Peter Pflaumer[1]

Evaluating the Accuracy of Population Forecasts

Paper Presented at the 3rd International Conference on Social Science Methodology, Trento, Italy, 22 – 26 June 1992

ABSTRACT: In this paper the accuracy of population forecasts is discussed. Various papers on errors of population forecasting are reviewed and summarized. The results are stated in six theses. The main findings show that no clear dominance of any one forecasting method can be determined, that the logarithmic forecast errors are more or less independent of the length of the forecast horizon, and that saturation models underestimate the population development in the long term, whereas geometric and polynomial trend models overestimate it. Finally, the accuracy of population forecasts is compared with the accuracy of short-term and long-term economic forecasts. It is found that the error of population forecasts is smaller than that of economic forecasts. However, the logarithmic error decreases with the length of the forecast period for most economic variables.

KEY WORDS: Demographic models, economic forecasts, forecast errors, RMSE

1. INTRODUCTION

The necessity of population forecasts is hardly disputed. In politics, in public administration and in business, far-reaching decisions are made which depend on the future development of the population. The reliability of population predictions is influenced, however, by a multitude of factors. If long-term population forecasts are to serve as a rational basis for decision-making, then one needs to have an idea of the uncertainty of such population forecasts. The smaller this uncertainty is, the more willing people will be to make decisions that are dependent on demographic factors (e.g. a decision to stabilize the financing of old age pensions). Several possibilities for taking the uncertainty into consideration exist: Sensitivity analyses (see, e.g., Pflaumer, 1988b), forecast intervals in time series models (see, e.g., Pflaumer, 1992), and stochastic component models (cf., e.g., Land, 1987 or Pflaumer, 1988b,c). A further method of describing the uncertainty in population forecasts is based on the calculation of forecast error measures. They serve the ex post evaluation of a forecast. Under certain conditions it is even possible to construct a forecast interval for future predictions using the distribution of past forecast errors as a basis (see, e.g., Keyfitz, 1981).

[1] Department of Statistics, University of Dortmund, Germany

2. MEASURING THE FORECAST ERRORS

The measures of error considered are based on the differences between the projected and the actual annual population growth rates. Keyfitz (1981) and Stoto (1983) have shown that this measure adjusts for the length of the projection period.
Let P_0 be the population at the beginning of the projection period, and P_T be the actual population T years later. It is easily shown that the actual average annual growth rate is

$$r = \frac{1}{T}\ln(P_T / P_0).$$

The projected average annual growth rate is

$$\hat{r} = \frac{1}{T}\ln\left(\hat{P}_T / \hat{P}_0\right),$$

where $\hat{P}_T$ is the projected population at time T and $\hat{P}_0$ is the estimated population at the beginning of the projection period. The following measures of error are considered for the analysis:

(1) the logarithmic forecast error

$$d = \hat{r} - r$$

(2) the average error or bias

$$BIAS = \frac{1}{n}\sum(\hat{r}_i - r_i) = \frac{1}{n}\sum d_i$$

(3) the mean square error

$$MSE = \frac{1}{n}\sum(\hat{r}_i - r_i)^2 = \frac{1}{n}\sum d_i^2$$

(4) the root mean square error

$$RMSE = \sqrt{\frac{1}{n}\sum(\hat{r}_i - r_i)^2} = \sqrt{\frac{1}{n}\sum d_i^2}$$

with

r_i = actual growth rate
$\hat{r}_i$ = projected growth rate
$d_i = \hat{r}_i - r_i$

Table 1: Error Measures of U.N. Population Projections by Continents (Percentage Points)

Forecasting Period	Africa		Asia		South Americ a		Rest	
	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE
1960-1965	-0.495	0.870	0.064	0.569	-0.153	0.515	-0.069	0.234
1960-1970	-0.700	1.104	-0.082	0.624	0.003	0.431	0.018	0.247
1960-1975	-0.616	0.993	0.044	0.505	0.120	0.424	0.041	0.276
1960-1980	-0.518	0.844	0.120	0.429	0.208	0.392	0.091	0.302
1965-1970	-0.173	0.606	0.216	0.704	0.329	0.705	0.154	0.292
1965-1975	-0.171	0.816	0.293	0.561	0.431	0.713	0.143	0.215
1965-1980	-0.244	0.783	0.308	0.517	0.465	0.626	0.205	0.279
1965-1985	-0.116	0.623	0.253	0.462	0.503	0.636	0.263	0.321
1970-1975	-0.024	0.507	0.026	0.671	0.152	0.753	-0.041	0.335
1970-1980	-0.073	0.588	0.173	0.620	0.258	0.816	0.043	0.308
1970-1985	-0.104	0.516	0.109	0.453	0.293	0.680	0.104	0.279
1975-1980	0.004	0.429	0.157	0.392	-0.006	0.513	0.007	0.298
1975-1985	-0.119	0.665	0.142	0.573	0.233	0.633	0.042	0.206
1980-1985	0.003	0.550	-0.009	0.415	0.117	0.423	0.076	0.300
Mean	-0.233	0.707	0.124	0.535	0.211	0.590	0.077	0.278

