Tests of Forecast Accuracy and Bias for County Population Projections

STANLEY K. SMITH*

This article deals with the forecast accuracy and bias of population projections for 2,971 counties in the United States. It uses three different projection techniques and data from 1950, 1960, 1970, and 1980 to make two sets of 10-year projections and one set of 20-year projections. These projections are compared with census counts to determine forecast errors. The size, direction, and distribution of forecast errors are analyzed by size of place, rate of growth, and length of projection horizon. A number of consistent patterns are noted, and an extension of the empirical results to the production of confidence intervals for population projections is considered.

KEY WORDS: Demographic projections and forecasts; Small area projections and forecasts; Projection techniques; Characteristics of population projections; Confidence intervals.

1. INTRODUCTION

Many decisions are determined by expectations of future population change. Planning for schools, hospitals, shopping centers, housing developments, electric power plants, and many other projects is strongly affected by expected population growth or decline. Indeed, the eventual success or failure of such plans often depends on whether expected population changes are realized over time. The distribution of funds for government programs and the granting of licenses and permits by various regulatory agencies are often determined by population projections as well. It is not surprising that population projections and forecasts are of so much interest to so many people.

What is surprising is that relatively little research has focused on the forecast accuracy of past population projections (Keyfitz 1981, p. 580). The research that has been done has frequently dealt with a small sample size and/or a single time period. Consequently the basis for distinguishing patterns in the distribution of forecast errors and judging the stability of those patterns over time has been limited. This article deals with these problems by using a very large data set and evaluating projections for two different 10-year time horizons and one 20-year horizon. The empirical results confirm some previously reported findings and present some new evidence on the forecast accuracy and bias of population projections.

Demographers often distinguish between the terms "projection" and "forecast." A projection is typically defined as the numerical outcome of a specific set of assumptions regarding future trends, whereas a forecast indicates the specific projection that the author believes is most likely to provide an accurate prediction of future population (e.g., Irwin 1977; Isserman and Fisher 1984; Keyfitz 1972; Morrison 1977). In this article the term "projection" refers to the future population implied by a particular technique and data set, and "forecast accuracy" refers to the percent difference between a projection and the census-enumerated population for the same year. In other words, projections are evaluated as if they were indeed forecasts of future population.

The focus of this article is the forecast accuracy and bias of population projections for countries. What are the average forecast errors for different projection techniques? What does the distribution of errors look like? Do some techniques provide more accurate forecasts than others? Is there a tendency for projections to be too high or too low? What factors influence the size, direction, and distribution of forecast errors? Can the study of past forecast errors help us predict future forecast errors? These are some of the questions considered in this article.

2. DATA AND TECHNIQUES

The data used in this study were taken from census counts for 1950, 1960, 1970, and 1980 for 2,971 counties in the United States. (For simplicity, parishes and other county equivalents are referred to as counties.) This sample included all counties for which comparable data were available in all four censuses. Excluded from the sample were Alaska and Hawaii, for which county data in 1950 and 1960 were not available on the computer tapes; Virginia, in which numerous changes in county boundaries occurred between 1950 and 1980; and five other counties in which consolidations or other problems made comparable data impossible to obtain. Those five counties were Menominee in Wisconsin, Armstrong in South Dakota, Carson and Ormsby in Nevada, and Yellowstone in Montana. All other counties in the United States were included in the sample. The data represent total population only; no analysis was performed on the age, sex, or race distribution of the population.

There are some problems with these data. As with all census data, there exists the possibility of errors caused by overcount and undercount. In addition, a number of counties experienced consolidations or boundary changes between 1950 and 1980. There is typically no documentation regarding the magnitude of these changes. Fortunately, changes have occurred only in a small proportion of all counties and their effects in terms of population have generally been very small. Although these changes undoubtedly have affected forecast errors in some individual

* Stanley K. Smith is Associate Professor of Economics and Associate Director, Bureau of Economic and Business Research, College of Business Administration, University of Florida, Gainesville, FL 32611. He thanks Michael Grant, Faith Sincich, and Ravi Bayya for excellent research support; Pam Middleton for skillful typing; and Andrew Isserman, the editors, and the anonymous referees for many helpful suggestions. An earlier version of this article was presented at the annual meeting of the Population Association of America, San Francisco, April 2–5, 1986. The full array of data underlying the figures and tables shown in this article is available from the author upon request.

counties, their impact on the results shown in this study is believed to be negligible.

Three primary projection techniques were used. First was linear extrapolation (LINE), which assumes that a population will increase (decrease) by the same number of persons in each future year as the average annual increase (decrease) during the base period:

$$P_f = P_b + x/y(P_b - P_a), \quad (1)$$

where P_f = county population projection for future year f, P_b = county population at end of base period (year b), P_a = county population at beginning of base period (year a), x = number of years in projection horizon (f–b), and y = number of years in base period (b–a).

The second technique was exponential extrapolation (EXPO), which assumes that a population will grow (decline) at the same rate in each future year as it grew (declined) per year during the base period:

$$P_f = P_b \exp(rx), \quad (2)$$

where r = average annual growth rate during base period.

In the third technique (SHIFT), county population data are expressed as shares of a state population for which a projection already exists. These shares were obtained from historical data and are extrapolated into the future by assuming that the average annual absolute change in the county's share of state population observed during the base period will continue into the future. The extrapolated county shares are then applied to an independent projection of state population to provide county population projections. The state projections used in this study were taken from Census Bureau publications. Projections based on 1950–1960 data used series I–D from the U.S. Bureau of the Census (1965), and projections based on 1960–1970 data used series I–E from the U.S. Bureau of the Census (1972). Mathematically, the SHIFT technique is defined as

$$P_f = P_{jf}\left[\frac{P_b}{P_{jb}} + x/y\left(\frac{P_b}{P_{jb}} - \frac{P_a}{P_{ja}}\right)\right], \quad (3)$$

where P_{jf} = state population projection for year f, P_{jb} = state population at end of base period (year b), and P_{ja} = state population at beginning of base period (year a).

Simple extrapolation techniques such as these are not held in high regard by many demographers (e.g., Birch 1977; Irwin 1977; Morrison 1977; Pittenger 1980; Schmitt and Crosetti 1952). They are widely viewed as simplistic, naive, and less accurate than more sophisticated techniques. They provide no information on the composition of the population (e.g., age, sex, race) and have no behavioral or theoretical content. Consequently, for many of the purposes for which population projections may be used—such as providing demographic characteristics, analyzing the effects of changes in specific components of population growth, or evaluating the demographic consequences of different economic scenarios—these simple techniques are *not* very useful.

If the only purpose of a projection is to forecast total population, however, these techniques can be very useful. A number of studies have found that simple extrapolation techniques produce short- to medium-term forecasts of total population that are at least as accurate as those produced by more sophisticated techniques (e.g., Ascher 1978; Greenberg 1972; Hajnal 1955; Kale, Voss, Palit, and Krebs 1981; Siegel 1972; Smith 1984; Stoto 1983). I know of no study showing more sophisticated techniques to consistently produce more accurate population projections than these simple techniques. Although future research may some day alter this conclusion, at the present time the projections coming from simple extrapolation techniques must be accepted as being at least as accurate as those coming from more sophisticated techniques.

Furthermore, minimal data requirements and ease of application make these techniques quite useful for projections of small areas, where the data required by more sophisticated techniques are often unavailable or outdated. Although seldom used at the state and national level, these and similar techniques are still occasionally used at the county level and very frequently used at the subcounty level (e.g., Federal–State Cooperative Program for Population Projections 1984; Irwin 1977). An additional advantage for the present study is that they can be applied retrospectively to a large number of counties to provide a set of projections based on identical assumptions. This avoids the methodological problem of comparing errors from a number of different projections in which the underlying assumptions differ or are unknown (e.g., Keyfitz 1981; Stoto 1983).

It is not my contention, of course, that the techniques used in this study are "better" than any other projection techniques. A number of other simple extrapolation techniques could have been used (e.g., Greenberg 1972; Isserman 1977; Voss and Kale 1985). Including additional simple techniques, however, would have expanded on already lengthy analysis and most likely would not have altered any of the main conclusions. I believe the techniques included in this study are satisfactory representatives of the general group of simple extrapolation techniques.

It also would have been desirable to include several more sophisticated techniques, such as cohort–component or economic-based. Unfortunately, no projections from these techniques have been produced for all counties in the United States for the time period covered by this study. Creating such a set specifically for this study would have been prohibitively expensive. Rather than limit the study to a smaller number of counties or a single time period, it was decided to leave cohort–component and economic-based techniques out of the analysis.

