in order to generate forecast intervals around the forecasted CCRs at time t+k. These intervals can then be translated directly to the actual population numbers forecasted for the age groups in question (Swanson and Beck, 1994). The forecast intervals generated were based on equation 4.2 found in chapter 4, Hyndman and Athanasopoulos (2012)

Empirical Evaluation

To empirically examine our regression-based method for developing forecast intervals around population forecasts by age generated using the Hamilton-Perry Method, we selected a sample made up of one state from each of the four Census Regions in the United States. The states selected are Georgia (the South Region), Minnesota (the Midwest Region), New Jersey (The Northeast Region) and Washington (The West Region). We then assembled census data for these four states for each census year from 1900 to 2010 (These data are available in the appendix). The data provide nine points in time at which the forecast intervals can be evaluated, 1930, 1940, 1950, 1960, 1970, 1980, 1990, 2000, and 2010. This sample provides a wide range of demographic information in terms of variation in population size, age-composition, and rates of change. Table 1 provides an overview of this range by displaying the population of each of the four states in 1900 and in 2010 and decennial rates of population change from 1900 to 2010.

(Table 1 about here)

Although we do not show a summary of the changes in age composition by state and census year, they are extensive as can be seen the appendix tables, which provide the age data by state and census year.

We proceed by constructing CCRs over two successive decennial periods (e.g., 1910-1920/1900-1910) over the entire period, using regression to estimate the CCR in the numerator from the CCR in the denominator. We then use the regression-based estimate of the CCR of the “current period “ (e.g., 1910-1920) to forecast the CCRs to the next period, the “launch year” (e.g., 1920-1930) and develop forecast intervals around the forecasted CCRs, which are then translated into the forecasted age groups for the “target year” (e.g., 1930). The forecast intervals are then examined to see if they contain the census age groups for the target year.

Because of the way data for the terminal open-ended age group are reported differently over the period for which we assemble census data, we used “75 years and over” for the entire period since it was the common denominator. This means that there are 16 age groups (0-4, 5-9, 70-74, and 75+) used in the empirical evaluation.

In terms of the width of the forecast interval we selected 66 percent because of prior research indicating that “low” and “high” scenarios constructed for the cohort-component method corresponded empirically to 66% confidence intervals (Stoto, 1983) as well as findings by Swanson and Beck (1994). Table 2 provides a summary of the results for all four states at each of the nine census test points. The table shows the number of times (out of 16) that the 66% forecast interval contained the corresponding census number for a given age group.

(Table 2 about here).

In examining the state of Georgia (South Census Region), we find that its population increased by almost five-fold between 1900 and 2010. In 1900 it had the largest population of any of the four sample states it retained that position in 2010. Its annual average growth rates (by decade) ranged from 0.0005 between 1920 and 1930 to .0168 between 2000 and 2010. Changes in its age composition are extensive (See Appendix Table 1) , with big impacts associated with the great depression, World War II, the baby boom, and immigration to the sunbelt states more recently. The 66 percent forecast intervals contain their corresponding age groups 76 times out of 144 observations, or 52.8 percent. Overall, Georgia has the lowest percent of census age groups within the 66 percent forecast interval.

The population of Minnesota tripled from 1900 to 2010. Its average annual growth rates ranged from a low of 0.0066 between 1940 and 1950 to a high of 0.0170 between 1900 and 1910, a period when the state was still receiving a large number of immigrants from Europe. As is the case for Georgia, changes in its age composition are extensive (See Appendix Table 2) , with big impacts associated with the restrictions placed on immigration by World War I and by US laws in the early 1920s, the great depression, World War II, the baby boom, and outmigration to sunbelt states in more recent decades. The 66 percent forecast intervals contain their corresponding age groups 113 times out of 144 observations, or 78.5 percent. Overall, Minnesota has the highest percent of census age groups within the 66 percent forecast interval.

For New Jersey, we see that its population grew from 1,879,890 in 1900 to 8,791,894 in 2010. New Jersey had the second highest population in 1900 and again in 2010. Its average annual growth rates ranged from a low of 0.0027 between 1970 and 1980 to a high of 0.0299 between 1900 and 1910. As is the case for Georgia and Minnesota, changes in its age composition are extensive (See Appendix Table 3) , with big impacts associated with the restrictions placed on immigration by World War I and by US laws in the early 1920s, the great depression, World War II, the baby boom, and outmigration to sunbelt states in more recent decades. The 66 percent forecast intervals contain their corresponding census age groups 106 times out of 144 observations, or 73.6 percent. Overall, New Jersey has the second-highest percent of census age groups within the 66 percent forecast interval.

