EMPIRICAL PREDICTION INTERVALS FOR COUNTY POPULATION FORECASTS

Stefan Rayer, University of Florida

Stanley K. Smith, University of Florida

Jeff Tayman, University of California, San Diego

ABSTRACT

Population forecasts entail a significant amount of uncertainty, especially for long-range horizons and for places with small or rapidly changing populations. This uncertainty can be dealt with by presenting a range of projections or by developing statistical prediction intervals. The latter can be based on models that incorporate the stochastic nature of the forecasting process, on empirical analyses of past forecast errors, or on a combination of the two. In this article, we develop and test prediction intervals based on empirical analyses of past forecast errors for counties in the United States. Using decennial census data from 1900 to 2000, we apply trend extrapolation techniques to develop a set of county population forecasts; calculate forecast errors by comparing forecasts to subsequent census counts; and use the distribution of errors to construct empirical prediction intervals. We find that empirically-based prediction intervals provide reasonably accurate predictions of the precision of population forecasts, but provide little guidance regarding their tendency to be too high or too low. We believe the construction of empirically-based prediction intervals will help users of small-area population forecasts measure and evaluate the uncertainty inherent in population forecasts and plan more effectively for the future.

Key words: Demography; small areas; population projections; forecast uncertainty; accuracy.

INTRODUCTION

Population forecasts entail a significant amount of uncertainty, especially for long-range horizons and for places with small or rapidly changing populations. Almost 40 years ago, Keyfitz (1972) argued that demographers should warn forecast users about this uncertainty. Warnings have typically been given by presenting a range of forecasts (e.g., Hollmann et al., 2000), but in recent years demographers have developed statistical prediction intervals that provide an explicit probability statement regarding the level of error expected to accompany a population forecast (e.g., Alho and Spencer, 1997; de Beer, 2000; Lee, 1999; Lutz and Goldstein, 2004).

Most of the research on probabilistic population forecasting has focused on large geographic areas. National-level analyses have been performed for Australia (Wilson and Bell, 2004), Austria (Lutz and Scherbov, 1998), Finland (Alho, 2002), the Netherlands (de Beer and Alders, 1999), Norway (Keilman et al., 2002), Poland (Matysiak and Nowok, 2007), Sweden (Cohen, 1986), and the United States (Lee and Tuljapurkar, 1994). Less research has been performed at the subnational level, but Miller (2002) produced a series of probabilistic forecasts for California; Rees and Turton (1998) investigated model input uncertainty for 71 regions in 12 countries in the European Union; Tayman et al. (2007) evaluated ARIMA models for four states in the United States; Wilson and Bell (2007) developed probabilistic forecasts for Queensland and the rest of Australia; and Gullickson and Moen (2001) used a probabilistic model for forecasting hospital admissions for two regions in Minnesota. Very little research has been done for small areas such as cities or counties.

We believe research on probabilistic forecasting methods at the subnational level is essential because state and local population forecasts are so widely used for planning, budgeting, and analytical purposes. Examples include planning for future water consumption in Texas

(Texas Water Development Board, 1997), choosing locations for new fire stations in San Diego (Tayman et al., 1994), evaluating the demand for hospital services in a southern metropolitan area (Thomas, 1994), developing conservation plans for a river basin in Arizona and Mexico (Steinitz et al., 2003), and projecting future enrollments for public school districts in Indiana (McKibben, 1996). Effective decision making for projects such as these cannot be accomplished without a clear understanding of the likely level of accuracy of the underlying population forecasts.

Probabilistic prediction intervals can be based on models that incorporate the stochastic nature of the forecasting process (e.g., Alho and Spencer, 1990; Cohen, 1986; Lutz et al., 1999; Pflaumer, 1992), on empirical analyses of past forecast errors (e.g., Bongaarts and Bulatao, 2000; Keilman, 1997; Keyfitz, 1981; Smith and Sincich, 1988; Stoto, 1983; Tayman et al., 1998), or on a combination of the two (e.g., de Beer, 1997). In this study, we focus on prediction intervals based on empirical analyses of past forecast errors.

The usefulness of the empirical approach rests heavily on the assumption that the distribution of population forecast errors remains relatively stable over time. Not all forecasters accept the validity of this assumption (e.g., Alho and Spencer, 1997), but few have tested it empirically. Perhaps the most comprehensive evaluation to date was conducted by Smith and Sincich (1988), who examined state-level population forecasts using data from 1900 to 1980. Following a methodology developed by Williams and Goodman (1971), they evaluated forecast errors for 10- and 20-year horizons and found that the means and variances of absolute forecast errors remained relatively stable over time, especially after 1920, and that the variances of algebraic forecast errors remained moderately stable over time but their means did not. They

concluded that the study of past forecast errors can be useful for predicting the level of precision of current population forecasts, but not for predicting their tendency to be too high or too low.

In this article, we follow the approach described by Smith and Sincich (1988) to investigate the construction and performance of empirical prediction intervals. Using 100 years of population data and a large sample of counties in the United States, we applied trend extrapolation techniques to develop a set of county population forecasts; calculated forecast errors by comparing forecasts to decennial census counts; used the distribution of past forecast errors to construct empirical prediction intervals; and investigated whether those intervals provided accurate predictions of subsequent forecast errors. We found that those intervals provided reasonably accurate predictions of the *precision* of subsequent forecasts, but we did not test predictions of *bias* because the evidence showed that past forecast errors provide little guidance regarding the tendency for subsequent forecasts to be too low or too high. We believe this and other studies of empirical prediction intervals will help forecast users measure and evaluate the uncertainty inherent in population forecasts and plan more effectively for the future.

DATA AND FORECASTING TECHNIQUES

We used decennial census data from 1900 to 2000 to construct and analyze population forecasts for counties (or county equivalents) in the United States. We restricted our analysis to the 2,482 counties for which no significant boundary changes occurred between 1900 and 2000; this group accounted for 79% of all current counties. Forecast errors for this subset were very similar to errors for a larger group of 2,978 counties whose boundaries did not change significantly after 1930 (not shown here). We used the smaller group with constant boundaries since 1900 because it permitted the analysis of a larger number of launch years and forecast horizons.

We use the following terminology to describe population forecasts:

1) Base year: the year of the earliest population size used to make a forecast.

2) Launch year: the year of the latest population size used to make a forecast.

3) Target year: the year for which population size is forecasted.

4) Base period: the interval between the base year and launch year.

5) Forecast horizon: the interval between the launch year and target year.

For example, if data from 1900 and 1920 were used to forecast population in 1930, then 1900 would be the base year, 1920 the launch year, 1930 the target year, 1900–1920 the base period, and 1920–1930 the forecast horizon.

We made forecasts of total population for each county using seven trend extrapolation techniques (see the Appendix for a description of these techniques). These forecasts were calculated using a 20-year base period, the base period producing the most accurate forecasts for counties in this data set (not shown here). The forecasts had launch years ranging from 1920 to 1990 and forecast horizons ranging from 10 to 30 years. The 21 combinations of launch year and forecast horizon – and their associated target years – are shown in Table 1.

[Table 1 about here.]

