Research Program on Forecasting

Evaluating Consensus Forecasts

Herman O. Stekler
Department of Economics
The George Washington University
Washington, DC 20052
Tel: 202-994-6150
Fax: 202-994-6147
hstekler@gwu.edu

RPF Working Paper No. 2008-007
http://www.gwu.edu/~forcpgm/2008-007.pdf

July 24, 2008

RESEARCH PROGRAM ON FORECASTING
Center of Economic Research
Department of Economics
The George Washington University
Washington, DC 20052
http://www.gwu.edu/~forcpgm

Evaluating Census Forecasts

Evaluating Census Forecasts

ABSTRACT

The Census Bureau makes periodic long-term forecasts of both the total US population and the population of each of the states. Previous evaluations of these forecasts were based on the *magnitude* of the discrepancies between the projected and actual population figures. However, it might be inappropriate to evaluate these long-term projections with the specific quantitative statistics that have been useful in judging short-term forecasts. One of the purposes of a long range projection of each state's population is to provide a picture of the **distribution** of the aggregate US population among the various states. Thus the evaluation should compare the projected **distribution** of the total US population by states to the actual distribution. This paper uses the dissimilarity index to evaluate the accuracy of the Census projected percentage distributions of population by states.

EVALUATING CENSUS FORECASTS

The Census Bureau makes periodic forecasts of the population of the United States 5, 10 or more years into the future. These forecasts are both for the total population and for the number of inhabitants of each of the states. The projections for each of the states are based on the cohort-component methodology. The projections for each state are then adjusted to the available state estimates and to the national population projections. (See Campbell (1996) for a complete explanation of the current methodology.)

There have been many evaluations of these state forecasts.[1] (For example, see Smith and Sinich, 1990, 1992; Campbell, 2002; Wang, 2002). The questions that were examined in these studies included whether the accuracy of these forecasts was affected by (1) the length of the base period used in estimating relationships, (2) the complexity of the forecasting technique and (3) the appropriateness of the statistical measures used to evaluate the projections. In all cases, the error measures were based on the *magnitude* of the discrepancies between the projected and actual population figures.

In addition to statistics that measure the quantitative errors, there are alternative procedures for evaluating these long-term projections. One of the purposes of a long range projection of each state's population is to provide a picture of the **distribution** of the aggregate

---

[1] There have also been evaluations of county forecasts and discussions of the appropriate statistical measure for evaluating projections where some of the errors are outliers, e.g. Swanson et al. (2000).

US population among the various states. There are many reasons why users of Census data might be concerned with the distribution of the population. Resources, such as highway construction funds, might be allocated on the basis of the expected future populations; politicians might be interested in knowing how the Congressional House seats will be allocated among the states; etc.

In these cases, the users of the population projections might not be concerned with the actual number of inhabitants of each state but rather with trends: whether the population of specific states was growing relative to that of other states or whether the population in a specific state was an increasing (decreasing) percent of the total national population. If one were only interested in knowing whether the projections captured the important trends that actually occurred, one might not be concerned with the magnitude of the errors. The accuracy of the quantitative projections of each state's total population is then not as relevant as an accurate depiction of major trends.

It is possible that the **share** of the nation's population that was in each state was predicted correctly, but that the national total and the estimates for each of the states were inaccurate by the same proportion. In that case, the projected distribution of the state populations would have exactly matched the observed distribution. Thus, the evaluation procedure that is suggested here does not focus on the specific numbers in the projections or the magnitude of the misestimates. Rather this evaluation asks whether the projected **share** of the total US population by states was similar to the actual distribution. Such an analysis enables one to determine whether the state distribution of the aggregate population was accurate even if the aggregate estimate is inaccurate.

The same issue, involved in evaluating long-run forecasts, has been examined in a different context: the accuracy of long-term labor-market forecasts. (Kolb and Stekler, 1992;

Stekler and Thomas, 2005). Both studies used statistics that directly measured whether major trends were predicted accurately. In evaluating the long-run projections, the first study used an information content statistic; the second used dissimilarity indexes.

This paper uses the dissimilarity index approach to evaluate the accuracy of the Census projected percentage distributions of population by states. The next section explains the methodology. This is followed by a description of the data, the results, and our interpretations and conclusions.