In 1900, Washington was largely a frontier state. It had the smallest population (511,844) of any of the four states in the sample. However, by 2010, it had grown to 6,724,540 which surpassed the population of Minnesota in 2010. Its annual rates of population change are somewhat more dramatic than the other states between the 1900-1910 period and the 2000-2010 period. Between 1900 and 1910 it posted an annual rate of 0.0797, the highest of any of the decennial growth rates in the sample. It also posted the second highest rate. Between 1940 and 1950 the state grew at an annual rate of 0.0314. The lowest rate of annual population change (0.0106) is found between 1930 and 1940. The 66 percent forecast intervals contain their corresponding census age groups 102 times, which represents 70.8 percent of the 144 observations.

Discussion

Overall, the 66 percent intervals contain their corresponding census age groups in 397 cases. This represents 68.9 percent of the 576 total observations. In terms of the nine census test points, the overall results show that in five of them (1960, 1970, 1990, 2000, and 2010) the forecast intervals contain the census age groups more than 66 percent of the time. In two test points (1930 and 1980), the intervals contain the census age groups more than 60 percent of the time. In the remaining two test points, 1940 and 1950, the intervals contain the census age 48 percent of the time. We note that the 1940 point encompasses the economic boom experienced in the 1920s and the economic depression during the 1930s and the large scale "baby bust" associated with it. The 1950 point encompasses the depression and baby bust period of the 1930s and the economic recovery stimulated by World War II and the initial part of the large scale "baby boom " from 1946 to 1950.

Although they are not shown here, the average width of the forecast intervals appears to us to be reasonable at the 66 percent level in that they are neither so wide as to be meaningless nor too narrow to be overly-restrictive. This is largely consistent with prior work by Swanson and Beck (1994) on confidence intervals derived from regression-based forecasts, which suggested that 66 percent confidence intervals were not only feasible, but also neither too wide to be meaningless nor to narrow as to be overly-restrictive. Also consistent with the work by Swanson and Beck, 1994), the fact the regression-based forecast intervals contain the actual numbers by age in 69 percent of the 576 observations provide further support that 66 percent forecast intervals based on the regression-estimated CCR approach is both useful and feasible. We find these results encouraging.

It is important to note that the forecast intervals are technically generated only for age- specific groups in this paper. Thus, they are not shown for summations of the low and high boundaries for age-specific groups, that is, for the total population. We do note, however, that Espenshade and Tayman (1982) describe a method whereby the forecast intervals for subgroups can be applied to the entire population.

At this point, we do not suggest using this method beyond a ten-year forecast horizon. This is consistent with observations about the use of the Hamilton-Perry method in general (Smith, Tayman, and Swanson, 2001; Swanson, Schlottmann and Schmidt, 2010) and, as such is not a major limitation. We also suggest that this approach to developing uncertainty measures be used with care when applied to small populations, such as those found at the county and sub-county levels. While our sample provides a wide range of demographic behavior in terms of size, age composition, and population changes, it is after all a sample of states, which means that some of the extreme conditions found at sub-state levels are not present (see, e.g., Swanson, Schlottmann, and Schmidt, 2010). We suggest that further research using this approach would be useful by examining both longer forecast horizons and smaller populations (i.e., the sub-state populations). Another area for further research would be to utilize Keyfitz's approach using Root Mean Square Errors in conjunction with the Hamilton-Perry Method.

Endnote

1. The appendix contains four tables, one for each state, that provide the input data used in this study. Space considerations prevent us all of the regression and evaluation results here. However, these results can be reconstructed from the tables in the appendix and, in addition, the authors will be pleased to provide them upon request. The authors are very grateful to anonymous reviewers who made valuable comments and suggestions on an earlier and quite different version of this paper. Their comments led us to considering how to develop a more simple, yet useful system for measuring uncertainty in a forecast method that is itself simple and useful.

References

Alders, M., N. Keilman and H. Cruijsen, H. (2007). Assumptions for long-term stochastic population forecasts in 18 European countries. *European Journal of Population* 23: 33-69.

Alho, J. (1984). Probabilistic Forecasts: The Case of Population Projections. *Scandinavian Housing and Planning Research* 1: 99-105.