Trend extrapolation techniques have a number of advantages compared to more complex forecasting methods. They require less base data, can be employed at lower cost, and can more easily be applied retrospectively to produce forecasts that are comparable over time. These characteristics are particularly important when making forecasts for a large number of geographic areas and historical time periods. Though simple in design, trend extrapolation techniques have been found to produce forecasts of total population that are at least as accurate as those produced by more complex methods (e.g., Ascher, 1978; Isserman, 1977; Long, 1995;

6

Murdock et al., 1984; Rayer, 2008; Smith and Sincich, 1992; Stoto, 1983). Similar evidence has been reported in other fields as well (e.g., Makridakis, 1986; Makridakis and Hibon, 1979; Mahmoud, 1984). Given these characteristics, trend extrapolation techniques provide a useful vehicle for assessing the stability of population forecast errors over time and for testing the validity of empirical prediction intervals.

We calculated the average of the forecasts from these seven techniques for each county (AV7) and the average after the highest and lowest forecasts were excluded (AV5). The latter measure reduces the impact of outliers and is often called a trimmed mean; we found it produced slightly smaller forecast errors than AV7. A number of studies have documented the benefits of combining forecasts, both in demography and other fields (e.g., Armstrong, 2001; Makridakis et al., 1998; Rayer, 2008; Smith et al., 2001; Webby and O'Connor, 1996). We focus primarily on the results for AV5 in this article; many of the results for AV7 and for individual techniques were similar to those reported here for AV5. To investigate how different functional forms affect the performance of empirical prediction intervals, we conclude the analysis with an evaluation of forecasts based on individual trend extrapolation techniques.

ANALYSIS OF FORECAST ERRORS

Forecasts for each county were made for each of the 21 launch year/forecast horizon combinations shown in Table 1 and were compared to census counts for each target year. The resulting differences are called forecast errors, although they may have been caused partly by errors in the census counts themselves. All errors are reported as percentages by dividing by census counts and multiplying by 100. We refer to errors that ignore the direction of error as absolute percent errors (APEs) and errors that account for the direction of error as algebraic

7

percent errors (ALPEs). The former reflect the *precision* of population forecasts and the latter reflect their *bias*.

Several summary measures were used to provide a general description of forecast accuracy. For precision, we start with the most commonly used measure in population forecasting, the mean absolute percent error (MAPE), which shows how close the forecasts were to population counts regardless of whether they were too high or too low. We also report the 90th percentile error (90[th]PE, calculated as the APE that was larger than exactly 90% of all APEs). The 90[th]PE provides a different perspective on precision by reducing the impact of outliers. For bias, we report the mean algebraic percent error (MALPE), which shows the tendency for forecasts to be too high or too low. Finally, we report the standard deviation, which measures the spread of APEs and ALPEs around the mean. These and similar measures have been used frequently to evaluate the accuracy of population forecasts (e.g., Isserman, 1977; Pflaumer, 1992; Rayer, 2007; Smith and Sincich, 1988; Tayman et al., 1998).

Table 2 shows county forecast errors for AV5 by target year and forecast horizon. Three patterns stand out regarding measures of precision. First, errors increased about linearly with the length of the forecast horizon. For each ten year increase in the forecast horizon, MAPEs rose by about 10 percentage points and 90[th]PEs rose by about 20 percentage points.

[Table 2 about here.]

Second, differences in errors across target years within each forecast horizon were relatively small. For 10-year horizons, for example, MAPEs between target years 1930 and 1980 varied only from 9.6 to 13.2, standard deviations varied from 9.7 to 13.9, and 90[th]PEs varied from 21.0 to 29.1. Forecast precision improved for target years 1990 and 2000, with all three measures showing lower values than in earlier years. The smaller errors for 1990 and 2000 can

8

largely be explained by the greater average population size and more moderate growth rates of counties toward the end of the 20th century. Averaged across all target years, the MAPE for 10-year horizons was 10.2, with a standard deviation of only 2.3. A similar degree of temporal stability can be seen for 20- and 30-year horizons.

Third, ratios of standard deviations to MAPEs were similar for all target years and forecast horizons, fluctuating in a narrow range around 1.0. For the average based on all target years within a particular horizon, this ratio was 0.980 for the 10-year horizon, 0.931 for the 20-year horizon, and 0.939 for the 30-year horizon. Even for target years in which MAPEs were considerably different than for other target years (e.g., 1990 and 2000 for 10-year horizons and 2000 for 20-year horizons), the ratios did not diverge substantially from 1.0. This implies that the distribution of errors around the mean remained fairly constant over time.

Unlike MAPEs, MALPEs varied considerably from one target year to another. There was a tendency for MALPEs to be positive for earlier target years and negative for later target years, but this relationship was fairly weak (except for 30-year forecast horizons). Standard deviations for ALPEs were much larger than their means and were consistently larger than standard deviations for APEs. Standard deviations for both APEs and ALPEs were generally smaller in later target years than in earlier target years, especially in 1990 and 2000 for 10-year horizons and 2000 for 20-year horizons. That is, the spread around the mean declined over time for both APEs and ALPEs. Averaged over all target years, MALPEs were very small for all three forecast horizons, suggesting that there was little systematic bias in the population forecasts.

We constructed histograms for each target year and forecast horizon to provide a more detailed look at the distribution of errors. Results for algebraic percent errors for AV5 for 20-year horizons are shown in Figure 1. The distributions were roughly normal for each target year,

9

although earlier years had more outliers than later years. The spread of errors around the mean declined slowly between 1940 and 1990 but declined substantially between 1990 and 2000. The center of the distribution fluctuated considerably over time, sometimes falling above zero and sometimes falling below zero. Histograms for the other forecast horizons showed similar patterns, although distributions were generally more compact for 10-year horizons and less compact for 30-year horizons (not shown here).

[Figures 1 and 2 about here.]

Figure 2 shows the corresponding results for absolute percent errors. The distributions were asymmetric and truncated at zero. Errors became slightly more concentrated near zero between 1940 and 1990 but became substantially more concentrated in 2000, when there were considerably more small errors and fewer large errors than in previous years. Distributions for 10- and 30- year horizons were similar to those shown here, but were concentrated closer to zero for the former than the latter (not shown here).

The results summarized in Table 2 and Figures 1 and 2 show that many of the forecasts analyzed in this study had large errors, especially for long forecast horizons. The prevalence of large errors may be disappointing news to forecast users, but it is an accurate reflection of reality and highlights the importance of developing measures of uncertainty to accompany small-area population forecasts. This is the topic we turn to next.

EMPIRICAL PREDICTION INTERVALS

Smith and Sincich (1988) conducted formal statistical tests of the stability of MAPEs, MALPEs, and their standard deviations over time. We replicated those tests using our data set and found only weak evidence of stability (not shown here). We believe our results differed from those

reported by Smith and Sincich because counties exhibit much more variability than states with respect to population size and growth rates; both of these factors influence forecast errors. Perhaps more important, Smith and Sincich's data set had only 50 units of analysis whereas our data set has almost 2,500. In a large data set, even small differences in means and variances (e.g., one percentage point) may be statistically significant.

The data shown in Table 2 and Figures 1 and 2 provide evidence of a fairly high degree of stability over time in MAPEs and the distribution of APEs. This suggests that – despite the lack of stability implied by formal statistical tests – past forecast errors may help us predict the level of precision of current forecasts. However, the data showed no stability in MALPEs and the center of the distribution of ALPEs, suggesting that past forecast errors are not likely to help us predict the tendency for current forecasts to be too high or too low. A number of previous studies have drawn similar conclusions (e.g., Isserman, 1977; Kale et al., 1981; Smith and Sincich, 1988; Tayman et al., 1998). We therefore focus on APEs in our efforts to develop and evaluate empirical prediction intervals.