3. TESTS OF ACCURACY AND BIAS

Using data from 1950, 1960, and 1970, the techniques described previously were used to make 10- and 20-year population projections for 1970 and 1980 for each of the 2,971 counties in the sample. These projections were then compared with census counts for 1970 and 1980. The resulting differences are called forecast errors, although they

might have been caused by errors in census enumeration as well as errors in the forecasts themselves.

Two measures of forecast accuracy and bias are reported in this article. Mean absolute percentage error (MAPE) is the average percentage error when the direction of the error is ignored. This provides a measure of accuracy. Mean algebraic percentage error (MALPE) is the average percentage error when the direction of error is accounted for. This provides a measure of bias: a positive error indicates a tendency for projections to be too high and a negative error indicates a tendency for projections to be too low. Throughout this article the term "error" refers to percentage error rather than numerical error.

Three other measures of forecast accuracy and bias were used in earlier stages of this research: root mean square percentage error, proportion of errors greater than 25%, and proportion of negative errors. The results from these three measures are not shown here but led to the same conclusions as those reported in this article (Smith 1986a).

3.1 Size of Place

The first set of projections used 1960 and 1970 data to project 1980 population. Table 1 shows the number of counties in each size category, and Figure 1 summarizes the forecast errors. The forecast accuracy for the entire sample was roughly the same for all three techniques, as MAPE's were 13.7% for LINE, 13.1% for EXPO, and 15.7% for SHIFT. There were some differences for counties of different sizes, however. For small counties, SHIFT had the largest errors and EXPO had the smallest. For large counties, EXPO had the largest errors and LINE had the smallest. These differences were generally quite small, especially for large counties.

For each technique there was a strong negative relationship between size of place and size of error: the larger the place, the smaller the error. MAPE's were about twice as large for counties with fewer than 5,000 population than for counties with 100,000 or more. A negative relationship between size of place and size of error is a common empirical finding for population projections (e.g., Irwin 1977; Isserman 1977; Smith 1984; White 1954). This relationship became fairly weak for counties of 25,000 or more, however. This too is consistent with several earlier studies showing no consistent relationship between size of place and size of error among large places (e.g., Schmitt and Crosetti 1951; Smith 1984).

Table 1. Number of Counties, by Population Size at End of Base Period

Population size	End of base period 1960	1970
<5,000	272	302
5,000–14,999	965	918
15,000–24,999	599	555
25,000–49,999	560	539
50,000–99,999	281	324
100,000+	294	333
Total	2,971	2,971

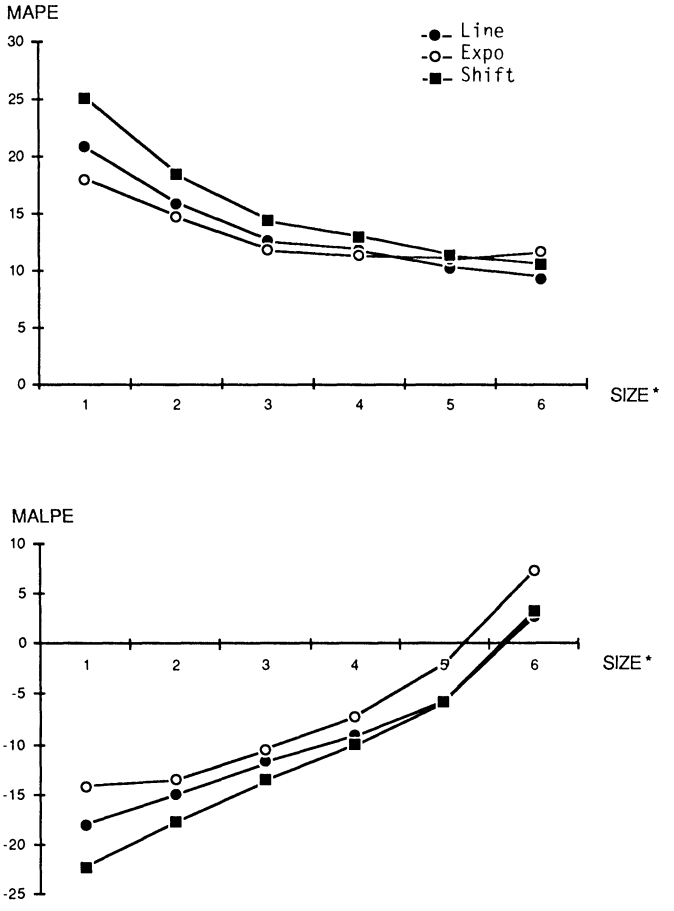

*Figure 1. MAPE and MALPE for 10-Year Projections to 1980, by Size of County Population in 1970. —●—, LINE; —○—, EXPO; —■—, SHIFT. * Size categories: (1) <5,000; (2) 5,000–14,999; (3) 15,000–24,999; (4) 25,000–49,999; (5) 50,000–99,999; (6) 100,000+.*

All three techniques showed a strong overall downward bias in their projections, as indicated by MALPE's of −10.6% for LINE, −8.3% for EXPO, and −12.3% for SHIFT. This tendency to produce low projections was strongly related to size of place: for all three techniques the MALPE became less strongly negative as population size increased. For counties with a population of 100,000 or more, positive errors actually outnumbered and outweighed negative errors. As will be shown later in this article, however, the relationship between bias and size of place appears to be spurious, being caused by differences in rates of population growth.

Were the patterns noted in Figure 1 simply the result of the particular historical context for which the projections were made, or would they be repeated if another time period were used? To answer this question a second set of projections was made, using 1950 and 1960 data to project population in 1970. These results are summarized in Figure 2. Although there were some minor differences, most of the patterns found in the first set of projections were also found in the second. The MAPE's for the entire sample were 11.6% for LINE, 12.9% for EXPO, and 15.8% for SHIFT; these errors were very similar to those

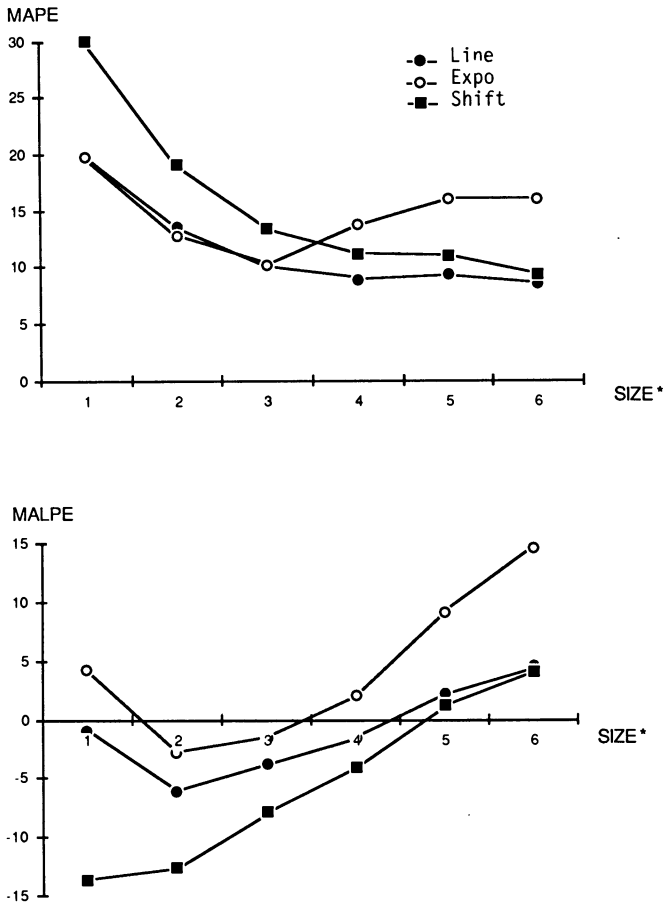

*Figure 2. MAPE and MALPE for 10-Year Projections to 1970, by Size of County Population in 1960. —●—, LINE; —○—, EXPO; —■—, SHIFT. * Size categories: (1) <5,000; (2) 5,000–14,999; (3) 15,000–24,999; (4) 25,000–49,999; (5) 50,000–99,999; (6) 100,000+.*

found in the first set of projections. Errors again declined as population size increased through the three smallest size categories, but then leveled off through the three largest categories (actually increasing for EXPO).

The only major difference between the two sets of projections was in the direction of error. Whereas the first set exhibited a strong overall downward bias for all three techniques, the second set exhibited much less downward bias (EXPO actually had a small upward bias). The MALPE's for the entire sample were −2.5% for LINE, 1.9% for EXPO, and −7.2% for SHIFT. A positive relationship between MALPE and size of place was found again, except for LINE and EXPO in the smallest two size categories. On the basis of this evidence it appears that—with the exception of overall bias—the characteristics noted in Figure 1 were not caused simply by a unique historical context.