Source: Own calculations.

Table 1 shows as an example error measures of U.N. population projections. The results are based on the medium variant of projections made by the United Nations between 1960 and 1985 for 101 countries with a population of one million or more. The results indicate that the forecasting performance depends on the year in which the projection was made. The correlation between jump-off year and forecast error was noted by other authors as well. Stoto (1983), e.g., showed that the jump-off year of the forecast is strongly correlated with the errors in population growth for forecasts made in the United States by the U.S. Bureau of the Census between 1946 and 1971. There is some evidence that the RMSE is nearly invariant with respect to the time period over which the projection is made. This result has already been found by Keyfitz (1981). Comparing the present results with Keyfitz's results, we see that the present RMSEs are larger. The reason for this is that more countries have been included, especially more African countries, which showed large forecast errors. Although the RMSE is nearly invariant with respect to the forecasting period, the variance of the RMSE increases with respect to the span of projections, which expresses the increasing uncertainty (see Jöckel/ Pflaumer, 1984). The independence of the forecast error and the length of the forecast horizon is of course only valid for errors which are based on growth rates. Other forecast errors generally increase with the length of the forcast horizon. Smith and Sincich (1991), e.g., find that in most instances there is a linear or nearly linear relationship between forecast accuracy and the length of the forecast horizon.

3. SUMMARY OF THE RESULTS OF THE STUDY OF THE ACCURACY OF POPULATION FORECASTS

Recently, several demographers have published papers on the subject of the ex post evaluation of population forecasts. These include, for example, the studies of Ascher (1979), Inoue/Yu (1979), Kale et al. (1981), Keyfitz (1981), Stoto/Schrier (1982), Stoto (1983), Jöckel/Pflaumer (1984), Keilman (1990), as well as Smith/Sincich (1991). It would go beyond the scope of this section if each of these studies were to be discussed separately. Therefore, the significant results of the studies will be summarized, supplemented by more results and presented in the form of theses that will be briefly discussed and explained. It should be mentioned that forecast error measures refer to measures that have been calculated with logarithmic forecast errors. In several of the papers cited, the forecast error measures were calculated with relative or absolute forecast errors, so that a comparison is only possible to a certain extent. In addition, it should be mentioned that a forecast error analysis using the previously given measures can only be done in the case of point forecasts. If alternative projections were carried out, then the forecast error always refers to the medium projection.

Summary 1 contains six theses concerning the accuracy of population forecasts which are explained in the following subsections.

Summary 1

Theses concerning the accuracy of population forecasts

1. In the case of population forecasts, no clear dominance of any one method can be determined.
2. In the long term, saturation models underestimate the population development, whereas geometric and polynomial trend models overestimate it.
3. The forecasting accuracy of the component method and that of other methods essentially depends on the time when the forecast was made.
4. Logarithmic forecast errors and the RMSE formed from them are more or less independent of the length of the forecast horizon.
5. The forecast error increases with the growth rate of the population.
6. The error of population forecasts is smaller than that of economic forecasts.

1. In the case of population forecasts, no clear dominance of any one method can be determined.

W. Ascher (1979) compared the accuracy of North American population forecasts of different demographers. The predictions, which were made between 1891 and 1972, include simple trend models (Pritchett, Sloane, Gannet), saturation curves (Pearl/Reed) and component methods (Whelpton, Dublin/Lotka, Whelpton/Thompson and U.S. Census Bureau). Concerning Ascher's results, the absolute values of the logarithmic forecast error d have been determined in table 2. Errors are given for a forecast horizon of 5, 10 and 20 years, as well as

Table 2
Absolute values of the logarithmic forecast error *d* (percentage points)
for various population forecasts of the U.S.A.