Smith and Sincich (1988) used information on the distribution of past APEs to predict the distribution of subsequent APEs. A major advantage of this approach is that it can accommodate any type of error distribution, including the asymmetric and truncated distributions characteristic of APEs. It also permits an assessment of the prediction intervals themselves; that is, we can compare the actual number of errors falling within the intervals with the predicted number. Following Smith and Sincich, we use 90[th]PEs to construct empirical prediction intervals.

Overall Results

We began by ranking APEs for AV5 for each of the 21 sets of forecasts and selecting the 90^{th}PE, as shown in Table 2. Then, we used the 90^{th}PE from target year t-n as the forecast of the 90^{th}PE in target year t, where n is the length of the forecast horizon. For example, if 1950 was the target year for a 10-year forecast based on launch year 1940, the 90^{th}PE for 1950 would be used to predict the 90^{th}PE for 1960 for a 10-year forecast based on launch year 1950. If error distributions remain relatively stable over time, 90^{th}PEs from past distributions will provide reasonably accurate predictions of future 90^{th}PEs. To assess the validity of this hypothesis, we compared the predicted with the actual 90^{th}PE for each target year and computed the percentage of APEs that fell within the predicted values.

Table 3 shows the percentage of APEs in each target year that was less than the predicted 90^{th}PE. The numbers can be interpreted as follows: A value of 90 reflects a perfect prediction. Values below 90 indicate that the 90^{th}PE for target year t was greater than the 90^{th}PE for target year t-n (i.e., fewer APEs fell within the predicted range). Values above 90 indicate the opposite. In addition to errors for each target year, Table 3 shows 90^{th}PEs averaged across all target years for each horizon, along with the standard deviation associated with the average 90^{th}PE.

[Table 3 about here.]

Table 3 reflects a high degree of stability for averages covering all the target years within a given forecast horizon: approximately 91% of APEs fell within the predicted 90^{th}PE for all three horizons. There was some variability when comparing results for individual target years within each horizon, but for the most part the values did not stray far from 90. For all three horizons, standard deviations were nearly identical and were small relative to their means, further demonstrating temporal stability. In this analysis, then, empirically-based prediction

12

intervals performed well: in most instances, intervals based on the distribution of past forecast errors encompassed approximately 90% of subsequent forecast errors.

It is possible that using data from several historical time periods would provide better results than using data from a single time period. To test this possibility, we evaluated the percentage of 90[th]PEs that was less than the average of the *two* previous target years (not shown here). This adjustment had little impact on the results, generally leading to errors that were slightly *larger* than those shown here. In this sample, then, data from a single time period were sufficient for constructing empirical prediction intervals.

In order to investigate the impact of the choice of cut-off points for prediction intervals, we replicated the analysis using 75th percentile errors (75[th]PEs) instead of 90th percentile errors (not shown here). For the averages covering all the target years within a given forecast horizon, between 75% and 77% of APEs fell within the predicted 75[th]PE for all three horizons, reflecting a fair amount of temporal stability. However, there was more variability from one target year to another than for 90[th]PEs, which was caused by the greater concentration of APEs around the 75[th]PE than the 90[th]PE. Consequently, small differences in the size of the predicted percentile error led to a larger difference in the percentage of APEs falling within the predicted value. This can be seen in Figures 1 and 2: the further the distance from the center of the error distribution, the lower the concentration of APEs around a particular percentile error.

Results by Population Size and Growth Rate

A number of studies have found that forecast precision improves with increases in population size and declines with increases in the absolute value of the growth rate (e.g., Keyfitz, 1981; Murdock et al., 1984; Smith and Sincich, 1992; Stoto, 1983; White, 1954). Others have found

bias to be unrelated to population size but positively related to the growth rate (e.g., Isserman, 1977; Rayer, 2008; Smith, 1987; Smith and Sincich, 1988; Tayman, 1996). To investigate the effects of these variables on the performance of prediction intervals, we extended our analysis to counties grouped by population size and growth rate.

Table 4 shows 90[th]PEs for AV5 for counties grouped by population size in the launch year. Confirming the results of previous studies, errors generally declined as population size increased for each target year and length of forecast horizon, with the largest declines typically occurring in the move from the smallest to the next-smallest size category. Furthermore, standard deviations generally declined relative to their means as population size increased, reflecting less decade-to-decade variation in errors for large counties than small counties.

[Table 4 about here.]

How do differences in population size affect the performance of empirical prediction intervals? Table 5 shows the percentage of APEs for AV5 that was less than the predicted 90[th]PE by population size in the launch year. In general, differences by population size were fairly small and followed no consistent pattern. For some combinations of target year and length of horizon, the percentages rose with population size; for others, they fell; and for some, they followed no clear pattern. The standard deviations were small and did not vary much among the four size categories or by length of forecast horizon. Although 90[th]PEs themselves varied considerably with differences in population size, it appears that differences in population size had no consistent impact on the predictability of 90[th]PEs.

[Table 5 about here.]

Table 6 shows 90[th]PEs for AV5 for counties grouped by the average per decade rate of population growth during the base period, for each combination of target year and forecast

14

horizon. Errors generally displayed a U-shaped pattern, with higher values for counties with large negative growth rates, smaller values for counties with moderate growth rates, and higher values for counties with large positive growth rates. These patterns are also consistent with those reported in previous studies. Standard deviations followed a U-shaped pattern for 10- and 20-year forecast horizons, but rose steadily with the growth rate for the 30-year horizon.

[Table 6 about here.]

Table 7 shows the percentage of APEs for AV5 that was less than the predicted 90^{th}PE by the average per decade rate of population growth during the base period. In contrast to differences in population size, differences in growth rates had a fairly consistent impact on the performance of prediction intervals: there was a tendency for the percentage of APEs that was less than the predicted 90^{th}PE to increase with the growth rate. Values were generally smallest for counties in the lowest growth category and rose with increases in the growth rate. Furthermore – as indicated by standard deviations that declined as growth rates increased for all three lengths of forecast horizon – values for individual target years varied most for counties with rapidly declining populations and varied least for counties with rapidly growing populations. That is, there was more consistency in the results across target years for rapidly growing populations than for rapidly declining populations. Future research may lead to techniques for modifying prediction intervals by accounting for differences in population growth rates.

[Table 7 about here.]

Results for Individual Techniques

15

Our analysis thus far has focused on AV5, the average of forecasts from the seven individual techniques after the highest and lowest were omitted. Would similar results be found for the individual techniques themselves? To answer this question, we replicated Table 2 for each individual technique (not shown here). We found many similarities in summary error statistics but several differences as well. For example, the exponential technique generally had larger MAPEs and 90[th]PEs than other techniques, especially for long forecast horizons, and displayed a strong upward bias while most of the other techniques displayed a downward bias.

In spite of these differences, prediction intervals based on the distribution of past errors performed well for most individual techniques (see Table 8). For 10-year horizons, the overall percentage of errors falling within the predicted 90[th]PE for individual techniques ranged only from 89.7 to 91.6; for 20-year horizons, from 89.5 to 91.6; and for 30-year horizons, from 84.2 to 91.3 (89.4 to 91.3 for all but the shift-share technique). There was more variability for individual target years, but in many instances the percentages were fairly close to 90. It appears that differences in the functional form of the forecasting technique had little impact on the performance of empirical prediction intervals.