It should perhaps be noted that the SHIFT technique contains two sources of error: one caused by errors in projecting the county's share of state population, and the other caused by errors in the state projections themselves. The latter source of error is minor. An alternative set of projections was made using actual state populations in 1970 and 1980 instead of projected state populations. The results from those projections differed only slightly from those reported in this article. State projections rather than census data were used in the present analysis because when the SHIFT technique is actually used for county projections, it must use projections of state population rather than census data.

To examine the effects of length of projection horizon on forecast errors, 20-year projections were made using 1950 and 1960 data to project 1980 population. These results are summarized in Figure 3. The relationships between errors and size of place were about the same as those reported earlier: MAPE's declined (except for EXPO) and MALPE's increased as county size increased. The magnitude of the errors, however, was much greater for the 20-year projections. The MAPE's for the 20-year projections were roughly twice as large as for the 10-year projections. For the entire sample, MAPE's were 25.9% for LINE, 30.8% for EXPO, and 35.4% for SHIFT. A strong positive relationship between length of projection horizon and size of error is a common empirical finding (e.g., Irwin 1977; Isserman 1977; Schmitt and Crosetti 1951, 1953; Siegel 1953; Smith 1984; Stoto 1983; White 1954).

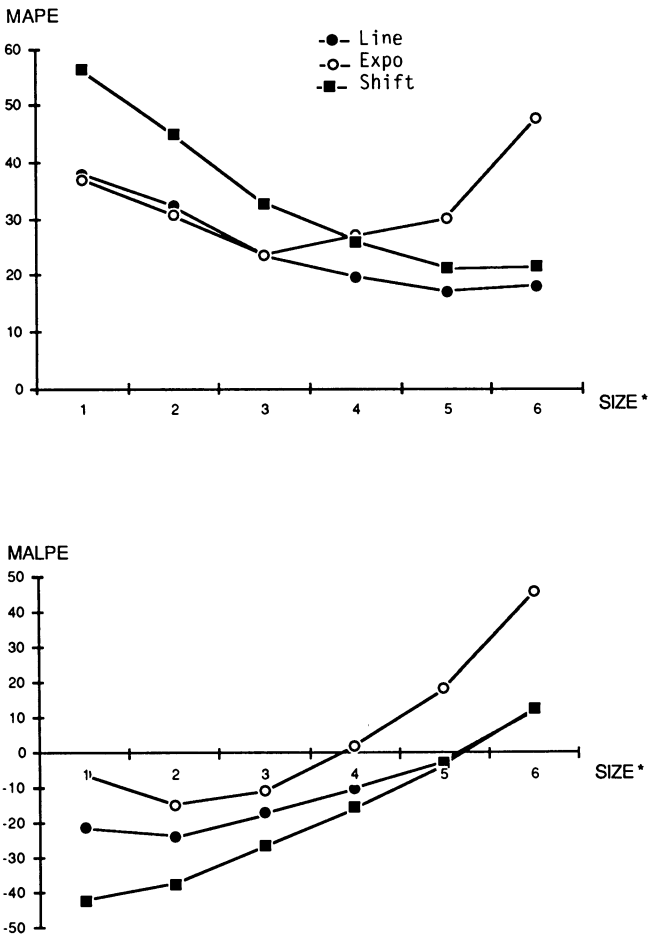

*Figure 3. MAPE and MALPE for 20-Year Projections to 1980, by Size of County Population in 1960. —●—, LINE; —○—, EXPO; —■—, SHIFT. * Size categories: (1) <5,000; (2) 5,000–14,999; (3) 15,000–24,999; (4) 25,000–49,999; (5) 50,000–99,999; (6) 100,000+.*

3.2 Rate of Growth

Several researchers have noted that forecast errors tend to be larger for rapidly growing places than slowly growing places (e.g., Isserman 1977; Keyfitz 1981; Schmitt and Crosetti 1951). Some of these studies, however, focused on the relationship between forecast errors and the rate of growth over the projection horizon. Since the rate of growth over the projection horizon is unknown at the time a projection is made, this information is not useful for predicting forecast errors. More useful information would come from studying the relationship between forecast errors and the rate of growth during the base period. Figures 4–6 summarize the results of such an analysis for projections covering the three time periods described earlier.

All three figures show a U-shaped relationship between rate of growth and size of forecast error. MAPE's were relatively large for counties that lost population during the base period, became smaller for counties with slow but positive growth rates, and then became steadily larger as growth rates increased. This relationship was found for all projection techniques in all three sets of projections. For the EXPO technique, errors were particularly large for counties with growth rates of 100% or more. Table 2 shows the number of counties in each growth rate category.

Substantial differences in forecast accuracy were found

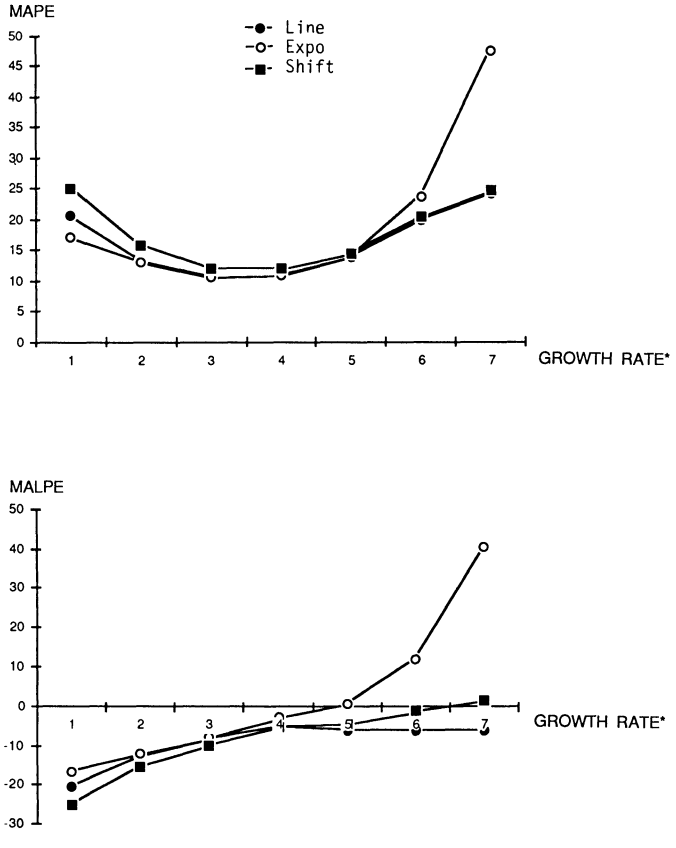

Figure 4. MAPE and MALPE for 10-Year Projections to 1980, by Rate of Population Growth, 1960–1970. —●—, LINE; —○—, EXPO; —■—, SHIFT.

** Growth rate categories: (1) < −10%; (2) −10%–0%; (3) 0%–10%; (4) 10%–25%; (5) 25%–50%; (6) 50%–100%; (7) 100%+.*

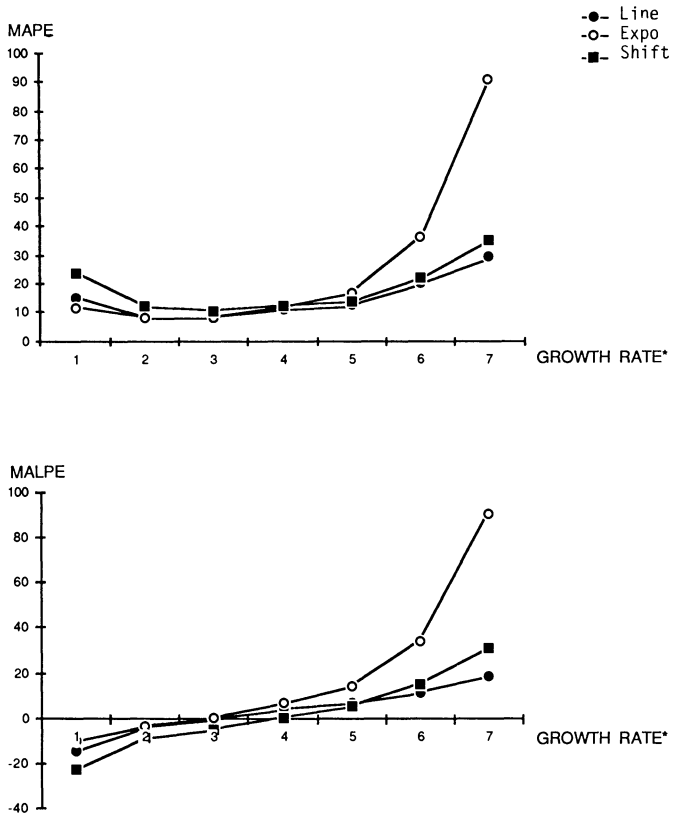

Figure 5. MAPE and MALPE for 10-Year Projections to 1970, by Rate of Population Growth, 1950–1960. —●—, LINE; —○—, EXPO; —■—, SHIFT.