Alho, J. and B. Spencer. (1985). Uncertain Population Forecasting. *Journal of the American Statistical Association* 80 (390): 306-314.

Alho, J. and B. Spencer. (1990). Error Models for Official Mortality Forecasts. *Journal of the American Statistical Associatio*n 85 (411): 609-616.

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

Alho, J. and B. Spencer. (2005). *Statistical Demography and Forecasting*. Dordrecht, The Netherlands: Springer.

Alkema, L., A. E. Raftery, P. Gerland, S. J. Clark, F. Pelletier, T. Buettner and G. K. Heilig. (2011). Probabilistic projections of the total fertility rate for all countries. *Demography* 48: 815-839.

Cameron, M.P. and J. Poot. (2011). Lessons from stochastic small-area population projections: The case of Waikato subregions in New Zealand. *Journal of Population Research* 28: 245-265.

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

Cheeseman- Day, J. (1992). *Population Projections of the United States, By Age, Sex, Race, and Hispanic Origin: 1992-2050*. Current Population Reports P-25, No. 1092. Washington, D.C.: U.S. Bureau of the Census.

Cochran, W. (1977). *Sampling Techniques, 3rd Edition*. Wiley: New York.

Cohen, J. (1986). Population Forecasts and Confidence Intervals for Sweden: A Comparison of Model-Based and Empirical Approaches. *Demography* 23 (1): 105-123.

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

Espenshade, T. and J. Tayman. (1982). Confidence Intervals for Postcensal Population Estimates. *Demography* 19 (2): 191-210.

George, M. V., S. Smith, D. Swanson, and J. Tayman. (2004). Population Projections. pp. 561-601 in J. Siegel and D. Swanson (eds.) *The Methods and Materials of Demography 2nd Edition*. New York, NY: Elsevier Academic Press.

Hamilton, C. H. and J. Perry. (1962). A Short Method for Projecting Population by age from one Decennial Census to Another. Social Forces, 41: 163-170.

Hyndman, R. J., and G. Athanasopoulos (2012) *Forecasting: Principles and Practice* (online at http://otexts.com/fpp/)

Hyndman, R.J. and H. Booth. (2008). Stochastic population forecasts using functional data models for mortality, fertility and migration. *International Journal of Forecasting* 24: 323-342.

Keyfitz, N. (1972). On Future Population. *Journal of the American Statistical Association* 67: 347-363.

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

Keyfitz, N. (1987). The Social and Political Context of Population Forecasting. pp. 235-258 in W. Alonso and P. Starr (Eds.) *The Politics of Numbers*. New York, NY: Russell Sage Foundation.

Land, K. (1986). Methods for National Population Forecasts: A Review. *Journal of the American Statistical Association* 81: 888-901.

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

Lee, R. (1992). Stochastic demographic forecasting. *International Journal of Forecasting* 8: 315-327.

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

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

Lee, R. and L. S. Tuljapurkar. (1994). Stochastic Population Projections for the United States: Beyond High, Medium, and Low. *Journal of the American Statistical Association* 89: 1175-1189.

Lutz, W., Sanderson, W., and Scherbov, S. (1999). Expert-based probabilistic population projections. pp. 139-155 in W. Lutz, J. Vaupel, and D. Ahlburg (Eds.) *Frontiers of Population Forecasting*. New York, NY: The Population Council. (A supplement to *Population and Development Review*, 24).

Makridakis, S., M. Hibon, E. Lusk, and M. Belhadjali. (1987). Confidence intervals: An empirical investigation of the series in the M-competition. *International Journal of Forecasting* 3: 489-508.

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

Moen, E. (1984). Voodoo Forecasting: Technical, Political, and Ethical Issues Regarding the Projection of Local Population Growth. *Population Research and Policy Review* 3: 1-25.

Murdock, S. and D. Ellis. (1991). *Applied Demography: An Introduction to Basic Concepts, Methods, and Data*. Denver, Colorado: Westview Press.

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

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

Pittenger, D. (1976). *Projecting State and Local Populations*. Cambridge, MA: Ballinger.

Pittenger, D. (1978). The Role of Judgment, Assumptions, Techniques and Confidence Limits in Forecasting Population. *Socio-Economic Planning Sciences* 12: 271-276.

San Diego County Water Authority. 2002. *Regional water facilities master plan. Appendix C: Development of probabilistic water demand forecasts*. San Diego, CA: San Diego County Water Authority.

Smith, S. and R. Bayya. (1992). An Evaluation of Population Forecast Errors for Florida and Its Counties. *Applied Demography* 7: 1-5.