[Table 8 about here.]

SUMMARY AND CONCLUSIONS

Under formal definitions, probability statements regarding the accuracy of population forecasts cannot be made because the distribution of future errors is unknown (and unknowable) at the time forecasts are made. However, if current forecasting methods are about as accurate as those used in the past, and if the degree of uncertainty will be about the same in the future as it was in the past, then we can assume that future forecast errors will be drawn from the same distribution

16

as past forecast errors (Keyfitz, 1981). If this is true, prediction intervals based on the distribution of past forecast errors will provide reasonably accurate measures of the uncertainty surrounding subsequent population forecasts. This is the issue we address in the present study.

We calculated population forecast errors for 2,482 counties in the United States throughout the 20th century and evaluated the performance of empirical prediction intervals based on the distribution of past forecast errors. We found that:

1) MAPEs and 90[th]PEs remained fairly constant over time, but declined over the last few decades in the century.

2) The standard deviations for APEs and ALPEs also remained fairly constant over time, but declined over the last few decades in the century.

3) MALPEs did not remain at all constant over time.

4) MAPEs and 90[th]PEs increased with the length of the forecast horizon, often in a nearly linear manner.

5) In most instances, the 90[th]PE from one time period provided a reasonably accurate forecast of the percentage of APEs falling within the predicted 90% interval in the following time period, even for long forecast horizons.

6) Differences in population size had little impact on the percentage of APEs falling within the predicted 90% interval, but differences in population growth rates had a fairly consistent impact.

Of particular interest is the finding that 90[th]PEs from previous error distributions provided reasonably accurate predictions of subsequent 90[th]PEs. Given the tremendous shifts in population trends that occurred during the 20th century, this is a notable finding. Drawing on these results and those reported in previous studies (e.g., Keyfitz, 1981; Smith and Sincich, 1988;

Stoto, 1983; Tayman et al., 1998), we conclude that empirical prediction intervals based on the distribution of past forecast errors can provide useful information regarding the likely level of precision of current population forecasts.

We did not construct empirical prediction intervals for measures of bias. As shown in Table 2 and Figure 1, MALPEs and the center of the distribution of ALPEs did not remain stable over time. Although bias has been found to be related to differences in population growth rates, it appears that past forecast errors do not provide useful information regarding the tendency for a particular set of forecasts to be too high or too low (e.g., Isserman, 1977; Kale et al., 1981; Smith and Sincich, 1988; Tayman et al., 1998).

The prediction intervals analyzed in this study were based on forecasts derived from seven simple trend extrapolation techniques. Would similar results be found for forecasts based on cohort-component models and other complex forecasting methods? The available evidence suggests this to be the case. A number of studies have found that average forecast errors for alternative methods tend to be similar when those methods are applied to the same geographic areas and time periods (e.g., Ascher, 1978; Isserman, 1977; Kale et al., 1981; Long, 1995; Morgenroth, 2002; Pflaumer, 1992; Rayer, 2008; Smith and Sincich, 1992; Smith and Tayman, 2003). These studies examined actual, published forecasts as well as simulated forecasts based on extrapolation techniques. Furthermore, the present study found that empirical prediction intervals worked well for almost every individual technique, including techniques based on very different functional forms. Additional research is needed, but it is likely that empirically-based prediction intervals can be usefully employed in conjunction with many different population forecasting methods.

Probabilistic prediction intervals can be based on models that incorporate the stochastic nature of the forecasting process, on empirical analyses of past forecast errors, or on a combination of the two. Which of these approaches is most useful for counties and other small areas? Model-based prediction intervals require a substantial amount of base data and are subject to errors in specifying the model, errors in estimating the model's parameters, and structural changes that invalidate the model's parameter estimates over time (Lee, 1992). In addition, many different models can be specified, each providing a different set of prediction intervals (e.g., Cohen, 1986; Keilman et al., 2002; Sanderson, 1995; Tayman et al., 2007). These limitations make model-based intervals more difficult to apply than empirically-based intervals and raise questions about their reliability. Although prediction intervals based on stochastic models may be useful for national and perhaps for state population forecasts, we believe an empirical approach will generally be more useful for counties and other small areas.

Empirically-based prediction intervals have their own limitations, of course. We found that more than 90% of APEs fell inside the 90% prediction intervals in some target years and less than 90% in other target years. Intervals based on 75[th]PEs did not perform quite as well as intervals based on 90[th]PEs. Perhaps more important, the empirical approach does not provide reliable forecasts of the likely direction of error for a particular set of population forecasts.

Further research on empirical prediction intervals is clearly needed. Can formal criteria be established for evaluating the stability of error distributions over time? How much historical data are needed to develop the most stable intervals? Can techniques be developed for adjusting prediction intervals to account for differences in population size, growth rate, geographic region, and perhaps other factors as well? How do differences in the choice of cut-off points (e.g., 90th vs. 75th percentile) affect the accuracy of forecast error predictions? Can information on the

19

distribution of errors for one geographic region be used to develop prediction intervals for another region? Would the performance of prediction intervals for other functional units, i.e. areas bound together by economic and political ties such as metropolitan areas or planning districts, differ from that shown here for counties? Can techniques be developed for predicting the tendency of forecasts to be too high or too low? Answers to these and similar questions promise to increase our understanding of population forecast errors and improve the performance of empirically-based prediction intervals.

APPENDIX: TREND EXTRAPOLATION TECHNIQUES

We used the following forecasting techniques: linear (LIN), modified linear (MLN), share-of-growth (SHR), shift-share (SFT), exponential (EXP), constant-share (COS), and constant (CON).

LIN: The linear technique assumes that the population will increase (decrease) by the same number of persons in each future decade as the average per decade increase (decrease) observed during the base period:

$$P_t = P_l + x/y \, (P_l - P_b),$$

where P_t is the population in the target year, P_l the population in the launch year, P_b the population in the base year, x the number of years in the forecast horizon, and y the number of years in the base period.

MLN: The modified linear technique initially equals the linear technique, but in addition distributes the difference between the sum of the linear county forecasts and an independent national forecast proportionally by population size at the launch year:

$$P_{it} = \text{LIN} + P_{il}/P_{jl} \, (P_{jt} - \Sigma \text{LIN}),$$

where i represents the county and j the nation.

SHR: The share-of-growth technique assumes that each county's share of population growth will be the same over the forecast horizon as it was during the base period:

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

SFT: The shift-share technique assumes that the average per decade change in each county's share of the national population observed during the base period will continue throughout the forecast horizon:

$$P_{it} = P_{jt} \, [P_{il}/P_{jl} + (x/y) \, (P_{il}/P_{jl} - P_{ib}/P_{jb})].$$

EXP: The exponential technique assumes the population will grow (decline) by the same rate in each future decade as during the base period:

$$P_t = P_l e^{rx}, \qquad r = [\ln (P_l / P_b)] / y,$$

where e is the base of the natural logarithm and ln is the natural logarithm.

COS: The constant-share technique assumes the county's share of the national population will be the same in the target year as it was in the launch year:

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

CON: The constant technique assumes that the county population in the target year is the same as in the launch year:

$$P_t = P_l.$$

Four of these techniques (MLN, SFT, SHR, and COS) require an independent national forecast for the target year population. Since no set of national forecasts covers all the launch years and forecast horizons used in this study, we constructed a set by applying the linear and exponential techniques to the national population. We used an average of these two forecasts as a forecast of the U.S. population.