** Growth rate categories: (1) < −10%; (2) −10%–0%; (3) 0%–10%; (4) 10%–25%; (5) 25%–50%; (6) 50%–100%; (7) 100%+.*

for different techniques in the various growth rate categories. For counties that lost population during the base period, EXPO had the smallest errors and SHIFT had the largest. For moderately growing counties, all three techniques had very similar errors. For the most rapidly growing counties, errors were smallest for LINE, somewhat larger for SHIFT, and by far the largest for EXPO. An explanation for these findings will be given in Section 4.

A distinct pattern can also be seen with respect to bias. For all techniques and all sets of projections, higher growth rates were associated with larger MALPE's. Projections for counties that lost population during the base period had a strong downward bias, indicating that those counties generally lost population less rapidly during the projection horizon than during the base period. Projections for counties that grew very rapidly during the base period had a strong upward bias, indicating that those counties generally grew less rapidly during the projection horizon than during the base period.

This finding suggests that county growth rates may have a tendency to regress toward the mean over time. To investigate this further, decade growth rates for each county were compared with the growth rates during the following decade. Two sets of comparisons were made, one comparing 1960–1970 rates with 1970–1980 rates and one com-

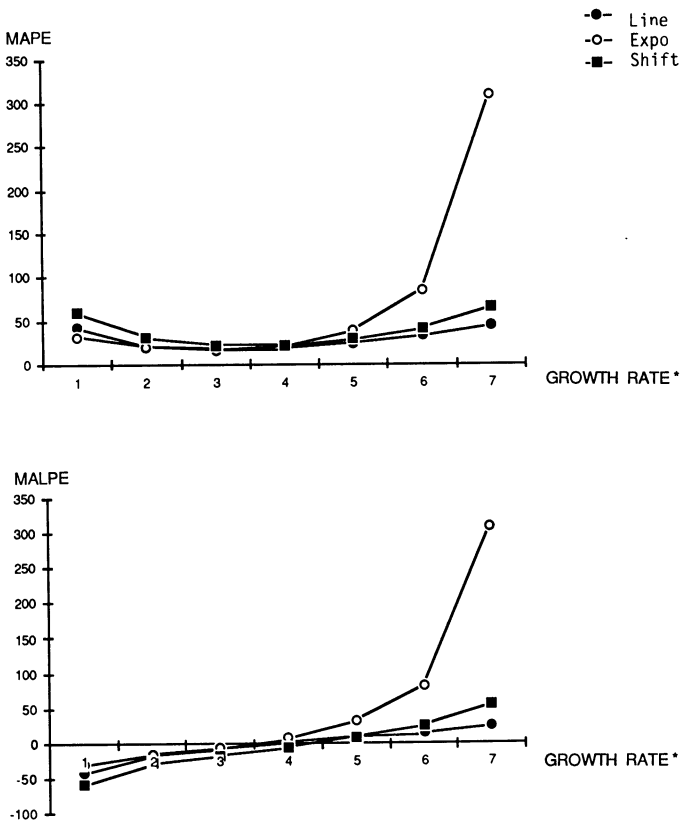

Figure 6. MAPE and MALPE for 20-Year Projections to 1980, by Rate of Population Growth, 1950–1960. —●—, LINE; —○—, EXPO; —■—, SHIFT.

** Growth rate categories: (1) < −10%; (2) −10%–0%; (3) 0%–10%; (4) 10%–25%; (5) 25%–50%; (6) 50%–100%; (7) 100%+.*

paring 1950–1960 rates with 1960–1970 rates. The results of these two sets were aggregated and are shown in Table 3. It is clear that extreme growth rates did tend to moderate over time: a large majority of the counties that declined or grew very slowly during one decade grew more rapidly the following decade, whereas a large majority of the counties that grew very rapidly during one decade grew more slowly the following decade. In this sample, then, growth rates considerably above or below average generally did *not* continue at such extreme levels over time. The implications of this finding for population projections will be discussed in Section 4.

Table 2. Number of Counties, by Rate of Population Growth During Base Period

Rate of growth	Base period 1950–1960	Base period 1960–1970
<−10%	755	516
−10%–0%	732	800
0%–10%	588	766
10%–25%	473	558
25%–50%	242	243
50%–100%	138	75
100%+	43	13
Total	2,971	2,971

Table 3. Comparison of Decade Growth Rates, 1950–1980

Growth rate, decade t	N	Number of counties with higher growth rate in decade t + 1 than in decade t
<−10%	1,271	1,125 (88.5)
−10%–0%	1,532	1,221 (79.7)
0%–10%	1,354	883 (65.2)
10%–25%	1,031	487 (47.2)
25%–50%	485	157 (32.4)
50%–100%	213	42 (19.7)
100%+	56	3 (5.4)
Total	5,942	3,918 (65.9)

NOTE: *N* = number of counties. Values in parentheses are percentages.

3.3 Size and Growth Rate

Figures 1–6 have shown some very strong relationships between forecast errors and size of place and rate of growth. Is it possible that some of these relationships were spurious, caused by a correlation between size of place and rate of growth? To answer this question, forecast errors must be considered for counties classified by size of place and rate of growth concurrently. Counties were thus divided into 16 groups based on four population-size categories and four rate-of-growth categories. The number of counties in each group is shown in Table 4. The projections from LINE, EXPO, and SHIFT were averaged, giving a fourth projection technique (AVE). Only the results from AVE are shown in this article; the results from the other three techniques were similar to those shown here (Smith 1986a).

The results from the two sets of 10-year projections for AVE are summarized in Figures 7 and 8. It is clear from these figures that size of place and rate of growth each had a substantial independent impact on forecast accuracy. For both sets of projections, there was a U-shaped relationship between MAPE and rate of growth within three

Table 4. Number of Counties, by Population Size at End of Base Period and Rate of Population Growth During Base Period

Group	Population size	Rate of growth	End of base period 1960	End of base period 1970
1	<5,000	<0%	206	241
2	<5,000	0%–10%	34	30
3	<5,000	10%–50%	25	25
4	<5,000	50+%	7	6
5	5,000–14,999	<0%	716	584
6	5,000–14,999	0%–10%	133	196
7	5,000–14,999	10%–50%	95	132
8	5,000–14,999	50+%	21	6
9	15,000–49,999	<0%	517	411
10	15,000–49,999	0%–10%	330	372
11	15,000–49,999	10%–50%	266	292
12	15,000–49,999	50+%	46	19
13	50,000+	<0%	48	80
14	50,000+	0%–10%	91	168
15	50,000+	10%–50%	329	352
16	50,000+	50+%	107	57

of the four size categories (Fig. 7). The only exception was for counties with less than 5,000 population, where the relationship was somewhat indistinct. (It should be noted that three of the four cells in this size category had a relatively small number of observations.) Within rate-of-growth categories (Fig. 8), there was a strong negative relationship between MAPE and population size for both sets of projections. The only exception was for the 1980 projections in counties growing by more than 50% between 1960 and 1970. (Again, the number of observations in three of the four cells was quite small.)

These general patterns for MAPE are the same as those reported earlier when size of place and rate of growth were considered independently. Several patterns, however, were stronger for some types of counties than for others. Differences in errors by growth rate were generally greater among small counties than large counties, and differences in errors by population size were generally greater for rapidly growing counties than slowly growing counties. The interaction between size of place and rate of growth clearly influenced forecast errors.

With respect to bias, Figures 7 and 8 show some similarities to the results reported earlier, but some differences as well. Within size categories there was generally a strong positive relationship between MALPE and the rate of growth. This is the same result as was reported in Figures 4–6. The only exceptions were the 1980 projections for the two middle size categories, where the MALPE declined rather than increased for the highest growth rate category. In both of these cases, the number of counties in the highest growth rate category was very small (see Table 4).