year	source	5 years	10 years	20 years	1960	1970	1975
1891	Pritchett		0.18	0.13	0.33	0.31	0.35
1900	Pritchett		0.25	0.38	0.38	0.35	0.39
1909	Sloane		0.39	0.31	0.28	0.31	
1909	Gannet		0.23	0.17	0.14	0.18	
1920	Pearl/Reed		0.04	0.17	0.30	0.37	0.35
1925	Pearl	0.12	0.02	0.16	0.34	0.42	0.39
1928	Whelpton		0.43	0.02	0.29	0.41	0.39
1930	Dublin/Lotka		0.34	0.29	0.60	0.69	
1933	Whelpton/Thompson	0.08	0.15	0.58	0.63	0.69	0.66
1938	Whelpton/Thompson	0.20	0.68	0.85	0.78	0.78	0.72
1940	Census Bureau, P-3, 15	0.60	0.67	0.94	0.94	0.90	0.80
1943	Whelpton/Thompson	1.04	1.12	1.05	0.97	0.90	0.82
1947	Census Bureau						
	Forecasts 1945-1975		0.97	0.98			0.87
1953	Census Bureau, P-25, 78	0.50	0.43	0.10	0.37	0.21	0.10
1958	Census Bureau, P-25, 187	0.00	0.21		0.20	0.18	0.44
1964	Census Bureau, P-25, 286	0.08	0.40			0.07	0.36
1970	Census Bureau, P-25, 448	0.20					0.20
1972	Census Bureau, P-25, 493						0.13

Source: Calculated from Ascher (1979).

for the target years 1960, 1970 and 1975. The fact is remarkable that the earlier forecasters obtained better results with their simple methods than the demographers of the 30's and 40's who used the component method, a new method for population forecasting, for the first time.

Forecasts have continually improved over the last 30 years, whereby the question must be left open whether this fact is due to the refinement of the forecast methods and more possibilities in the data processing sector or whether the demographic environment is nowadays easier to forecast because the variations of the components of the population growth have decreased. If one compares the accuracy of the U.S. Census Bureau forecasts with those of the earlier forecasts, one does not find any great differences. The prediction error of the yearly growth rates varies in both cases between 0.1 and 0.4 percentage points.

In order to get an idea of the accuracy of the Box-Jenkins method in the case of population forecasts, ARIMA(2,2,0) models of the total population development for the United States have been identified, estimated and used as forecasts for various time periods (cf. Pflaumer, 1988b and Pflaumer, 1992). Long-term ex ante forecasts were made, and subsequently the logarithmic forecast error *d* was calculated for all ARIMA models (cf. table 3).

The errors with the same forecast horizon are found in the diagonals, whereas the errors for the same target or base year are listed in the columns or lines.

Table 3
Logarithmic forecast errors of the Box-Jenkins method with U.S. population forecasts (percentage points)

estimation interval of the model	forecast year	1900	1910	1920	1930	1940	1950	1960	1970	1980
1790-1900			-0.29	0.09	-0.08	0.09	0.04	-0.06	-0.08	-0.04
1790-1910				0.48	0.36	0.52	0.40	0.23	0.19	0.21
1790-1920					-0.44	-0.08	-0.19	0.35	-0.35	-0.28
1790-1930						0.42	0.05	-0.21	-0.23	-0.16
1790-1940							-0.58	-0.78	-0.69	-0.53
1790-1950								-0.16	0.01	0.18
1790-1960									0.35	0.50
1790-1970										0.20

Source: Own calculations.

If one compares the results in table 2 with those in table 3, no great differences can be seen for the absolute values of the logarithmic errors. Other authors also arrive at the conclusion that the Box-Jenkins method is in principle just as good or just as bad as all of the other methods of population forecasting. Kale et al. (1981) assign a reliability to simple ARIMA(1,1,0) models that is equal to that of other methods of population forecasting.

Voss/Palit (1981) investigated the forecasting properties of ARIMA(1,1,0) models for the prediction of the population in 48 states of the U.S.A.. They too came to the conclusion that the Box-Jenkins model can by all means compete with the component method. Thus they wrote, "The strongest defense (of using ARIMA models) lies in our ex post evaluation. We have demonstrated that this strategy produces population forecasts at least as reliable as more traditional demographic models."

Stoto/Schrier (1982), who also studied the accuracy of population forecasts in 48 states of the U.S.A., discovered that simpler methods did not predict the future population development any worse than more complicated methods did. The fact that the far more complicated model, a demo-econometric forecasting model, had the worst forecasting properties, was commented on by the authors as follows: "Complexity simply does not pay off." In the following table, Stoto/Schrier compare error measures of the geometric model with those of the Census Bureau projections. This comparison is based on the entire population of the United States. The logarithmic forecast error *d* was used for the calculation of the error measures.