REFERENCES

Alho, J. M. (2002). The population of Finland in 2050 and beyond. Discussion Paper No. 826.

    The Research Institute of the Finnish Economy. Helsinki, Finland.

Alho, J. M., & Spencer, B. D. (1990). Error models for official mortality forecasts. *Journal of the*

    American Statistical Association, 85, 609–616.

Alho, J. M., & Spencer, B. D. (1997). The practical specification of the expected error of

    population forecasts. *Journal of Official Statistics*, 13, 203–225.

Armstrong, J. S. (2001). *Principles of forecasting*. Boston: Kluwer Academic Publishers.

Ascher, W. (1978). *Forecasting: An appraisal for policy makers and planners*. Baltimore: Johns

    Hopkins University Press.

Bongaarts, J., & Bulatao, R. A. (eds.). (2000). *Beyond six billion: Forecasting the world's*

    population. Washington DC: National Academy Press.

Cohen, J. E. (1986). Population forecasts and confidence intervals for Sweden: A comparison of

    model-based and empirical approaches. *Demography*, 23, 105–126.

de Beer, J. (1997). The effect of uncertainty of migration on national population forecasts: The

    case of the Netherlands. *Journal of Official Statistics*, 13, 227–243.

de Beer, J. (2000). Dealing with uncertainty in population forecasting. Demographic Working

    Paper, Statistics Netherlands.

de Beer, J., & Alders, M. (1999). Probabilistic population and household forecasts for the

    Netherlands. Paper for the European Population Conference EPC99, 30 August – 3

    September 1999, The Hague, The Netherlands.

Gullickson, A., & Moen, J. (2001). The use of stochastic methods in local area population

    forecasts. Paper presented at the annual meeting of the Population Association of

    America, Washington DC.

Hollmann, F. W., Mulder, T. J., & Kallan, J. E.. (2000). Methodology and assumptions for the

    population projections of the United States: 1999 to 2100. Population Division Working

    Paper No. 38. Washington DC: US Census Bureau.

Isserman, A. M. (1977). The accuracy of population projections for subcounty areas. *Journal of*

    the American Institute of Planners, 43, 247–259.

Kale, B. D., Voss, P. R., Palit, C. D., & Krebs, H. C. (1981). On the question of errors in

    population projections. Paper presented at the annual meeting of the Population

    Association of America, Washington DC.

Keilman, N. (1997). Ex-post errors in official population forecasts in industrialized countries.

    Journal of Official Statistics, 13, 245–277.

Keilman, N., Pham, D. Q., & Hetland, A. (2002). Why population forecasts should be

    probabilistic – illustrated by the case of Norway. *Demographic Research*, 6, 409–453.

Keyfitz, N. (1972). On future population. *Journal of the American Statistical Association*, 67,

    347–362.

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

    579–593.

Lee, R. D. (1992). Stochastic demographic forecasting. *International Journal of Forecasting*, 8,

    315–327.

Lee, R. D. (1999). Probabilistic approaches to population forecasting. In W. Lutz, J. W. Vaupel, & D. A. Ahlburg (Eds.), *Frontiers of Population Forecasting* (pp. 156–190). New York: The Population Council. (A supplement to Population and Development Review, 24.)

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

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

Lutz, W., & Goldstein, J. R. (2004). Introduction: How to deal with uncertainty in population forecasting? *International Statistical Review*, 72, 1–4.

Lutz, W., & Scherbov, S. (1998). An expert-based framework for probabilistic national population forecasts: The example of Austria. *European Journal of Population*, 14, 1–17.

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

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

Makridakis, S. (1986). The art and science of forecasting: An assessment and future directions. *International Journal of Forecasting*, 2, 15–39.

Makridakis, S., & Hibon, M. (1979). Accuracy of forecasting: An empirical investigation. *Journal of the Royal Statistical Society A*, 142, 97–145.

Makridakis, S., Wheelwright, S., & Hyndman, R. (1998). *Forecasting: Methods and applications*. New York: Wiley.

Matysiak, A., & Nowok, B. (2007). Stochastic forecast of the population of Poland, 2005–2050. *Demographic Research*, 17, 301–338.

McKibben, J. N. (1996). The impact of policy changes on forecasting for school districts. *Population Research and Policy Review*, 15, 527–536.

Miller, T. (2002). California's uncertain population future. Technical Appendix for Lee, Miller, and Edwards (2003) The Growth and Aging of California's Population: Demographic and Fiscal Projections, Characteristics and Service Needs. Center for the Economics and Demography of Aging. CEDA Papers: Paper 2003-0002CL. University of California, Berkeley.

Morgenroth, E. (2002). Evaluating methods for short to medium term county population forecasting. *Journal of the Statistical and Social Inquiry Society of Ireland*, 31, 111–136.

Murdock, S. H., Leistritz, F. L., Hamm, R. R., Hwang, S. S, & Parpia, B. (1984). An assessment of the accuracy of a regional economic-demographic projection model. *Demography*, 21, 383–404.

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

Rayer, S. (2007). Population forecast accuracy: Does the choice of summary measure of error matter? *Population Research and Policy Review*, 26, 163–184.

Rayer, S. (2008). Population forecast errors: A primer for planners. *Journal of Planning Education and Research*, 27, 417–430.

Rees, P., & Turton, I. (1998). Investigation of the effects of input uncertainty on population forecasting. Paper presented at the 3rd International Conference on GeoComputation, September 17–19, 1998, Bristol, UK.

Sanderson, W. C. (1995). Predictability, complexity, and catastrophe in a collapsible model of

  population, development, and environmental interactions. *Mathematical Population*

  Studies, 5, 259–279.

Smith, S. K. (1987). Tests of forecast accuracy and bias for county population projections.

  Journal of the American Statistical Association, 82, 991–1003.

Smith, S. K., & Sincich, T. (1988). Stability over time in the distribution of forecast errors.

  Demography, 25, 461–473.

Smith, S. K., & Sincich, T. (1992). Evaluating the forecast accuracy and bias of alternative

  population projections for states. *International Journal of Forecasting*, 8, 495–508.

Smith, S. K., & Tayman, J. (2003). An evaluation of population projections by age.

  Demography, 40, 741–757.

Smith, S. K., Tayman, J., & Swanson, D. A. (2001). *State and local population projections:*

  Methodology and analysis. New York: Kluwer Academic/Plenum Publishers.

Steinitz, C., Arias, H., Bassett, S., Flaxman, M., Goode, T., Maddock, T., Mouat, D., Peiser, R.,

  & Shearer, A. (2003). *Alternative futures for changing landscapes. The Upper San Pedro*

  River Basin in Arizona and Sonora. Washington DC: Island Press.

Stoto, M. A. (1983). The accuracy of population projections. *Journal of the American Statistical*

  Association, 78, 13–20.

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, 85-98.

Tayman, J., Parrott, B., & Carnevale, S. (1994). Locating fire station sites: The response time

  component. In H. J. Kintner, P. R. Voss, P. A. Morrison, & T. W. Merrick (Eds.), *Applied*

Demographics: A Casebook for Business and Government (pp. 203–217). Boulder, CO: Westview Press.

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

Tayman, J., Smith, S. K., & Lin, J. (2007). Precision, bias, and uncertainty for state population forecasts: An exploratory analysis of time series models. *Population Research and Policy Review*, 26, 347–369.