Within rate-of-growth categories, however, there was no consistent relationship between population size and MALPE. In some rate-of-growth categories the MALPE increased with population size, whereas in others it declined. This is quite different from the results shown in Figures 1–3. In this sample, then, population size had no

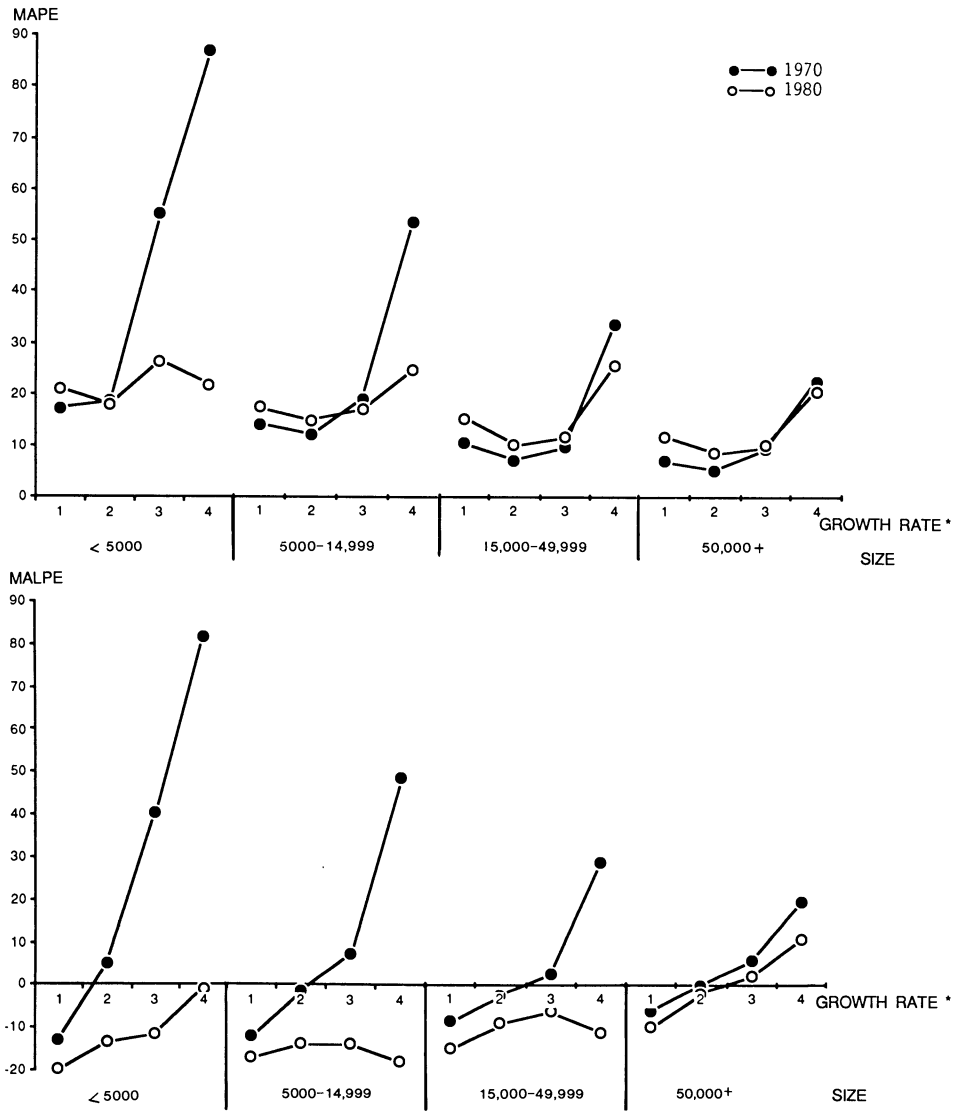

Figure 7. MAPE and MALPE for Two Sets of 10-Year Projections (AVE) by Rate of Population Growth, Controlling for Size of County. ●—●, 1970; ○—○, 1980.

** Growth rate categories: (1)<0%; (2) 0%–10%; (3) 10%–50%; (4) 50%+.*

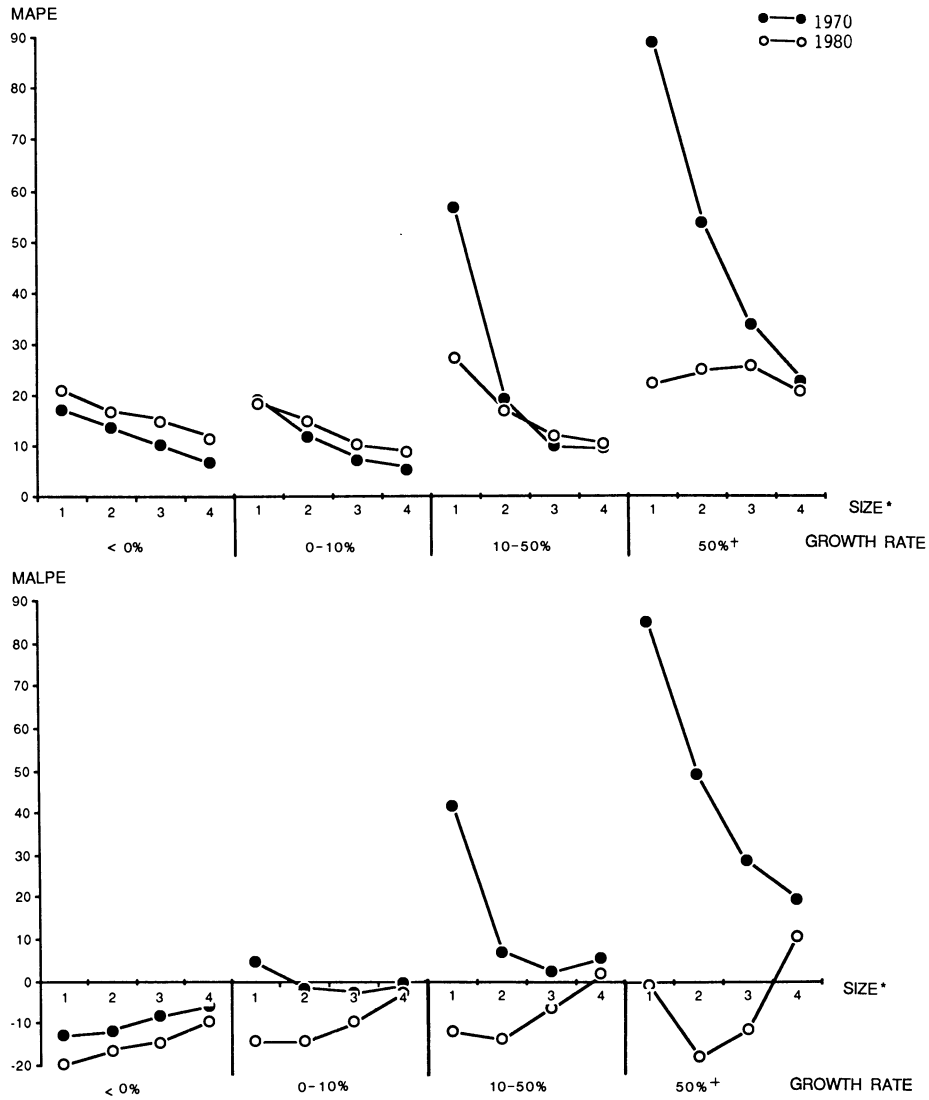

Figure 8. MAPE and MALPE for Two Sets of 10-Year Projections (AVE) by Size of County, Controlling for Rate of Population Growth. ●—●, 1970; ○—○, 1980.

** Size categories: (1)<5,000; (2) 5,000–14,999; (3) 15,000–49,999; (4) 50,000+.*

consistent independent impact on bias, once differences in growth rates were accounted for.

Why did Figures 1–3 show a positive relationship between MALPE and size of place? The most likely explanation is that in this sample county sizes and growth rates were strongly correlated. Table 5 shows this relationship clearly: for both sets of 10-year projections, the larger the population size category, the higher the growth rate during the previous decade. Thus the relationship between population size and bias shown in Figures 1–3 was spurious, being caused by the correlation between population size and growth rate.

Table 5. Average Decade Growth Rates for Counties, by Population Size at End of Decade

Population size	Decade growth rate	
	1950–1960	1960–1970
<5,000	−5.3	−7.6
5,000–14,999	−4.3	−1.5
15,000–24,999	1.3	3.3
25,000–49,999	10.7	9.1
50,000–99,999	24.4	17.4
100,000+	37.0	21.2

4. DISCUSSION

The data in Section 3 confirm a number of conclusions regarding forecast accuracy that have been drawn in previous studies: forecast errors tend to increase as the length of the projection horizon increases; forecast errors are generally larger for places with high growth rates than for places with low-to-moderate growth rates; and forecast errors are generally larger for small places than for large places. These findings are so well established that I believe they can be taken as general characteristics of the forecast accuracy of population projections.

This study has also presented some new evidence on forecast accuracy and bias. A number of studies have concluded that population projections generally have no predictable upward or downward biases (e.g., Kale et al.