Table 4

Error measures for population forecasts in 48 States of the U.S.A. from 1955 - 1975 (percentage points)

projection	forecast horizon (years)	bias	standard deviation	RMSE
geometric model	5	-0.067	1.03	1.032
geometric model	10	0.252	0.87	0.906
census bureau projections	5	-0.200	0.87	0.897
census bureau projections	10	-0.143	0.85	0.861

Source: Stoto/Schrier (1982).

Stoto/Schrier conclude from these results that it is appropriate to use the geometric projection model, since it does not require any considerable effort and yet is no more uncertain than the component method. However, if demographic or economic details about the future development are needed, then the geometric projection model is of course no substitute for the more complicated methods, since geometric models only forecast the entire population (cf. Stoto/Schrier, 1982, p. 29 ff.).

2. In the long term, saturation models underestimate the population development, whereas geometric and polynomial trend models overestimate it.

Previous experience with the logistic function has shown that these models have caused the population to be underestimated in the long term: in the 1920's, Pearl and Reed (1920) predicted a saturation level of approximately 196 million for the U.S.A.. As is well known, this population level was surpassed as early as the mid 1960's.

Keyfitz (1979) fit the four countries France, Sweden, England and the U.S.A. each with a logistic function and a geometric trend curve within different time periods and produced forecasts that extended from 10 to 50 years. In the overwhelming majority of the cases, the population was significantly underestimated by the logistic function after as little as 30 years. With a forecast period of this length, the geometric trend curve showed no clear tendency towards over- or underestimating the population. In almost 70% of the cases, the accuracy of the geometric model was greater than that of the logistic one. In the case of short- and medium-term forecasts, the geometric model can by all means compete with other methods, as the explanations in the previous section have shown. In the long term, the application of the geometric method with a constant growth rate leads to great overestimations of the population.

Forecasts by Wigglesworth (1775) or by Elster (1891) are the best examples for this. In the year 1775, a clergyman named E. Wigglesworth, Professor at Harvard College in Cambridge, Massachusetts, published a paper titled "Calculations on American Populations". In this study he made a long-term forecast for the population of the "British Colonies". He had observed that the population of his country doubled approximately every 25 years, which corresponded to a yearly growth rate of 2.8%. If one assumes, as Wigglesworth did, an initial population of 2.5 million and a continued constant growth rate of 2.8%, then one arrives at a forecast of 640 million people for the year 1975. With this figure, however, Wigglesworth considerably overestimated the actual population, which was less than 220 million in 1975.

One of the first population forecasts for Germany was published by E. Elster in the year 1891 as part of an article in the Handbook of Political Science (cf. Elster, 1891). Based on the population development in Germany from 1871 to 1880, he calculated population growth rates taking and not taking emigration into consideration. Assuming constant growth rates, he estimated the population of Germany to be either 165 or 207 million in the year 2000, depending on whether emigration was considered or not. The reason for the overestimation is the fact that the growth rates of the populations have decreased in the past.

However, polynomial trend models, which take the decreasing growth rates into consideration, have also always led to an overestimation of the population after a certain period of time. The most well-known model of this kind is the Pritchett model: Pritchett (1891) criticizes the geometric growth model for the assumption of constant growth rates, since fertility sinks in all industrialized countries with progressive development. In the study of population data from 1790 to 1880, he comes to the conclusion that the growth development of the U.S.A. can be best described with a polynomial function of degree three. The growth rates now decrease with increasing time, but the polynomial function has not a saturation level. The Pritchett model, which produced excellent forecasts for the population of the U.S. until 1980, estimates a population of 376 million for the year 2000, a figure that is undoubtably too high. For the year 2050, even 600 million residents of the U.S.A. are forecasted. Obaidullah (1976), however, used a combination of linear and exponential trend curves to forecast the population of Bangladesh. He came to the conclusion that this model yielded the best approximation of the actual figures.

3. The forecasting accuracy of the component method and that of other methods essentially depends on the time the forecast was made.