Smith, S.K. and S. Rayer. (2012). An evaluation of population forecast errors for Florida and its counties, 1980-2010. Paper presented at the Conference on Applied Demography, San Antonio, TX.

Smith, S. and T. Sincich. (1988). Stability over Time in the Distribution of Population Forecast Errors. *Demography* 25: 461-474.

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

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

Stoto, M. (1983). The Accuracy of Population Projections. *Journal of the American Statistical Association* 78: 13-20.

Swanson, D. and D. Beck. (1994). A New, Short-Term County Population Projection Method. *Journal of Economic and Social Measurement* 20: 1-26.

Swanson, D. and J. Tayman. (1995). Between A Rock And A Hard Place: The Evaluation of Demographic Forecasts. *Population Research and Policy Review* 14: 233-249.

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

Swanson, D., A. Schlottmann, and R. Schmidt. (2010). Forecasting the Population of Census Tracts by Age and Sex: An Example of the Hamilton–Perry Method in Action. *Population Research and Policy Review* 29: 47-63.

Tayman, J. E. Schaefer, 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.

Tayman, J., S. Smith, and J. Lin. (2007). Precision, Bias, and Uncertainty for State Population Forecasts: An Exploratory Analysis of Time Series Models. *Population Research and Policy Review* 26 (3): 347-369.

Tayman, J., Smith, S.K., and Rayer, S. (2011). Evaluating population forecast accuracy: A regression approach using county data. *Population Research and Policy Review* 30: 235-262.

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

Torri, T. and J. W. Vaupel. (2012). Forecasting life expectancy in an international context. *International Journal of Forecasting* 28: 519-531.

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

U.S. Census Bureau. (1982). *1980 Census of Population, General Population Characteristics,* (by State). Washington, DC: Government Printing Office.

U. S. Census Bureau. (1992). *1990 Census of Population, General Population Characteristics* (by State). Washington, DC: Government Printing Office.

U.S. Census Bureau (1973). *1970 Census of Population, General Population Characteristics* (by State). Washington, D.C., Government Printing Office

Vaupel, J. and A. Yashin. (1985). Heterogeneity's Ruses: Some Surprising Effects of Selection on Population Dynamics. *The American Statistician* 39 (August): 176-185.

Williams, W., and M. Goodman. (1971). A simple method for the construction of empirical confidence limits for economic forecasts. *Journal of the American Statistical Association* 66: 752-754.

Wilson, T. (2012). Forecast Accuracy and Uncertainty of Australian Bureau of Statistics State and Territory Population Projections. *International Journal of Population Research*. Article ID 419824, 16 pages, 2012. doi:10.1155/2012/419824

Wilson, T. and M. Bell. (2004). Probabilistic regional population forecasts: The example of Queensland, Australia. *Geographical Analysis* 39: 1-25.

Table 1. The Total Population of each State in 1900 and 2010 and Annual Rates of Change from 1900 to 2010 by Decade*					
	STATE				
Census Year	GEORGIA	MINNESOTA	NEW JERSEY	WASHINGTON	
1900	2,209,974	1,747,292	1,879,890	511,844	Population
1910	0.0164	0.0170	0.0299	0.0797	Average annual Rate of Change over the Decade
1920	0.0105	0.0141	0.0219	0.0175	Average annual Rate of Change over the Decade
1930	0.0005	0.0072	0.0247	0.0144	Average annual Rate of Change over the Decade
1940	0.0072	0.0086	0.0030	0.0106	Average annual Rate of Change over the Decade
1950	0.0098	0.0066	0.0150	0.0314	Average annual Rate of Change over the Decade
1960	0.0135	0.0135	0.0227	0.0183	Average annual Rate of Change over the Decade
1970	0.0152	0.0108	0.0167	0.0178	Average annual Rate of Change over the Decade
1980	0.0174	0.0069	0.0027	0.0192	Average annual Rate of Change over the Decade
1990	0.0170	0.0071	0.0048	0.0164	Average annual Rate of Change over the Decade
2000	0.0234	0.0117	0.0085	0.0192	Average annual Rate of Change over the Decade
2010	0.0168	0.0075	0.0044	0.0132	Average annual Rate of Change over the Decade
2010	9,687,653	5,303,925	8,791,894	6,724,540	Population
* The 1900 population totals exclude those for whom age was not reported.					