1981; Smith 1984; Stoto 1983). The present study does not contradict this conclusion as it relates to overall bias: for any given set of projections, there is no way to know in advance whether forecast errors are likely to turn out to be predominantly positive or negative. This study does suggest, however, that a strong relationship exists between bias and the rate of population growth during the base period. For all projection techniques and all sets of projections, there was a strong positive relationship between the rate of population growth and the MALPE. This relationship was found in the 10-year projections for 1980, which exhibited a strong overall downward bias, and in the 10-year projections for 1970, which exhibited only a slight overall downward bias. This suggests that extrapolative projection techniques may generally forecast too low for counties with very low or negative growth rates during the base period and too high for counties with very high growth rates, *regardless* of the overall tendency for a particular set of projections to be too high or too low.

Another common conclusion is that the choice of projection technique has little impact on the forecast accuracy of short- to medium-term population projections, once the base period and length of projection horizon have been fixed (e.g., Ascher 1978; Greenberg, Krueckeberg, and Michaelson 1978; Kale et al. 1981; White 1954). The results reported here support that conclusion as it relates to overall accuracy but not when counties are broken down into size or growth rate categories. SHIFT consistently had larger errors and a stronger downward bias than LINE or EXPO for small counties and for counties with negative growth rates. EXPO consistently had larger errors and a stronger upward bias than LINE and SHIFT for large counties and for counties with high growth rates.

Why did these techniques exhibit these particular characteristics? I believe the answer lies with the "regression toward the mean" phenomenon described earlier. For counties that lost population during the base period, the LINE technique projects the *same* annual numerical decline over the forecast horizon as during the base period, the EXPO technique projects a *smaller* annual numerical decline (because the same percentage decline is applied to a smaller beginning population), and the SHIFT technique projects a *larger* annual numerical decline (unless the state population is projected to grow much more rapidly over the projection horizon than during the base period). If extreme growth rates tend to moderate over time, EXPO should produce the smallest errors and least downward bias for counties that lost population during the base period and SHIFT should produce the largest errors and greatest downward bias. That is exactly what was found in the empirical analysis.

For counties with high growth rates during the base period, the EXPO technique projects increasingly larger numerical increases over time because a constant growth rate is applied to a constantly expanding population, the LINE technique projects constant numerical increases, and the SHIFT technique projects increases that are generally between LINE and EXPO, depending on the state projection. If extreme growth rates tend to moderate over time, EXPO would be expected to have the largest errors and greatest upward bias for counties with high growth rates and LINE would be expected to have the smallest errors and least upward bias. This result is clearly shown in Figures 4–6 for counties with growth rates of 50% or higher.

On the basis of these results, it appears that some projection techniques may be *consistently* more accurate and less biased than other techniques for counties with particular size and/or growth rate characteristics. Thus for population projections that will be used primarily as forecasts, it may be advisable to use certain techniques for some places and different techniques for other places. For example, one might exclude linear and shift–share techniques from projections of places that lost population during the base period and exclude exponential techniques from projections of places that grew very rapidly. This "composite" approach has been tested before and was found to improve forecast accuracy (Isserman 1977). Research is needed on techniques other than those included in this study, but I believe the use of different projection techniques for places with different size and/or growth rate characteristics has the potential to significantly improve the forecast accuracy of population projections.

Table 3 showed a strong tendency for extreme growth rates to moderate over time. What caused this tendency? Migration is the demographic variable primarily responsible for differences in growth rates among states and local areas; differences in rates of natural increase are generally not large (Shryock and Siegel 1973, p. 793). If a county is growing very rapidly, then, the reason is most likely a high rate of net in-migration. Conversely, if a county is losing population, the reason is most likely net out-migration. An explanation for the tendency for extreme growth rates to moderate over time, therefore, must focus on changes in migration rates.

In order for rapidly growing areas to maintain constant growth rates, levels of net in-migration must continue increasing year after year. Yet if out-migration rates are based on the size of an area's population and in-migration rates are based on the size of the population outside that area, and if these rates remain constant over time, then levels of net in-migration must *decline* because the source of out-migrants is growing more rapidly than the source of in-migrants (Smith 1986b). Similarly, for areas losing population, levels of net out-migration must eventually decline because the source of out-migrants is growing more slowly than the source of in-migrants. Therefore, it is very unlikely that an area will maintain an extremely high or low population growth rate for an extended period of time.

Theoretical/behavioral explanations can also be given for expecting extreme migration rates to moderate over time. Rapidly growing areas have increasing numbers of "migration-prone" people who are likely to move again, whereas declining areas have declining numbers of such highly mobile persons (e.g., Miller 1967). Some in-migrants may become disenchanted with their new locations and return to their former homes (e.g., Eldridge 1965). It could also be argued that migration itself is a self-equili-

brating mechanism that causes the comparative advantage of one area over another to fade with time, eventually leading to declining rates of in- or out-migration for rapidly growing or declining areas (e.g., Borts and Stein 1964).

Thus there are theoretical reasons as well as empirical evidence to suggest that extreme growth rates for counties are likely to moderate over time. Further testing of this "regression toward the mean" phenomenon is necessary, especially on the timing of changes in rates of population growth, the correlates of those changes, and the effects of using a longer time period than that covered by this study. If additional evidence supports the findings reported here, the implications would be very important for population projections used as forecasts: since virtually all projection techniques (including cohort–component and economic-based) are dependent upon some type of extrapolation of past trends, it may be very useful to develop adjustments to moderate the projected rates of loss or increase for areas that lost population or grew very rapidly during the base period. Additional research must be undertaken to determine exactly how these adjustments might be accomplished (the "composite" approach mentioned earlier is one possibility), but its potential for improving accuracy and reducing bias appears to be great.

The empirical results shown in this article were based on three simple projection techniques. Would similar results be found for other commonly used projection techniques, such as cohort–component or economic-based? I believe there is a good chance the results would be similar in most respects. First, even though the underlying growth assumptions for the LINE, EXPO, and SHIFT techniques differed considerably from each other, the error characteristics for these three techniques were frequently much alike. Except for EXPO in large or rapidly growing counties and SHIFT in small or declining counties, the three techniques were quite similar in terms of the accuracy and bias of their projections. More important, the relationships between forecast errors and size of place, rate of growth, and length of projection horizon were much the same for all three techniques. Second, more sophisticated techniques are themselves typically based on extrapolations of one type or another (e.g., migration rates, birth rates, and survival rates for cohort–component projections; employment trends for economic-based projections). The functional forms of these extrapolations are often similar to those of the simple projection techniques (e.g., cohort–component projections using net migration data are frequently similar to exponential growth rate extrapolations). If applied to the same base period, then, the results of the more sophisticated extrapolations would most likely be

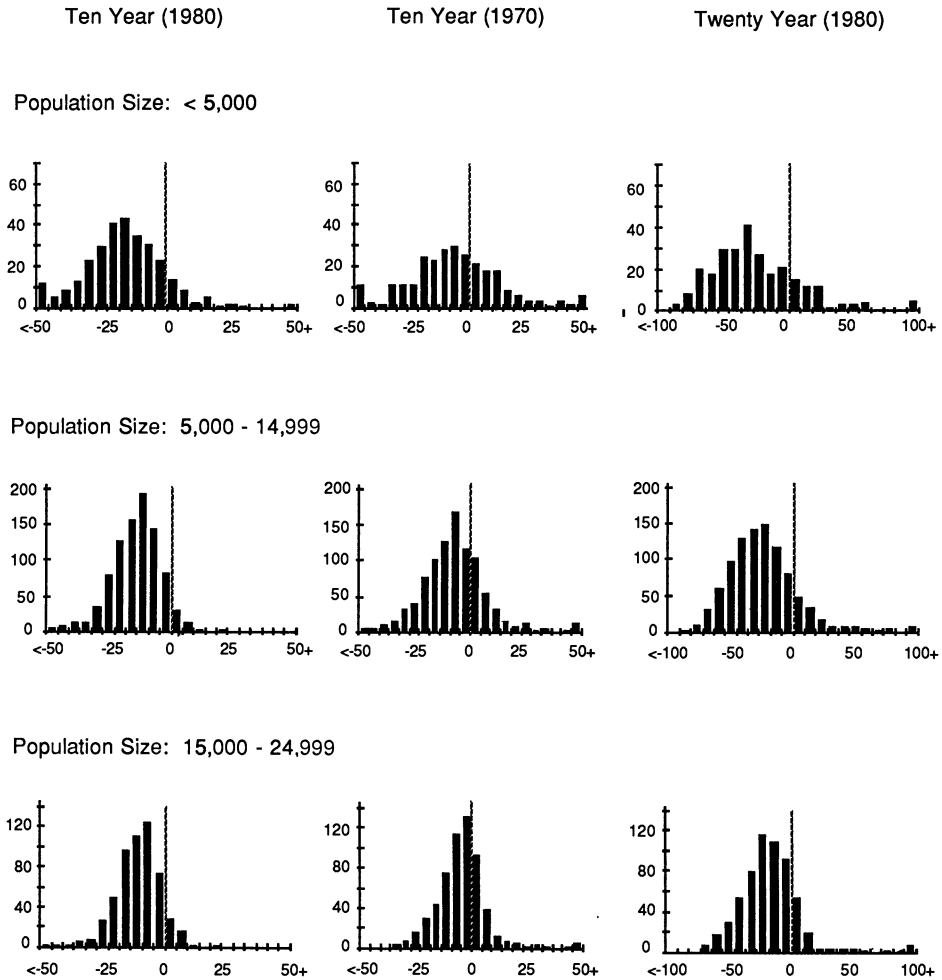

Figure 9. Frequency Distribution of Algebraic Percentage Errors for County Population Projections: AVE. The horizontal axis shows percentage errors; the vertical axis shows number of counties.