When the logarithmic forecast errors in tables 2 and 3 are compared, it becomes apparent that the variations within the columns are larger than those within the lines. Concretely, this means that the forecast error primarily depends on the point of time at which the forecast was made and not so much on the length of the forecast horizon. A forecast becomes especially difficult when unpredictable things occur which strongly influence the population growth. These interference factors can lie in the demographic as well as in the non-demographic areas. Demographically-caused forecast errors result above all when the birth development or the size of immigration and emigration have been incorrectly estimated. The demographers of the 30's and 40's greatly underestimated the future population development despite their comparatively sophisticated methods. The reason for the incorrect predictions lay in the assumption that the low fertility level of these years would continue in the future. The baby boom of the 50's and the early 60's was not predicted and therefore not included in any calculations either. Thus, for example, the respected demographer Whelpton, who produced his forecast in 1943, underestimated the yearly population growth on the average by one percentage point (cf. table 2). The forecasters of the 60's and 70's also estimated the future population incorrectly, just as the forecasters of the 40's had done. The population estimations had to be continually revised downwards because forecasters had assumed high birth rates for too long. For example, in 1970 the Federal Office of Statistics in Germany forecasted a resident population of 66 million in West Germany for the year 1990. However, the actual population was only 60 million.

In conclusion, it can be determined that demographers were successful with their population forecasts in those years in which no special changes in fertility or in migration occurred after the predictions were made. Demographers were inaccurate with their forecasts when these changes did come about. If one compares forecasters to archers, then the one archer would hit the bull's-eye when the target happened to be standing still while he aimed, whereas another archer would miss the bull's-eye because the target just happened to move at the moment that he aimed (Keyfitz, 1979).

4. Logarithmic forecast errors and the RMSE formed from them are more or less independent of the length of the forecast horizon.

Keyfitz (1981) was the first to point out that when the logarithmic forecast error is used in the analysis, the accuracy of a population forecast is practically independent of the time period about which the prediction is made. In his study he analyzed the errors of 1100 population forecasts. These forecasts had been made by the United Nations in the time period between 1958 to 1968 for all countries of the world with a population of more than one million. The most important results of Keyfitz' study can be seen in table 5.

Table 5

RMSE of population forecasts of the United Nations (percentage points)

Base year of the forecast	forecast period	population growth rate			all population growth rates
		low	medium	high	
1958	1955-1960	0.320	0.555	0.793	0.589
	1955-1965	0.338	0.545	0.713	0.553
	1955-1970	0.323	0.534	0.698	0.541
	1955-1975	0.310	0.538	0.691	0.536
	all	0.323	0.543	0.725	0.555
1963	1960-1965	0.236	0.421	0.527	0.413
	1960-1970	0.248	0.447	0.487	0.408
	1960-1975	0.250	0.481	0.596	0.465
	all	0.245	0.450	0.539	0.429
1968	1965-1970	0.298	0.359	0.398	0.354
	1965-1975	0.245	0.418	0.390	0.359
	all	0.273	0.390	0.394	0.357
all base years		0.288	0.478	0.604	0.476

Source: Keyfitz (1981).

Keyfitz constructs the hypothesis that the RMSE of population forecasts is on the average circa 0.4 percentage points. In the case of less developed countries, this error is somewhat higher, whereas it is somewhat lower in the case of more developed countries. He especially emphasizes that the error is not dependent on the length of the forecast period as long as this is less than 25 years. In other words, the length of the forecast period does not affect the accuracy of the prediction. Similar results have been obtained by Pflaumer (1988a). The constancy of the RMSE calculated from logarithmic forecast errors is of course not the same as the constancy of the RMSE calculated from absolute forecast errors. As can be easily shown, the absolute error increases with the length of the forecast horizon (cf. Smith/Sincich, 1991).

In connection with the study of Keyfitz, Jöckel/Pflaumer (1984) analyzed the errors of 360 population forecasts of the United Nations from the year 1958. They too arrived at the conclusion that the forecast error and the RMSE are relatively stable, but that the variance of the RMSE increases with the length of the forecast horizon (cf. table 6). Indeed, this result also seems to be plausible. The increasing variance is an expression of the growing uncertainty that is caused by an extension of the forecast period.

The stability of the error in relation to the length of the forecast period can also be seen in table 2. Whether the population is estimated for 5, 10, 20, or more years, the magnitude of the logarithmic forecast error does not change that significantly. This fact is especially noticable in the forecast of Pritchett from the year 1900. The forecast error is 0.25 percentage points in the ten-year forecasts and slightly less than 0.4 percentage points in all other forecast periods.

Table 6
Error measures for population forecasts of the United Nations
of the year 1958 (percentage points)

forecast period	RMSE	standard deviation of the RMSE
1955-1960	0.58	0.17
1955-1965	0.55	0.19
1955-1970	0.54	0.24
1955-1975	0.54	0.32

Source: Calculated from Jöckel/Pflaumer (1984).