I. Methodology

A. Decomposing the Errors

Assume that $\boldsymbol{x_t^a}$ is the actual aggregate population of the US at time t and $\boldsymbol{x_t^f}$ is the aggregate value that was projected for time t. The error in the aggregate projection is

$$e_t = \boldsymbol{x_t^a} - \boldsymbol{x_t^f} .$$

In addition it is also possible to examine the errors associated with the population projections for each of the i states. Accordingly the proportions (p_i) of the predicted and actual (a_i) aggregated population associated with each of the i states are:

$$\boldsymbol{x_{i,t}^f} = (p_{i,t})\, \boldsymbol{x_t^f} ; \quad \boldsymbol{x_{i,t}^a} = (a_{i,t})\, \boldsymbol{x_t^a} ; \quad \sum p_{i,t} = 1, \quad \sum a_{i,t} = 1,$$

The forecast error for each state is

$$e_{i,t} = (a_{i,t})\, \boldsymbol{x_t^a} - (p_{i,t})\, \boldsymbol{x_t^f} .$$

If the aggregate forecast is absolutely accurate, the quantitative error for each state would be

$$e_{i,t} = (a_{i,t} - p_{i,t})\ \boldsymbol{x_{i,t}^a} ,$$

which is the difference between the actual and forecast proportions of the aggregate population which is in each state. The same holds true if the aggregate forecast is inaccurate. If $\boldsymbol{x_t^a} \neq \boldsymbol{x_t^f}$,

$$e_{i,t} = (a_{i,t} - p_{i,t})\ \boldsymbol{x}_t^a + p_{i,t}(\boldsymbol{x}_t^a - \boldsymbol{x}_t^f)$$

Thus the quantitative forecast error for each state, $e_{i,t}$, is the sum of two components. The first represents the error in predicting the proportion of the population in each state. The second measures the error in failing to predict the aggregate correctly. In order to evaluate these long term population forecasts, we will focus on the first term, using the dissimilarity measure as our statistic.

B. Dissimilarity Index

Suppose that one has the following data: the population of each state and the national total. Then one can calculate the percentage of the national population that resided in each state. This calculation can be made for both the projections and the actual numbers, yielding two distributions of the population by states. A dissimilarity index can be used to compare the projected and actual distributions.[2]

Specifically, the dissimilarity index measures the amount by which the projected distribution would have to change to be identical to the actual distribution. Using our notation, the formula for the dissimilarity index for every period is:

$$D_t = 0.5 \sum | (\boldsymbol{x}_{i,t}^f / \boldsymbol{x}_t^f) - (\boldsymbol{x}_{i,t}^a / \boldsymbol{x}_t^a) |, \text{ or } D_t = 0.5 \sum | p_{i,t} - a_{i,t} |$$

D is bounded in the interval 0 to 100 percent.[3] The smaller the value of D, the smaller is the difference between the predicted and actual distributions, i.e. the more accurate is the forecast.

---

[2] Davis (1994) noted that dissimilarity indexes can be used to evaluate county population estimates but did not publish any results.

[3] Percentages are used, in interpreting the results, even though p_i and a_i are defined as proportions. The dissimilarity index is bounded in the interval 0 to1 for proportions.

C. Benchmark Comparisons

For purposes of evaluation, the Census projections are compared with a benchmark. The selected benchmark must only use data that were available at time t, the date when the projection was issued. The benchmark in this case is a naïve model. We assume that the naïve projection of the states' shares of the US population for year t+h is identical to the known distribution that is available from either the Census count or from the population estimate in year t, the year from which the projection was extended.

II. Data

We evaluate the Census state population projections that were made between 1970 and 1996 for the years 1975-2005.[4] There are seven such sets of projections. The length of the forecasting horizon varied between 2 and 25 years. The naïve projections were made using the same starting points and horizons. These projections were compared either with the actual Census counts for 1980, 1990, and 2000 and or with the population estimates that the Census Bureau made for 1975, 1985, 1995 and 2005.