Texas Water Development Board. (1997). Water for Texas: A consensus-based update to the State Water Plan. Volume 2, Technical Planning Appendix, GF-6-2: Austin, TX.

Thomas, R. K. (1994). Using demographic analysis in health services planning: A case study in obstetrical services. In H. J. Kintner, P. R. Voss, P. A. Morrison, & T. W. Merrick (Eds.), *Applied Demographics: A Casebook for Business and Government* (pp. 159–179). Boulder, CO: Westview Press.

Webby, R., & O'Connor, M. (1996). Judgmental and statistical time series forecasting: A review of the literature. *International Journal of Forecasting*, 12, 91–118.

White, H. R. (1954). Empirical study of the accuracy of selected methods for projecting state populations. *Journal of the American Statistical Association*, 49, 480–498.

Williams, W. H., & Goodman, M. L. (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., & Bell, M. (2004). Australia's uncertain demographic future. *Demographic Research*, 11, 195–234.

Wilson T., & Bell, M. (2007). Probabilistic regional population forecasts: The example of

Queensland, Australia. *Geographical Analysis*, 39, 1–25.

Table 1. Target Years by Launch Year and Forecast Horizon

Forecast Horizon (Years)

Launch Year	10	20	30
1920	1930	1940	1950
1930	1940	1950	1960
1940	1950	1960	1970
1950	1960	1970	1980
1960	1970	1980	1990
1970	1980	1990	2000
1980	1990	2000	---
1990	2000	---	---

Table 2. Selected Measures of Forecast Error Distributions for U.S. Counties

Target Year	Horizon Length	Absolute % Error			Algebraic % Error	
		Mean	Std. Dev.	90th PE	Mean	Std. Dev.
1930	10	12.2	13.9	29.1	2.2	18.3
1940	10	11.2	12.4	23.3	0.4	16.7
1950	10	11.2	11.1	24.9	2.9	15.5
1960	10	10.3	9.9	23.2	0.3	14.3
1970	10	9.6	9.7	21.0	-2.4	13.4
1980	10	13.2	10.2	26.3	-9.5	13.7
1990	10	7.8	6.6	15.6	4.0	9.4
2000	10	6.2	6.0	13.9	-3.5	7.9
Average	10	10.2	10.0	22.2	-0.7	13.6
Std. Dev.	10	2.3	2.7	5.2	4.4	3.5
1940	20	20.2	22.8	47.2	5.9	29.9
1950	20	19.9	20.0	43.6	3.6	28.0
1960	20	23.0	21.3	50.9	6.4	30.7
1970	20	16.7	16.3	37.3	-0.5	23.3
1980	20	21.4	17.3	45.7	-12.1	24.7
1990	20	19.4	15.0	39.6	-9.3	22.7
2000	20	11.4	10.7	24.6	0.7	15.6
Average	20	18.9	17.6	41.3	-0.8	25.0
Std. Dev.	20	3.8	4.1	8.6	7.3	5.2
1950	30	33.1	37.4	78.7	14.0	48.0
1960	30	32.9	34.5	68.1	8.5	47.0
1970	30	31.9	31.1	68.3	9.0	43.6
1980	30	22.1	19.9	49.3	-9.8	28.1
1990	30	29.3	23.0	60.6	-11.9	35.3
2000	30	27.8	20.3	55.3	-14.5	31.2
Average	30	29.5	27.7	63.4	-0.8	38.9
Std. Dev.	30	4.2	7.6	10.5	12.6	8.5

Table 3. Percentage of APEs Less Than the Predicted 90th Percentile Error

Target Year	Horizon Length	Percentage
1940	10	93.0
1950	10	88.4
1960	10	90.9
1970	10	92.2
1980	10	80.1
1990	10	98.5
2000	10	92.4
Average	10	90.8
Std. Dev.	10	5.6
1950	20	91.5
1960	20	84.9
1970	20	95.3
1980	20	83.4
1990	20	94.0
2000	20	97.6
Average	20	91.1
Std. Dev.	20	5.7
1960	30	92.9
1970	30	89.9
1980	30	96.9
1990	30	82.0
2000	30	93.1
Average	30	91.0
Std. Dev.	30	5.6

Table 4. 90th Percentile Errors by Population Size

Target Year	Horizon Length	< 5,000	5,000 − 15,000	15,000 − 50,000	> 50,000	All
1930	10	59.9	30.1	22.8	21.9	29.1
1940	10	59.0	28.8	17.0	18.4	23.3
1950	10	45.7	26.0	22.8	22.3	24.9
1960	10	31.8	23.5	21.5	26.7	23.2
1970	10	34.8	23.4	18.7	17.4	21.0
1980	10	36.8	28.7	24.1	19.0	26.3
1990	10	19.0	16.1	14.9	15.4	15.6
2000	10	21.0	14.7	12.3	12.5	13.9
Average	10	38.5	23.9	19.3	19.2	22.2
Std. Dev.	10	15.5	5.8	4.2	4.4	5.2
1940	20	86.7	57.0	35.9	35.0	47.2
1950	20	76.7	50.2	35.4	31.6	43.6
1960	20	80.8	56.1	48.1	43.2	50.9
1970	20	51.5	35.5	35.5	37.7	37.3
1980	20	66.8	52.6	38.8	32.0	45.7
1990	20	53.2	41.9	34.6	34.6	39.6
2000	20	32.3	25.9	21.8	23.8	24.6
Average	20	64.0	45.6	35.7	34.0	41.3
Std. Dev.	20	19.4	11.7	7.7	6.0	8.6
1950	30	124.7	96.5	61.9	54.6	78.7
1960	30	100.3	80.7	59.4	50.5	68.1
1970	30	115.1	73.9	64.0	55.5	68.3
1980	30	68.9	50.5	46.1	42.8	49.3
1990	30	78.1	67.4	51.7	49.3	60.6
2000	30	71.5	58.5	47.9	49.2	55.3
Average	30	93.1	71.3	55.2	50.3	63.4
Std. Dev.	30	23.7	16.4	7.6	4.6	10.5

Table 5. Percentage of APEs Less Than the Predicted 90th Percentile Error by Population Size

Target Year	Horizon Length	< 5,000	5,000 − 15,000	15,000 − 50,000	> 50,000	All
1940	10	90.5	91.4	96.3	94.1	93.0
1950	10	95.0	92.2	82.8	86.5	88.4
1960	10	96.1	91.1	91.2	87.6	90.9
1970	10	87.7	89.7	93.2	98.1	92.2
1980	10	85.8	78.2	80.3	86.5	80.1
1990	10	97.6	98.4	99.0	96.2	98.5
2000	10	86.8	91.1	93.8	94.7	92.4
Average	10	91.3	90.3	90.9	92.0	90.8
Std. Dev.	10	4.8	6.0	6.9	4.9	5.6
1950	20	93.0	93.4	90.9	94.8	91.5
1960	20	88.7	87.2	80.8	81.0	84.9
1970	20	98.7	97.5	95.4	92.1	95.3
1980	20	78.4	72.2	87.2	95.0	83.4
1990	20	97.4	96.3	93.4	86.9	94.0
2000	20	97.6	97.9	97.3	97.5	97.6
Average	20	92.3	90.7	90.8	91.2	91.1
Std. Dev.	20	7.8	9.9	6.0	6.2	5.7
1960	30	88.4	94.6	91.1	93.0	92.9
1970	30	90.6	91.9	88.0	88.0	89.9
1980	30	98.0	98.4	96.9	95.0	96.9
1990	30	80.7	74.5	86.9	85.0	82.0
2000	30	94.7	95.7	91.9	90.6	93.1
Average	30	90.5	91.0	91.0	90.3	91.0
Std. Dev.	30	6.6	9.5	3.9	4.0	5.6