Table 2. Summary Results of the Regression-based 66% forecast Intervals by State and Year of Test (The number in each cell shows how many times the census value fell within the forecast interval)						
	STATE					
TEST YEAR	GEORGIA	MINNESOTA	NEW JERSEY	WASHINGTON	TOTAL	PERCENT (N/64)
1930	9	12	8	13	42	65.63%
1940	3	5	11	12	31	48.44%
1950	10	14	4	3	31	48.44%
1960	13	14	14	8	49	76.56%
1970	6	12	14	13	45	70.31%
1980	7	12	12	10	41	64.06%
1990	13	14	14	14	55	85.94%
2000	8	15	14	15	52	81.25%
2010	7	15	15	14	51	79.69%
TOTAL	76	113	106	102	397	
PERCENT (N/144)	52.78%	78.47%	73.61%	70.83%	68.92%	
					N=(144*4)=576 = (64*9)	

APPENDIX TABLE 1. DATA FOR GEORGIA

AGE GROUP	1900 CENSUS POPULATION	1910 CENSUS POPULATION	1920 CENSUS POPULATION	1930 CENSUS POPULATION	1940 CENSUS POPULATION	1950 CENSUS POPULATION	1960 CENSUS POPULATION	1970 CENSUS POPULATION	1980 CENSUS POPULATION	1990 CENSUS POPULATION	2000 CENSUS POPULATION	2010 CENSUS POPULATION
Total Population: 0 to 4 years	325,473	376,641	363,229	316,404	313,122	422,486	471,901	421,709	414,935	495,535	595,150	686,785
Total Population: 5 to 9 years	313,524	347,369	382,373	353,910	319,056	355,208	440,198	470,311	446,831	483,952	615,584	695,161
Total Population: 10 to 14 years	277,865	315,217	365,312	338,860	325,009	311,293	411,650	480,924	469,598	466,614	607,759	689,684
Total Population: 15 to 19 years	241,478	280,383	307,549	334,836	328,410	291,806	331,554	442,571	530,773	497,152	596,277	709,999
Total Population: 20 to 24 years	229,199	260,140	272,814	288,126	304,638	276,193	271,211	416,949	516,084	522,634	592,196	680,080
Total Population: 25 to 29 years	172,819	214,250	230,373	222,930	277,500	276,270	251,770	330,790	481,276	589,952	641,750	673,935
Total Population: 30 to 34 years	127,782	169,314	180,749	183,399	236,138	255,385	256,351	273,995	448,765	584,944	657,506	661,625
Total Population: 35 to 39 years	111,711	152,232	185,500	186,959	209,545	254,264	260,063	256,934	356,263	531,619	698,735	698,059
Total Population: 40 to 44 years	97,256	109,644	140,477	151,156	174,120	219,640	244,981	260,140	291,069	484,079	654,773	699,481
Total Population: 45 to 49 years	78,565	85,850	125,849	133,154	156,489	182,855	229,397	252,278	266,793	374,918	573,017	722,661
Total Population: 50 to 54 years	78,307	96,240	106,175	131,455	134,244	153,118	196,204	232,825	261,211	294,033	506,975	668,591
Total Population: 55 to 59 years	46,756	61,442	66,256	84,633	102,773	126,309	161,507	207,126	246,907	259,735	375,651	573,551
Total Population: 60 to 64 years	42,863	55,526	64,125	67,562	83,965	100,096	125,668	175,565	215,869	238,779	285,805	496,006
Total Population: 65 to 69 years	27,942	35,469	44,269	45,142	75,095	95,556	113,144	137,744	188,897	218,078	236,634	356,007
Total Population: 70 to 74 years	18,887	21,911	29,550	33,738	42,732	60,606	81,647	97,362	141,977	169,973	199,061	250,422
Total Population: 75+	19,547	23,349	28,292	34,398	40,887	63,493	95,870	132,352	185,857	266,219	349,580	425,606
Total Population	**2,209,974**	**2,604,977**	**2,892,892**	**2,906,662**	**3,123,723**	**3,444,578**	**3,943,116**	**4,589,575**	**5,463,105**	**6,478,216**	**8,186,453**	**9,687,653**
Age Not Reported	**6,357**	**4,144**	**2,940**	**1,844**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Total Population including those not reporting Age	**2,216,331**	**2,609,121**	**2,895,832**	**2,908,506**	**3,123,723**	**3,444,578**	**3,943,116**	**4,589,575**	**5,463,105**	**6,478,216**	**8,186,453**	**9,687,653**