5. The forecast error increases with the growth rate of the population.

Empirical studies verify that the higher the growth rate of the population, the higher the forecast error. Keyfitz (1981) differentiates in table 5 between countries with low, medium and high growth rates. In the fast growing regions, the RMSE of 0.6 is more than twice as high as that in the slowly growing regions, which is only 0.288. Stoto/Schrier (1982) also speak of a positive relationship between the forecast error and the growth rate of the population. In the case of the population forecast for 48 states of the U.S.A., they calculate a correlation coefficient of about 0.8 between the growth rate and the RMSE. Pflaumer (1988a) also concludes that population growth tends to increase the forecast error.

6. The error of population forecasts is smaller than that of economic forecasts.

Without having undertaken empirical comparisons of economic and population forecasts, E. Günther, a German statistican, had already supposed in 1935 that there was no process in human life that could be predicted so exactly over such a long period of time as the population development. "A farmer or forester, a businessman, a minister of finance or transportation who could make his cost estimates nearly as accurately for 36 months instead of 36 years would be very fortunate". Günther comes to his euphoric conclusions from an ex post estimation of the population of Germany from 1871 to 1930, which he had done in 1931. When the logarithmic forecast error is applied to his results, it shows a maximum of about 0.2 percentage points.

In the following, the thesis that population forecasts are better than economic forecasts will be proven with data. In order to be able to correctly classify the results, we once again refer to the study of Keyfitz, which determined that the RMSE of population forecasts is roughly 0.4 percentage points and that this is independent of the length of the forecast period. The comparison of the accuracy of economic and population forecasts is divided into two parts. First, the accuracy of short-term economic forecasts will be evaluated. A multitude of studies exist for this purpose. These will have to be restricted, however, to the studies of McNees (1981) and Pflaumer (1986), since their forecast error measures are the only ones that were calculated from logarithmic errors. If the other studies were to be included, a direct comparison with the previous results would no longer be possible. After this, attention will be turned to the evaluation of long-term economic forecasts. There is little literature concerning the ex post evaluation of long-term economic forecasts (cf. eg. Ascher, 1979, and the studies cited therein). For this reason, a few new calculations have been carried out for U.S. data.

An overview of the accuracy of short-term forecasts can be found in the publications of McNees (1981), who investigates the accuracy of the forecasts of various organizations of the U.S.A. at regular intervals. The organizations in question are consulting firms, university institutes, government posts and planning divisions of companies. The comparison consists of a total of 13 institutions that regularly produce and publish forecasts. Some of these organizations limit themselves to predicting only a few economic variables, whereas others predict up to 10,000. The forecast horizon ranges from one month to up to four years in the future. The forecast methods used by the organizations under consideration are econometric models, time-series methods, data analyses and personal judgements. The period of investigation for the comparison extends from the 1st quarter of 1976 to the 3rd quarter of

1980. In the evaluation, McNees had to analyze roughly 36,000 errors. He concentrated the material with several forecast error measures and published it in a number of tables. One appendix is of special interest. This contains data concerning the RMSE calculated from logarithmic errors. In the original, McNees calculated a separate RMSE for each forecasting organization. These data were used to calculate an average RMSE, and the results thereof can be seen in table 7.

Table 7

RMSE for various economic variables of the U.S.A. (percentage points)

	Forecast horizon (quarters)							
	1	2	3	4	5	6	7	8
real national product (in 1972 prices)	3.3	2.4	1.7	1.2	0.9	0.9	1.0	1.2
gross national product	6.1	3.8	2.7	2.4	2.0	1.9	1.7	1.7
personal consumption expenditure (durables)	11.2	8.5	5.2	4.2	4.0	4.2	3.9	3.9
personal consumption expenditure (nondurables)	2.6	2.5	2.2	2.1	2.2	2.4	2.2	2.2
nonresidential fixed investment	9.0	6.0	5.3	5.2	4.8	4.6	4.2	4.2
investment in residential structures	18.4	13.3	11.3	11.2	10.8	9.9	9.9	10.7
federal government purchases	8.9	6.6	5.6	4.8	4.0	3.7	3.3	2.5
consumer price index (1967=100)	1.9	2.2	2.7	2.9	3.2	3.4	3.5	3.6
civilian employment	1.3	1.3	1.2	1.1	1.0	1.0	1.0	1.1

Source: Compiled from McNees (1981).