similar to the results of the simpler extrapolations. Finally, a number of previous studies have found the error characteristics for simple ratio and extrapolation techniques to be about the same as those for more sophisticated techniques (e.g., Kale et al. 1981; Siegel 1953; Smith 1984; White 1954). Although there would certainly be some differences, I believe the error characteristics reported in this article would be much like those found for most other projection techniques applied to the same set of counties and time periods.

5. EXTENSION: CONFIDENCE INTERVALS

A number of studies in recent years have focused on the distribution of forecast errors and the production of confidence intervals for population projections (e.g., Alho 1984; Cohen 1986). Some have used time series models in which historical population data were fit to autoregressive or moving average processes and future population values were made to depend on a weighted average of past values and a random error component (e.g., Kale et al. 1981; Lee 1974; Saboia 1974). Others have focused directly on the errors of projections made in the past (e.g., Keyfitz 1981; Stoto 1983). The data generated in the present study can be used for a simple application of this second approach.

Williams and Goodman (1971) suggested a method for constructing "empirical confidence limits" based on the distribution of past forecast errors. This method can accommodate any error distribution, including asymmetric and truncated. The critical assumption underlying this method is that the distribution of errors remains stable over time. Is this a reasonable assumption? Figure 9 shows the frequency distribution of forecast errors for the AVE technique, by size of place, for the two sets of 10-year projections and one set of 20-year projections discussed in this article. (Figures were also constructed for the three primary techniques and for errors by rate of growth, but they are not shown here.) A number of the patterns men-

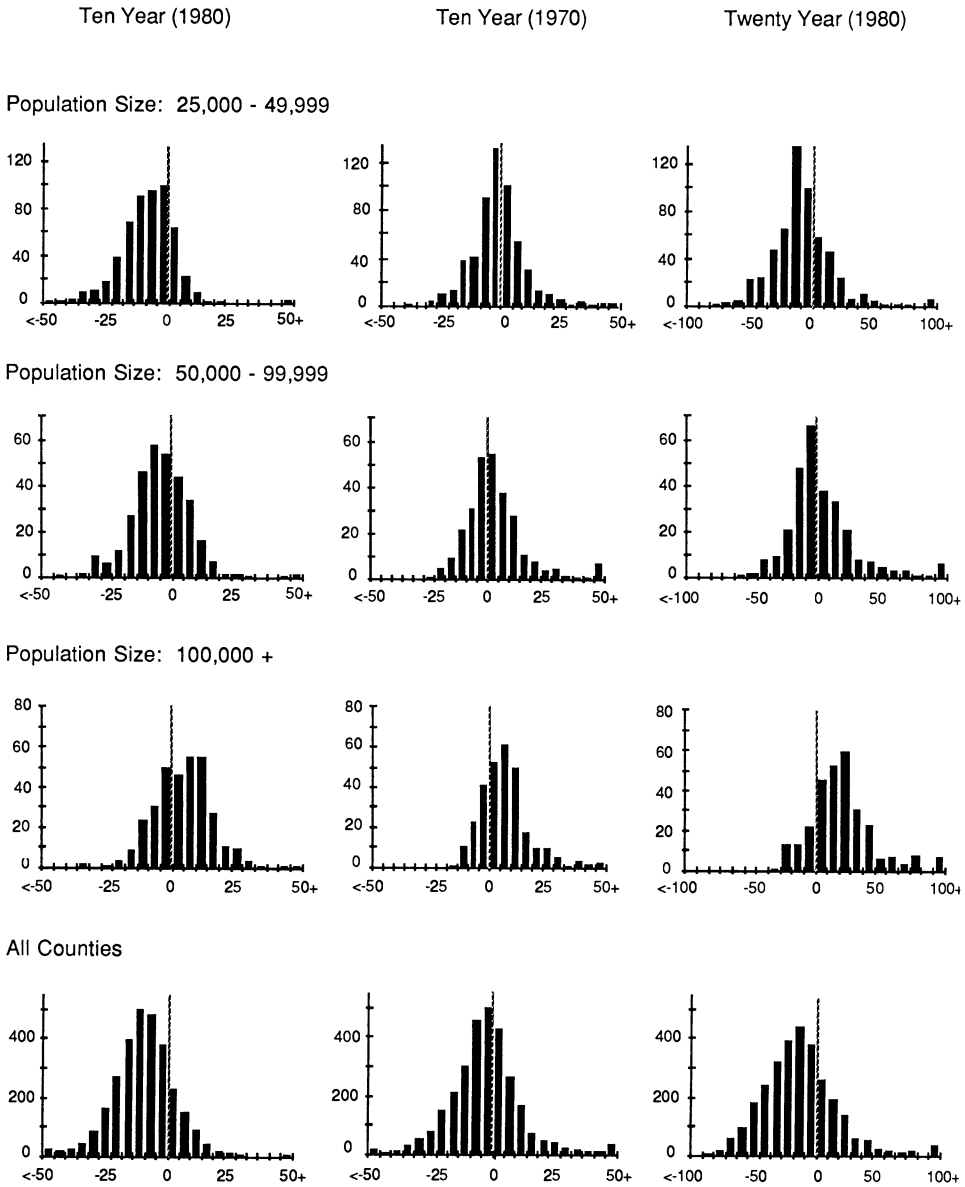

Figure 9 (continued).

Table 6. 90th Percentile Forecast Errors by Population Size for 10-Year Projections for 1970 and 1980

	LINE		EXPO		SHIFT		AVE	
Base population	*1970*	*1980*	*1970*	*1980*	*1970*	*1980*	*1970*	*1980*
<5,000	35.0	37.2	34.1	34.1	58.3	47.9	40.3	39.3
5,000–14,999	28.6	28.4	26.0	26.7	39.9	33.8	30.8	29.2
15,000–24,999	23.0	23.7	21.3	22.2	28.3	27.9	22.6	24.4
25,000–49,999	19.4	23.3	21.1	22.2	24.1	26.2	20.7	23.3
50,000–99,999	19.7	20.7	30.7	22.2	23.6	23.4	20.9	21.0
100,000+	17.2	17.7	34.8	22.9	20.9	21.2	23.7	20.7
Total	25.3	26.7	26.2	25.6	34.2	31.5	27.1	27.8

tioned before are evident in Figure 9, such as a downward bias stronger for the 1980 than the 1970 10-year projections, a downward bias stronger for small counties than large counties, and a dispersion of errors wider for 20-year than 10-year projections. It is also clear that the shapes of the frequency distributions were quite similar for the three sets of projections.

Following the Williams and Goodman approach, absolute percent forecast errors were ranked for each of the two sets of 10-year projections. The 90th percentile error (i.e., the absolute percent error that was larger than exactly 90% of all absolute percent errors) was noted by size of place for each technique and each time period. These results are shown in Table 6. There was a very high degree of similarity in the errors from the two sets of projections. Therefore, 1970 data on 90th percentile errors would have provided very accurate predictions of 90th percentile errors in 1980. For example, if the AVE technique had been chosen to provide forecasts for 1980, it would have been predicted that 90% of all projections would have errors of less than 27.1%. Actually, 90% of the 1980 AVE projections had errors of less than 27.8%, very close to the predicted value. A similar degree of accuracy would have been obtained using either LINE or EXPO, whereas SHIFT would have been slightly less accurate.

This analysis could have been refined even further. The prediction from AVE that all counties (regardless of size) would have 90% of errors less than 27.1% in 1980 would have been too low for small counties and too high for large counties. Data on errors by size of place could have been used to predict that 9 out of 10 projections would have forecast errors smaller than 40% for counties with a population less than 5,000, smaller than 30% for counties with 5,000–14,999, and smaller than 22% for counties with 15,000 or more. These predictions would have turned out to be very accurate. Similar refinements could be made by separating counties by rate of population growth during the base period or by size of place and rate of growth simultaneously.