2010 data are from Table QT-P1, 2010 Decennial Census

2000 data are from Table QT-P1, 2000 Decennial Census

1990 Data are from Table 19, *General Population Characteristics* ,1990 Decennial Census

1980 Data are from Table 19, *General Population Characteristics* , 1980 Decennial Census,

THE 1900 THROUGH 1970 POPULATION BY AGE DATA ARE FROM TABLE 21 IN "CHARACTERISTICS OF THE POPULATION, VOLUME 1" (BY STATE), 1970 CENSUS OF POPULATION

GEORGIA (VOL 1, PART 12) MARCH 1973

APPENDIX TABLE 2: DATA FOR MINNESOTA												
AGE GROUP	1900 CENSUS POPULATION	1910 CENSUS POPULATION	1920 CENSUS POPULATION	1930 CENSUS POPULATION	1940 CENSUS POPULATION	1950 CENSUS POPULATION	1960 CENSUS POPULATION	1970 CENSUS POPULATION	1980 CENSUS POPULATION	1990 CENSUS POPULATION	2000 CENSUS POPULATION	2010 CENSUS POPULATION
Total Population: 0 to 4 years	228,290	226,840	261,394	231,001	230,057	332,460	416,005	331,771	307,249	336,800	329,594	355,504
Total Population: 5 to 9 years	217,447	220,233	248,599	256,751	220,176	267,652	380,650	402,635	296,295	345,840	355,894	355,536
Total Population: 10 to 14 years	192,064	214,402	233,961	253,788	238,918	223,787	324,710	415,021	333,378	313,297	374,995	352,342
Total Population: 15 to 19 years	170,177	215,148	219,609	239,946	257,349	207,460	251,352	373,405	399,818	297,609	374,362	367,829
Total Population: 20 to 24 years	160,674	216,670	217,919	214,432	245,592	213,712	194,883	292,037	393,566	316,046	322,483	355,651
Total Population: 25 to 29 years	148,607	187,438	213,646	193,469	225,097	220,780	193,160	249,516	363,435	381,759	319,826	372,686
Total Population: 30 to 34 years	131,055	153,195	189,778	189,705	204,311	212,765	206,487	206,769	313,104	397,984	353,312	342,900
Total Population: 35 to 39 years	121,193	135,612	168,540	192,934	192,452	205,447	211,163	192,863	246,356	361,274	412,490	328,190
Total Population: 40 to 44 years	100,646	117,256	135,353	172,980	187,196	189,729	204,868	202,710	202,860	304,810	411,692	352,904
Total Population: 45 to 49 years	72,042	105,289	122,435	147,143	182,525	176,212	194,149	202,904	187,051	237,050	364,247	406,203
Total Population: 50 to 54 years	57,896	88,110	105,208	122,171	162,931	170,805	176,190	193,956	193,199	191,410	301,449	401,695
Total Population: 55 to 59 years	45,293	59,272	87,437	100,813	129,941	157,690	159,840	177,011	189,457	173,066	226,857	349,589
Total Population: 60 to 64 years	35,137	45,188	69,827	84,372	103,137	134,854	146,056	155,454	170,638	171,220	178,012	279,775
Total Population: 65 to 69 years	28,251	34,825	45,827	69,079	82,635	105,188	131,315	130,155	149,114	160,036	153,169	202,570
Total Population: 70 to 74 years	19,424	23,536	30,188	48,256	60,455	73,705	102,086	110,251	121,034	134,486	142,656	151,857
Total Population: 75+	19,096	27,696	34,751	46,145	69,528	90,237	120,950	168,513	209,416	252,412	298,441	328,694
Total Population	**1,747,292**	**2,070,710**	**2,384,472**	**2,562,985**	**2,792,300**	**2,982,483**	**3,413,864**	**3,804,971**	**4,075,970**	**4,375,099**	**4,919,479**	**5,303,925**
Age Not Reported	**4,102**	**4,998**	**2,653**	**968**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Total Population including those not reporting Age	**1,751,394**	**2,075,708**	**2,387,125**	**2,563,953**	**2,792,300**	**2,982,483**	**3,413,864**	**3,804,971**	**4,075,970**	**4,375,099**	**4,919,479**	**5,303,925**