One sees that except for the consumer price index, the RMSE decreased for all variables with the length of the forecast period. The error is roughly three times as large in quarterly forecasts than in two-year forecasts. The short-term forecasting of durable consumer goods and for investment in residential structures seems to to be especially difficult; the RMSE is as high as 11.2 and 18.4 percentage points respectively. In the long term, however, durable consumer goods can be predicted with a signifcantly smaller error.

The uncertainty of investments in residential structures remains high even in the case of longer forecast periods. On the other hand, the forecast accuracy of the real national product, an explanation factor for consumer spending, is significantly better. With short-term forecasts, the RMSE is 3.3 percentage points and then sinks to roughly one percentage point as the forecast period increases. A different movement is observed with the price index. The RMSE increases with the length of the forecast period. This means that not only the absolute error, but also the logarithmic error grows in the course of the forecast period.

The accuracy of the one-year forecasts of the „Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung" (council of experts for the evaluation of the total economic development) between 1970 and 1983 has been studied by Pflaumer (1986). The actual growth rates of several economic variables were compared with the forecasted growth rates. Table 8 shows the RMSE for the individual variables. Here as well, the great variations in forecast accuracy can be seen. The accuracy of some variables is unsatisfactory. These variables include investments, exports and imports, as well as income from entrepreneurial activity and savings. The RMSE of the forecast of employed persons is noticeably small. It only amounts to 0.68. Thus, it is not much higher than the RMSE determined by Keyfitz for the population forecasts. The error in the forecast of employed persons is certainly so low because, in addition to economic factors, employment is above all determined by demographic factors that can be predicted more accurately.

In order to understand the relationship between the RMSE and the forecast horizon better, the forecast accuracy of long-term economic forecasts will be discussed briefly.

Table 8

RMSE for various economic variables of West Germany (percentage points)

variable	RMSE	variable	RMSE
private consumption	1.57	gross national product	1.97
government consumption	1.96	price development of the gross national product	1.07
capital investment	5.02	employed persons	0.68
nonresidential fixed investment	6.03	gross income from dependent employment	2.23
investment in residential structures	4.77	gross income from entrepreneurial activity and savings	4.14
exports	7.00		
imports	6.38		

Source: Pflaumer (1986).

Ascher (1979) evaluated the forecast accuracy of five institutions that have been producing long-term forecasts for the real national product of the U.S.A. since 1950. The errors from 5-year, 10-year and 15-year forecasts were determined. The 5-year forecasts had 17 observation pairs available, the 10-year forecasts had 24 and the 15-year forecasts had 14. The RMSE was calculated from the data of Ascher. This is 1.9 percentage points in the case of the 5-year forecasts, 0.9 in the case of the 10-year forecasts, and 0.6 in the case of the 15-year forecasts. In order to be able to determine the forecast error for the individual components of the national product as well, forecasts from "Predicasts" between 1960 and 1975 were analyzed. "Predicasts" is a commercial publication that collects and publishes short- and long-term forecasts of all possible North American economic data from the most known institutions. Each forecast in Predicasts (publication: quarter-yearly) is given along with the source, so that the original forecast can be referred to if necessary. The RMSE for the most important components of the national product of the U.S.A. was calculated for five-, ten-, and fifteen-year forecasts from the data in Predicasts. The number of observations for one variable varies between 20 and 100. As with the short-term forecasts, a reduction of the RMSE dependent on the forecast period is determined in table 9. The results from tables 7 and 9 are not readily comparable, since the calculations originate from different periods and sources. However, the course of the errors leads to the assumption that for many economic variables, the RMSE calculated from logarithmic errors shows a decreasing course dependent on the length of the forecast period.

Table 9

RMSE for components of the gross national product of the U.S.A. (percentage points)

	forecast horizon		
	5 years	10 years	15 years
gross national product	1.8	1.0	0.7
personal consumption	1.5	0.9	0.5
durable goods	2.8	1.1	1.2
nondurable goods	1.0	1.0	0.8
services	1.5	0.8	0.5
gross private domestic investment	6.1	5.0	3.4
residential	11.2	6.0	4.4
nonresidential	2.5	2.4	1.1
government purchases	0.6	1.6	1.9
federal	0.9	3.8	3.1
state & local	1.4	0.5	1.1

Source: Own calculations from Predicasts 1960-1975.