Table 6. 90th Percentile Errors by Population Growth Rate

| Target Year	Horizon Length	< -10%	-10% to 10%	10% to 25%	> 25%	All
1930	10	35.6	17.3	27.2	44.5	29.1
1940	10	40.7	16.7	21.6	54.1	23.3
1950	10	19.6	21.6	27.6	37.8	24.9
1960	10	16.1	19.2	29.7	37.9	23.2
1970	10	23.9	17.2	17.5	31.2	21.0
1980	10	31.8	24.1	23.3	30.8	26.3
1990	10	12.4	13.5	16.1	20.6	15.6
2000	10	17.3	10.7	13.6	22.6	13.9
Average	10	24.7	17.5	22.1	34.9	22.2
Std. Dev.	10	10.2	4.3	5.9	11.1	5.2
1940	20	69.7	25.4	44.8	86.7	47.2
1950	20	61.0	32.2	51.0	81.5	43.6
1960	20	44.9	45.4	57.8	74.0	50.9
1970	20	30.1	30.5	51.1	60.2	37.3
1980	20	56.8	35.2	34.7	50.5	45.7
1990	20	47.7	32.8	38.6	50.9	39.6
2000	20	27.9	19.5	25.5	37.4	24.6
Average	20	48.3	31.6	43.4	63.0	41.3
Std. Dev.	20	15.5	8.1	11.1	18.2	8.6
1950	30	71.9	43.5	82.6	145.0	78.7
1960	30	71.2	54.3	87.8	139.6	68.1
1970	30	61.6	60.6	82.6	136.6	68.3
1980	30	56.1	40.4	55.8	68.2	49.3
1990	30	72.2	46.5	53.0	70.1	60.6
2000	30	66.0	46.5	52.9	69.9	55.3
Average	30	66.5	48.6	69.1	104.9	63.4
Std. Dev.	30	6.6	7.5	16.8	39.0	10.5

Table 7. Percentage of APEs Less Than the Predicted 90th Percentile Error by Population Growth Rate

Target Year	Horizon Length	< -10%	-10% to 10%	10% to 25%	> 25%	All
1940	10	84.1	91.1	96.1	91.5	93.0
1950	10	93.8	82.9	85.0	98.6	88.4
1960	10	84.7	93.0	88.4	88.2	90.9
1970	10	75.3	90.5	95.8	93.0	92.2
1980	10	72.4	76.4	83.9	95.6	80.1
1990	10	99.5	98.6	97.9	97.0	98.5
2000	10	78.7	93.2	93.8	89.1	92.4
Average	10	84.1	89.4	91.6	93.3	90.8
Std. Dev.	10	9.8	7.4	5.7	4.0	5.6
1950	20	93.7	84.7	93.6	93.1	91.5
1960	20	92.1	79.1	88.3	94.1	84.9
1970	20	91.0	96.7	95.5	91.5	95.3
1980	20	52.5	78.5	97.0	94.2	83.4
1990	20	95.7	91.7	90.1	93.2	94.0
2000	20	98.1	97.2	97.7	97.7	97.6
Average	20	87.2	88.0	93.7	94.0	91.1
Std. Dev.	20	17.2	8.4	3.8	2.1	5.7
1960	30	89.7	85.4	95.4	92.7	92.9
1970	30	92.7	86.1	95.4	93.8	89.9
1980	30	91.0	96.8	98.9	98.6	96.9
1990	30	67.8	77.2	90.3	89.5	82.0
2000	30	95.7	89.8	94.1	92.9	93.1
Average	30	87.4	87.1	94.8	93.5	91.0
Std. Dev.	30	11.2	7.1	3.1	3.3	5.6

Table 8. Percentage of APEs Less Than the Predicted 90th Percentile Error by Forecasting Technique

Target Year	Horizon Length	LIN	MLN	SHR	SFT	EXP	COS	CON	AV7	AV5
1940	10	92.6	92.7	92.4	90.9	93.2	95.0	96.5	92.9	93.0
1950	10	89.6	89.6	90.8	94.4	92.3	79.2	79.1	87.1	88.4
1960	10	91.3	91.5	91.4	91.2	92.0	89.5	88.2	90.9	90.9
1970	10	89.9	90.5	87.8	77.1	91.0	91.8	94.7	93.2	92.2
1980	10	81.7	82.3	84.1	88.6	81.1	96.5	86.3	80.8	80.1
1990	10	98.7	98.6	98.8	99.2	97.1	87.2	96.0	98.4	98.5
2000	10	91.8	92.9	93.2	92.3	94.8	95.8	87.2	91.9	92.4
Average	10	90.8	91.2	91.2	90.5	91.6	90.7	89.7	90.8	90.8
Std. Dev.	10	5.0	4.9	4.6	6.8	5.1	6.1	6.4	5.5	5.6
1950	20	90.7	90.8	91.0	91.1	94.7	83.2	89.0	91.6	91.5
1960	20	86.3	86.1	87.6	92.7	89.6	84.9	78.3	85.0	84.9
1970	20	94.9	95.6	95.1	92.8	95.5	94.8	93.0	95.9	95.3
1980	20	79.8	81.3	78.0	66.6	81.2	93.3	92.1	84.2	83.4
1990	20	93.9	93.6	94.2	94.7	92.7	91.1	91.3	93.8	94.0
2000	20	97.9	97.7	98.4	99.1	95.7	93.9	93.3	97.5	97.6
Average	20	90.6	90.8	90.7	89.5	91.6	90.2	89.5	91.3	91.1
Std. Dev.	20	6.6	6.2	7.2	11.6	5.6	4.9	5.7	5.6	5.7
1960	30	91.5	92.1	92.1	88.9	95.0	84.9	85.6	92.9	92.9
1970	30	90.3	90.9	92.0	93.5	91.7	90.9	85.9	90.0	89.9
1980	30	95.5	96.3	95.2	90.0	96.8	97.3	94.4	97.3	96.9
1990	30	80.5	81.3	77.2	65.8	79.5	83.7	90.7	82.8	82.0
2000	30	94.2	93.8	94.4	83.0	93.7	95.5	90.6	93.3	93.1
Average	30	90.4	90.9	90.2	84.2	91.3	90.5	89.4	91.3	91.0
Std. Dev.	30	5.9	5.7	7.4	11.0	6.9	6.1	3.7	5.4	5.6