2010 data are from Table QT-P1, 2010 Decennial Census

2000 data are from Table QT-P1, 2000 Decennial Census

1990 Data are from Table 19, *General Population Characteristics* ,1990 Decennial Census

1980 Data are from Table 19, *General Population Characteristics* , 1980 Decennial Census,

THE 1900 THROUGH 1970 POPULATION BY AGE DATA ARE FROM TABLE 21 IN "CHARACTERISTICS OF THE POPULATION, VOLUME 1" (BY STATE), 1970 CENSUS OF POPULATION
MINNESOTA (VOL 1, PART 23) JANUARY 1973

APPENDIX TABLE 3: DATA FOR NEW JERSEY

AGE GROUP	1900 CENSUS POPULATION	1910 CENSUS POPULATION	1920 CENSUS POPULATION	1930 CENSUS POPULATION	1940 CENSUS POPULATION	1950 CENSUS POPULATION	1960 CENSUS POPULATION	1970 CENSUS POPULATION	1980 CENSUS POPULATION	1990 CENSUS POPULATION	2000 CENSUS POPULATION	2010 CENSUS POPULATION
Total Population: 0 to 4 years	206,446	266,942	338,696	329,668	256,264	458,906	642,197	589,226	463,289	532,637	563,785	541,020
Total Population: 5 to 9 years	196,725	242,279	322,958	380,918	280,722	371,826	582,212	692,648	508,447	493,044	604,529	564,750
Total Population: 10 to 14 years	174,347	228,695	291,236	384,342	337,776	290,544	524,380	710,409	605,841	480,983	590,577	587,335
Total Population: 15 to 19 years	166,746	236,541	255,161	364,396	375,112	295,859	396,363	611,831	670,665	505,388	525,216	598,099
Total Population: 20 to 24 years	178,228	250,613	271,042	350,402	376,912	350,403	321,054	509,198	614,828	566,594	480,079	541,238
Total Population: 25 to 29 years	176,408	236,172	286,617	332,810	361,291	409,890	362,373	463,164	574,135	668,917	544,917	553,139
Total Population: 30 to 34 years	158,858	213,082	263,733	331,332	340,976	409,434	435,080	403,475	563,758	691,734	644,123	556,662
Total Population: 35 to 39 years	144,124	199,647	251,252	338,222	322,760	393,917	472,429	413,929	479,749	622,963	727,924	588,379
Total Population: 40 to 44 years	117,887	166,638	207,122	291,871	315,720	357,760	446,139	465,492	400,074	573,696	707,182	649,918
Total Population: 45 to 49 years	92,115	136,295	185,551	246,388	297,595	318,504	406,721	477,978	394,038	466,481	611,357	704,516
Total Population: 50 to 54 years	78,915	112,003	151,688	205,434	259,570	305,235	350,531	439,103	432,520	376,528	547,541	674,680
Total Population: 55 to 59 years	60,248	75,739	108,505	157,128	198,622	263,516	304,112	380,677	430,048	355,677	423,338	565,623
Total Population: 60 to 64 years	49,226	62,678	86,297	124,676	158,024	215,546	262,777	314,045	367,660	363,521	330,646	480,542
Total Population: 65 to 69 years	33,955	45,948	56,135	88,449	119,172	164,921	222,457	245,757	303,670	340,232	293,196	350,972
Total Population: 70 to 74 years	23,186	31,193	38,149	58,951	80,239	109,441	163,149	194,112	227,037	269,960	281,473	260,462
Total Population: 75+	22,476	29,946	39,197	53,643	79,410	119,627	174,808	257,120	329,064	421,833	538,467	574,559
Total Population	**1,879,890**	**2,534,411**	**3,153,339**	**4,038,630**	**4,160,165**	**4,835,329**	**6,066,782**	**7,168,164**	**7,364,823**	**7,730,188**	**8,414,350**	**8,791,894**
Age Not Reported	**1,128**	**662**	**792**	**244**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Total Population including those not reporting Age	**1,881,018**	**2,535,073**	**3,154,131**	**4,038,874**	**4,160,165**	**4,835,329**	**6,066,782**	**7,168,164**	**7,364,823**	**7,730,188**	**8,414,350**	**8,791,894**

2010 data are from Table QT-P1, 2010 Decennial Census

2000 data are from Table QT-P1, 2000 Decennial Census

1990 Data are from Table 19, *General Population Characteristics* ,1990 Decennial Census

1980 Data are from Table 19, *General Population Characteristics* , 1980 Decennial Census,

THE 1900 THROUGH 1970 POPULATION BY AGE DATA ARE FROM TABLE 21 IN "CHARACTERISTICS OF THE POPULATION, VOLUME 1" (BY STATE), 1970 CENSUS OF POPULATION
NEW JERSEY (VOL 1, PART 32) MARCH 1973