The similarity in the distribution of forecast errors for the two sets of 10-year projections is striking, particularly in light of the major demographic changes that occurred between 1950 and 1980. Fertility rates rose during the 1950s and declined dramatically during the 1960s and early 1970s. Mortality rates declined steadily over the entire period. Levels of foreign immigration increased substantially. A number of regional net migration flows reversed direction and some long-standing trends in urban and rural population growth changed considerably. Yet the distribution of forecast errors for the two sets of 10-year projections remained remarkably stable. This stability supports the hypothesis that data on the distribution of past forecast errors may be very useful in predicting the distribution of future forecast errors. Firm conclusions cannot be drawn until research is conducted on other projection techniques and on error distributions covering more time periods than were covered in this study, but these preliminary results provide an indication of the potential usefulness of developing some type of "confidence limit" to accompany population forecasts.

6. SUMMARY AND CONCLUSIONS

Demographers frequently claim they are not in the business of predicting population. They refer to their estimates of future population as "projections" rather than "forecasts" and often produce a series of projections rather than a single set. This reluctance to predict is not surprising, given the degree to which many past forecasts have been wide of the mark. But users want forecasts, not projections. They want the author's prediction of what will actually happen in the future, not a series of hypothetical scenarios. In fact, users will generally interpret projections as forecasts, regardless of the author's intentions and whatever terminology or disclaimers might be used. As soon as projections reach the public, they become forecasts.

What can be said about the forecast accuracy of population projections? Some demographers have argued that testing forecast accuracy is pointless because what is really being tested are particular sets of assumptions within particular historical contexts (e.g., Pittenger 1978). I would argue that tests of forecast accuracy are essential because projections are so commonly used as forecasts and are so heavily relied upon for planning and budgeting purposes. It is certainly true that we cannot "know" what the future will be like, regardless of our knowledge of the past. But we can measure and evaluate the performance of population projections made in the past, and uncover as many consistent patterns as possible. If no consistent patterns are found, then we must conclude that very little can be said about the expected degree of forecast accuracy of population projections. That in itself would be useful (al-

beit disappointing) information for users of population projections.

This and other studies, however, have shown that consistent patterns do emerge from the empirical evidence. These patterns have been found for a number of different projection techniques and for different base periods and projection horizons. Although this evidence does not prove that future errors will follow the same patterns as past errors, at least it does not cause one to reject such a hypothesis. The past may not be a foolproof guide to the future, but it does provide a useful set of clues.

This study has documented several strong relationships between population forecast errors and the size of the base population, the rate of growth during the base period, and the length of the projection horizon. Some interesting questions have been answered (at least conditionally), but many others remain to be explored. Would other projection techniques or time periods yield results similar to those shown here? What other consistent patterns might be found, such as differences in forecast errors between central cities and suburban areas or urban and rural areas? What effect would changing the length of the base period (e.g., 5 years, 20 years) have on forecast errors? Would controlling county projections to independent state projections improve forecast accuracy? Would the selective application of the forecaster's personal judgement to the mechanistic projection techniques used in this study tend to raise or lower forecast errors? These are just a few of the questions that require additional study. Future research will surely provide some interesting and useful answers.

[*Received April 1985. Revised April 1987.*]

REFERENCES

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

Ascher, William (1978), *Forecasting: An Appraisal for Policy-Makers and Planners,* Baltimore: Johns Hopkins University Press.

Birch, David L. (1977), "Comments on Long-Term Population and Employment Forecasting," in *Report of the Conference on Economic and Demographic Methods for Projecting Population,* Washington, DC: American Statistical Association, pp. 148–152.

Borts, George H., and Stein, Jerome L. (1964), *Economic Growth in a Free Market,* New York: Columbia University Press.

Cohen, Joel (1986), "Population Forecasts and Confidence Intervals for Sweden: A Comparison of Model-Based and Empirical Approaches," *Demography,* 23, 105–126.

Eldridge, Hope T. (1965), "Primary, Secondary and Return Migration in the United States, 1955–1960," *Demography,* 2, 444–455.

Federal–State Cooperative Program for Population Projections (1984), "State Survey on Population Projections," Sacramento: California Department of Finance, Population Research Unit.

Greenberg, Michael (1972), "A Test of Combinations of Models for Projecting the Population of Minor Civil Divisions," *Economic Geography,* 48, 179–188.

Greenberg, Michael, Krueckeberg, Donald, and Michaelson, Connie (1978), *Local Population and Employment Projection Techniques,* New Brunswick: Center for Urban Policy Research.

Hajnal, John (1955), "The Prospects for Population Forecasts," *Journal of the American Statistical Association,* 50, 309–322.

Irwin, Richard (1977), "Guide for Local Area Population Projections," Technical Paper 39, Washington, DC: U.S. Bureau of the Census.

Isserman, Andrew (1977), "The Accuracy of Population Projections for Subcounty Areas," *Journal of the American Institute of Planners,* 43, 247–259.

Isserman, Andrew, and Fisher, Peter (1984), "Population Forecasting and Local Economic Planning: The Limits on Community Control Over Uncertainty," *Population Research and Policy Review,* 3, 27–50.

Kale, Balkrishna, Voss, Paul, Palit, Charles, and Krebs, Henry (1981), "On the Question of Errors in Population Projections," unpublished paper presented at the annual meeting of the Population Association of America, Washington, DC.

Keyfitz, Nathan (1972), "On Future Population," *Journal of the American Statistical Association,* 67, 347–363.

——— (1981), "The Limits of Population Forecasting," *Population and Development Review,* 7, 579–593.

Lee, Ronald D. (1974), "Forecasting Births in Post-Transition Populations: Stochastic Renewal With Serially Correlated Fertility," *Journal of the American Statistical Association,* 69, 607–617.

Miller, Ann (1967), "The Migration of Employed Persons To and From Metropolitan Areas of the United States," *Journal of the American Statistical Association,* 62, 1418–1432.

Morrison, Peter A. (1977), "Forecasting Population of Small Areas: An Overview," in *Population Forecasting for Small Areas,* Oak Ridge, TN: Oak Ridge Associated Universities, pp. 3–13.

Pittenger, Donald (1978), "The Role of Judgment, Assumptions, Techniques and Confidence Limits in Forecasting Population," *Socioeconomic Planning Sciences,* 12, 271–276.

——— (1980), "Some Problems in Forecasting Population for Government Planning Purposes," *The American Statistician,* 34, 135–139.

Saboia, Joao (1974), "Modeling and Forecasting Population by Time Series: The Swedish Case," *Demography,* 11, 483–492.

Schmitt, Robert, and Crosetti, Albert (1951), "Accuracy of the Ratio Method for Forecasting City Population," *Land Economics,* 27, 346–348.

——— (1952), "Accuracy of the Ratio Method: Rejoinder," *Land Economics,* 28, 183–184.

——— (1953), "Short-Cut Methods of Forecasting City Population," *Journal of Marketing,* 17, 417–424.

Shryock, Henry, Siegel, Jacob, and associates (1973), *The Methods and Materials of Demography,* Washington, DC: U.S. Government Printing Office.

Siegel, Jacob S. (1953), "Forecasting the Population of Small Areas," *Land Economics,* 29, 72–88.

——— (1972), "Development and Accuracy of Projections of Population and Households in the United States," *Demography,* 9, 51–68.

Smith, Stanley K. (1984), *Population Projections: What Do We Really Know?,* Monograph No. 1, University of Florida, Bureau of Economic and Business Research.

——— (1986a), "Population Projections in the Real World: Tests of Forecast Accuracy," unpublished paper presented at the annual meeting of the Population Association of America, San Francisco.

——— (1986b), "Accounting for Migration in Cohort–Component Projections of State and Local Populations," *Demography,* 23, 127–135.

Stoto, Michael (1983), "The Accuracy of Population Projections," *Journal of the American Statistical Association,* 78, 13–20.

United States Bureau of the Census (1965), "Illustrative Projections of the Population of States: 1970 to 1985," *Current Population Reports,* Ser. P-25, No. 301.

——— (1972), "Preliminary Projections of the Population of States: 1975 to 1990," *Current Population Reports,* Ser. P-25, No. 477.

Voss, Paul R., and Kale, Balkrishna D. (1985), "Refinements of Small-Area Population Projection Models: Results of a Test Based on 128 Wisconsin Communities," unpublished paper presented at the annual meeting of the Population Association of America, Boston.

White, Helen R. (1954), "Empirical Study of the Accuracy of Selected Methods of Projecting State Population," *Journal of the American Statistical Association,* 49, 480–498.

Williams, W. H., and 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.