Figure 1. Distribution of Algebraic Percent Errors by Target Year, 20 Year Horizon.
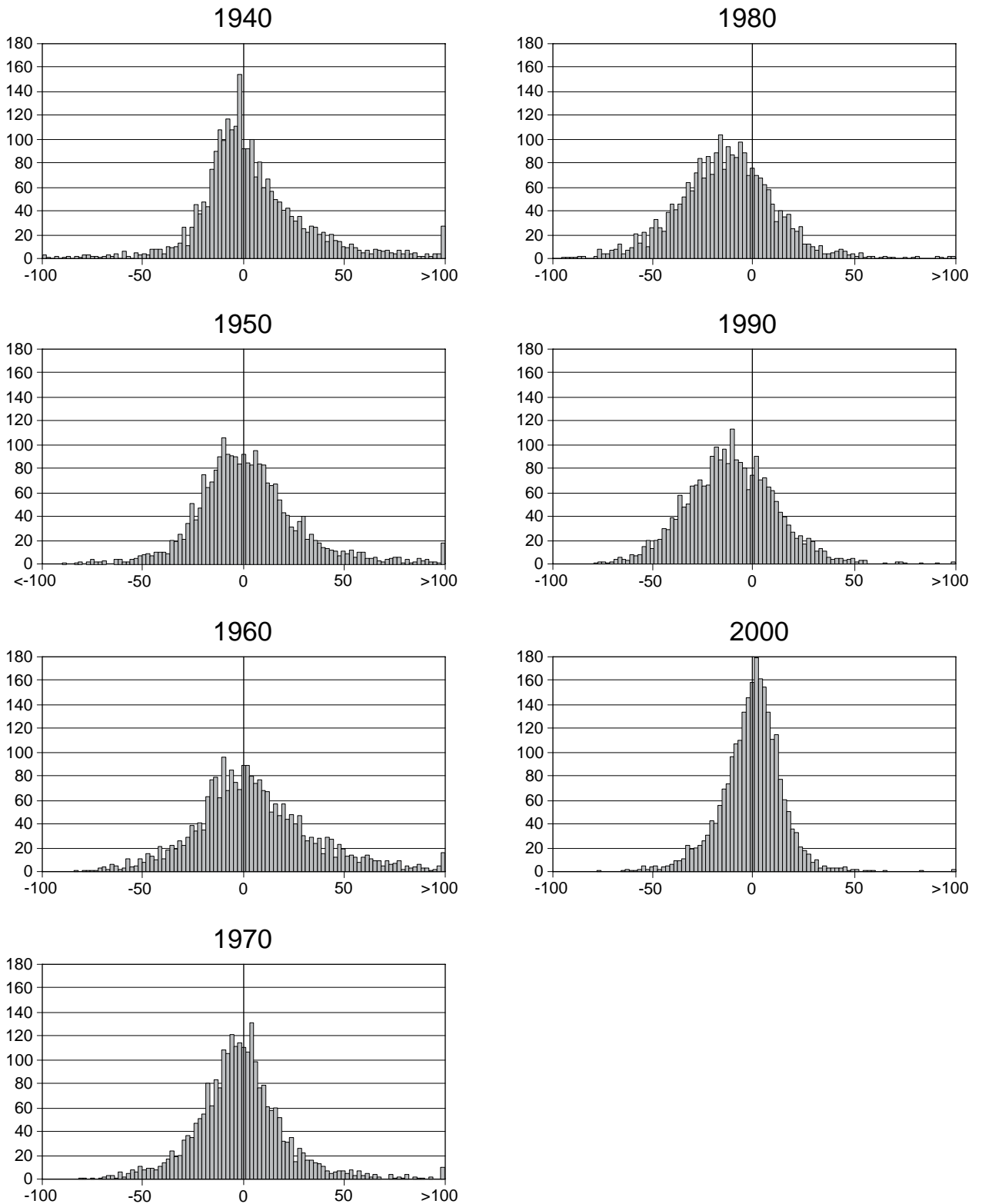
Note: Number of counties on vertical axis, algebraic percent errors on horizontal axis.

Figure 2. Distribution of Absolute Percent Errors by Target Year, 20 Year Horizon.
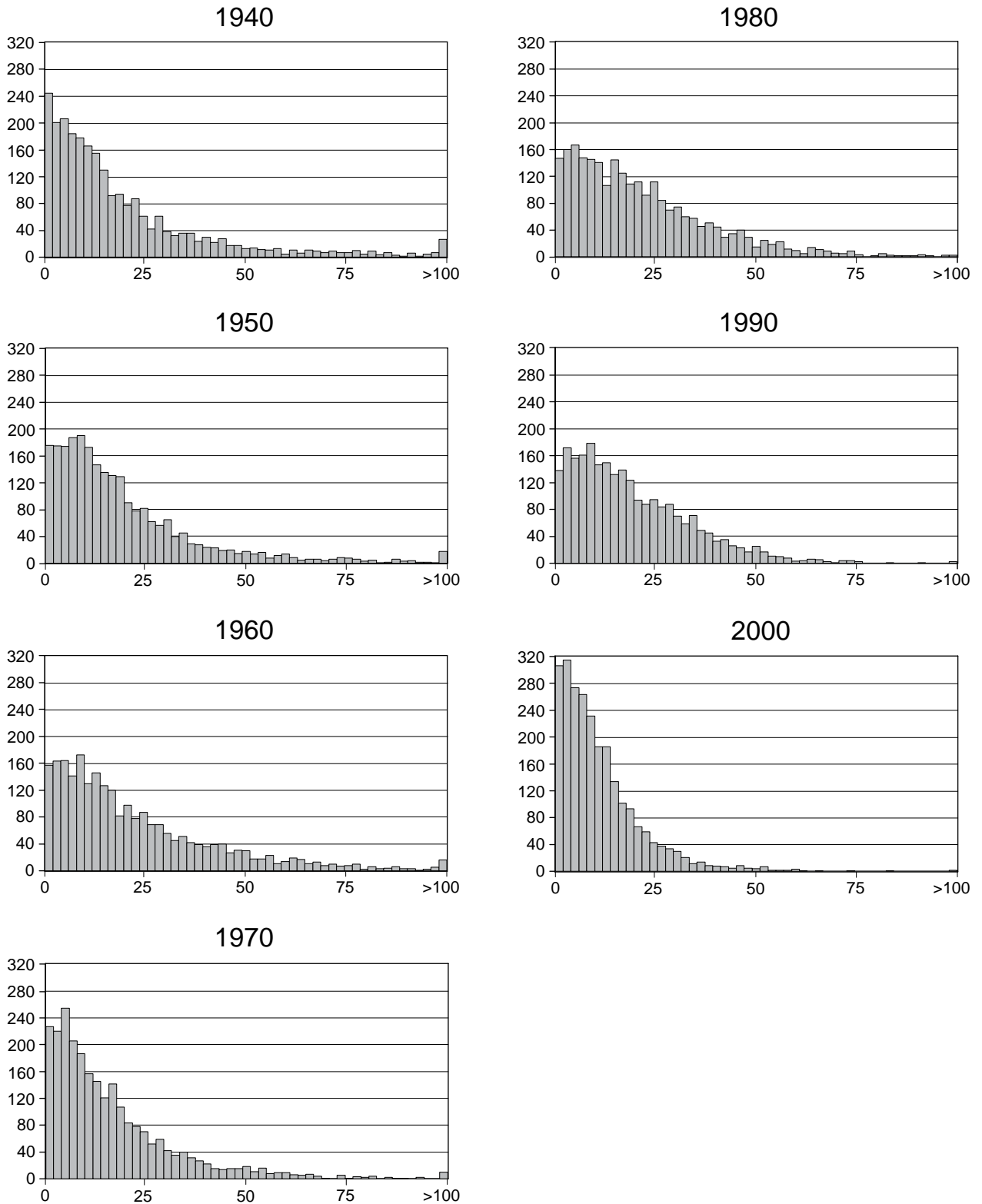
Note: Number of counties on vertical axis, absolute percent errors on horizontal axis.