APPENDIX TABLE 4: DATA FOR WASHINGTON

AGE GROUP	1900 CENSUS POPULATION	1910 CENSUS POPULATION	1920 CENSUS POPULATION	1930 CENSUS POPULATION	1940 CENSUS POPULATION	1950 CENSUS POPULATION	1960 CENSUS POPULATION	1970 CENSUS POPULATION	1980 CENSUS POPULATION	1990 CENSUS POPULATION	2000 CENSUS POPULATION	2010 CENSUS POPULATION
Total Population: 0 to 4 years	53,243	108,756	126,434	114,854	121,918	263,326	315,633	280,442	306,123	366,780	394,306	439,657
Total Population: 5 to 9 years	56,423	99,678	128,258	136,013	116,762	203,786	301,051	328,397	296,011	371,093	425,909	429,877
Total Population: 10 to 14 years	48,233	92,802	117,553	138,393	127,842	159,695	275,510	348,892	321,995	337,662	434,836	438,233
Total Population: 15 to 19 years	44,104	99,647	106,485	137,922	146,725	157,695	208,575	329,903	369,023	322,711	427,968	462,128
Total Population: 20 to 24 years	46,403	122,058	111,014	130,401	148,867	175,619	173,804	295,964	400,542	351,680	390,185	461,512
Total Population: 25 to 29 years	46,093	126,074	120,421	120,651	146,594	195,087	166,376	238,704	389,997	411,822	403,652	480,398
Total Population: 30 to 34 years	47,118	106,963	119,446	115,448	134,757	188,636	179,899	193,398	354,645	443,366	437,478	453,383
Total Population: 35 to 39 years	46,368	90,149	117,587	122,833	124,990	180,749	198,495	181,020	273,382	427,690	483,950	448,607
Total Population: 40 to 44 years	37,863	77,286	95,805	118,105	118,525	159,090	189,191	192,828	213,832	376,073	491,137	459,698
Total Population: 45 to 49 years	26,027	64,992	81,764	108,280	117,709	136,714	176,071	203,880	193,473	284,674	454,223	492,909
Total Population: 50 to 54 years	20,754	52,413	69,451	90,223	112,915	125,939	150,495	188,774	198,548	216,869	391,749	495,296
Total Population: 55 to 59 years	14,127	33,661	55,053	69,260	96,698	115,306	129,003	166,878	203,986	191,602	285,505	453,078
Total Population: 60 to 64 years	10,407	24,144	42,352	57,530	77,569	103,916	110,066	138,028	179,037	189,382	211,075	382,087
Total Population: 65 to 69 years	7,195	16,585	27,298	44,440	57,963	86,551	98,659	107,008	151,324	186,679	176,225	270,474
Total Population: 70 to 74 years	4,161	10,374	16,647	30,075	41,943	59,655	80,938	84,335	112,023	149,355	160,941	186,746
Total Population: 75+	3,325	9,614	16,266	26,988	44,414	65,199	99,448	130,718	168,215	239,254	324,982	370,457
Total Population	**511,844**	**1,135,196**	**1,351,834**	**1,561,416**	**1,736,191**	**2,376,963**	**2,853,214**	**3,409,169**	**4,132,156**	**4,866,692**	**5,894,121**	**6,724,540**
Age Not Reported	**6,259**	**6,794**	**4,787**	**1,980**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Total Population including those not reporting Age	**518,103**	**1,141,990**	**1,356,621**	**1,563,396**	**1,736,191**	**2,376,963**	**2,853,214**	**3,409,169**	**4,132,156**	**4,866,692**	**5,894,121**	**6,724,540**

2010 data are from Table QT-P1, 2010 Decennial Census

2000 data are from Table QT-P1, 2000 Decennial Census

1990 Data are from Table 19, *General Population Characteristics* ,1990 Decennial Census

1980 Data are from Table 19, *General Population Characteristics* , 1980 Decennial Census,

THE 1900 THROUGH 1970 POPULATION BY AGE DATA ARE FROM TABLE 21 IN "CHARACTERISTICS OF THE POPULATION, VOLUME 1" (BY STATE), 1970 CENSUS OF POPULATION
WASHINGTON (VOL 1, PART 49) JANUARY 1973