III. Results

The dissimilarity indexes derived from both the Census and naïve projections are presented in Table 1. The longer was the projection horizon, the larger was the size of the dissimilarity index that was associated with the projections, i.e. the less accurate the projected distribution. This result is similar to findings about the relationship between quantitative errors and the length of the forecast horizon. As indicated above, the size of these indexes measures the amount by which the projected distribution would have to change to be identical to the actual

---

[4] The data were obtained from U.S. Bureau of the Census, Current Population Reports, Series P25, Nos. 477, 735, 937, 1017, 1044, 1053, and 1111.

distribution. This was less than 1% for the very short projections to more than 5% for some of the longer horizons.

Moreover, the projections seem to have improved over time. For the 5 year projections, the values of the dissimilarity indexes declined from more than 1.5% to less than 1%. The magnitude of the index for the 10 year projection made in 1970 was almost 4%; the similar numbers for the projections made in the late 1980s and 1990s were all less than 1.5%. A similar trend was observed in the more recent 20 year projections.

Nevertheless, the Census forecasts associated with the distributions of the state population forecasts are inferior to the naïve forecasts (See Table 1). In all but one case, the dissimilarity indexes associated with the naïve forecasts are smaller than the ones derived from the comparable Census projections. The exception is the five year projection made in 1975. While the conventional way for comparing two sets of forecasts is to test whether there is a statistically significant difference, this would be difficult in this case because the distribution of the dissimilarity index is not known. Bootstrapping would have provided an alternative procedure for testing the significance of the results, but this was not necessary in this case because the naïve forecasts were an order of magnitude superior to the Census projections.

IV. Conclusions

The customary method for evaluating disaggregated long-term population projections has been concerned with the magnitude of the errors made in forecasting the absolute size of each state's population. This paper has presented an alternative evaluation method based on the difference between the predicted and actual distributions of the state projections. This evaluation of US Bureau of Census population projections made over the 1970-1996 period for forecast horizons ranging from two to twenty years yielded the following results: (a) the accuracy of the projections decreased with the length of the horizon, (b) the quality of the projections has improved over time, and (c) naïve forecasts were much superior to Census projections at all horizons. These results are consistent with previous findings that simple methods for making long-term population projections were more accurate than complex procedures. (Smith and Sinich, 1992).

References

Campbell, P. R., (1996), *Population Projections for States by Age Sex, Race and Hispanic Origin: 1995-2025,* U.S. Bureau of the Census, Population Division, PPL-47.

Campbell, P. R., (2002), Evaluating forecast error in state population projections using Census 2000 counts, U.S. Bureau of the Census, Population Division, Working Paper Series No. 57.

Davis, S.T., (1994), Evaluation of postcensal county estimates for the 1980s, U.S. Bureau of the Census, Population Division, Working Paper Series No. 5.

Kolb, R. A. and Stekler, H. O., (1992), Information Content of long-term employment forecasts, *Applied Economics*, 24, 593-596.

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

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

Stekler, H. O. and Thomas, R., (2005), Evaluating BLS labor force, employment, and occupation projections for 2000, *Monthly Labor Review*, July 2005, 46-56.

Swanson, D. A., Tayman, J., and Barr, C. F., (2000), A note on the measurement of accuracy for he subnational demographic estimates, *Demography*, 37,193-201.

Wang, C., (2002), Evaluation of Census Bureau’s 1995-2005 state population projections, Working Paper no. 67, US Census Bureau.

Table 1
Values of Dissimilarity Index (percentage points)
Census and Naive Forecasts

Date
Projections
Made

Date of Projection

	1975	1980	1985	1990	1995	2000	2005
1970	a 1.7 b 1.7 [0.2]	a 3.9 b 3.9 [0.4]	a 5.5 b 5.5 [0.7]	a 6.5 b 6.4 [0.8]			
1975		1.8 [2.7]		3.7 [0.7]		5.3 [0.9]	
1980				2.5 [0.4]		4.2 [0.7]	
1986				1.0 [0.2]	1.2 [0.3]	1.9 [0.4]	2.3 [0.4]
1988				a 0.6 b 0.8 [0.1]		a 2.5 b 1.7 [0.4]	
1992					0.4 [0.1]	1.4 [0.2]	2.0 [0.4]
1996						0.8 [0.2]	1.2 [0.3]

Notes: Numbers in [ ] are for naive (benchmark) projections.

There were two sets of projections issued in 1970 and 1988. They are denoted a and b.