The reason for the decreasing development of the RMSE in economic forecasts can be seen in the great cyclical behavior of the economic variables. In the past, demographic as well as economic variables have followed an exponential growth path. The economic variable varies more around the trend than the demographic one does. These short-term variations make economic forecasts relatively difficult. In the case of long-term forecasts, the short-term variations do not weigh heavily, and as with the demographic forecast, the most important thing is to correctly predict the trend. Thus, the difference in RMSE between long-term demographic and long-term economic forecasts is no longer that great.

4. CONCLUSION

A population forecast fulfils its purpose when it helps to reduce the extent of the uncertainty about future demographic development. Equally, the calculation of forecast error measures fulfils its purpose when it helps to quantify the extent of the uncertainty. Only when a forecast error measure is supplied with the publication of a forecast does the user become aware of the unavoidable uncertainty of the estimation, and it becomes possible for him to calculate the risk of his individual planning, which is partially or wholly based on the published population forecast.

As long as the forecaster does not present any convincing reasons for a reduction of the future forecast error, the user should apply the forecast error derived from the past values as an indicator of the uncertainty of a forecast. This holds true not only for population forecasts, but also for economic forecasts. Politicians, planners and businessmen would perhaps not have made many decisions that later proved to be mistaken if they had been aware of the uncertainty of their forecasts. However, what politician, planner or businessman inquires about such statistical details as forecast error measures?

5. REFERENCES

Ascher, W. (1979), Forecasting, an appraisal for policy-makers and planners, Baltimore-London.

Elster, L. (1891), Bevölkerungwesen, Handwörterbuch der Staatswissenschaften, 2. Bd., Jena, 465-528.

Günther, E. (1935), Wert oder Unwert der Vorausberechnung der künftigen Bevölkerung, Allgemeines Statistisches Archiv 25, 404-415.

Inoue, S. and Yu, Y.C. (1979), United Nations new population projections and analysis of ex post errors, Paper presented at the Annual Meeting of the Population Association of America, Philadelphia.

Jöckel, K.-H. and Pflaumer, P. (1984), Calculating the variance of errors in population forecasting, Statistics and Probability Letters 2, 211-213.

Kale, B.D., Voss, P.R., Palit, C.D. and 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, D.C.

Keilman, N.W. (1990), Uncertainty in national population forecasting: Issues, backgrounds, analyses, recommendations, Swets & Zeitlinger, Amsterdam.

Keyfitz, N. (1979), Long-range projections: models, pitfalls, possible break-throughs, United Nations, Prospects of population methodology and assumptions, Population Studies, No. 67, New York, 107-121.

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

Land, K.C. (1987), Methods for national population forecast: A review, Journal of the American Statistical Association, 81, 888-901.

McNees, S.K. (1981), The recent record of thirteen forecasters, New England Economic Review, Sept./Oct., 5-21.

Obaidullah, M. (1976), Expo-linear models for population growth, Rural Demography 3.

Pearl, R. and Reed, L.-J. (1920), On the rate of growth of the population of the United States since 1790 and its mathematical representation, Proceedings of the National Academy of Science, 6, 275-288.

Pflaumer, P. (1986), Messung und statistische Analyse der Fehler in den Prognosen des Sachverständigenrats, Allgemeines Statistisches Archiv 70, 368-387.

Pflaumer, P. (1988a), The accuracy of U.N. population projections, Proceedings of the American Statistical Association, Social Statistics Section, Alexandria, Virginia, 299-304.

Pflaumer, P. (1988b), Methoden der Bevölkerungsvorausschätzung, Duncker & Humblot, Berlin.

Pflaumer, P. (1988c), Confidence Intervals for Population Projections based on Monte Carlo Methods, International Journal of Forecasting 4.1988, 135-142.

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

Pritchett, H.S. (1891), A formula for predicting the population of the United States, Publication of the American Statistical Association (Boston), 2, 14, 22-30.

Smith, S.K. and Sincich, T. (1991), An empirical analysis of the effect of the length of forecast horizon on population forecast errors, Demography, 28, 261-274.

Stoto, M.A. (1983), The accuracy of population projections, Journal of the American Statistical Association, 78, 13-20.

Stoto, M.C. and Schrier, A.P. (1982), The accuracy of state population projections, Discussion Paper Series, Kennedy School of Government, Harvard University.

Voss, P.R. and Palit, C.D. (1981): Forecasting state population using ARIMA time series techniques, Technical Series 70-60, University of Wisconsins, Madison.

Wigglesworth, E. (1775), Calculations on American Populations, Cambridge.