ARTIKLER 61

LEVELS OF ERROR IN POPULATION FORECASTS

By Jan M. Hoem

USIKKERHETSNIVÅER
VED BEFOLKNINGSPROGNOSER

OSLO 1973

STATISTISK SENTRALBYRÅ

ARTIKLER FRA STATISTISK SENTRALBYRÅ NR. 61

LEVELS OF ERROR IN POPULATION FORECASTS

By Jan M. Hoem

USIKKERHETSNIVÅER VED BEFOLKNINGSPROGNOSER

With an Appendix by Leo Törnqvist

OSLO 1973

ISBN 82 - 537 - 0317 - 1

PREFACE

The Socio-Demographic Research Unit of the Central Bureau of Statistics of Norway is engaged in improving methods for prior calculation of the size and composition of a future population. This article presents some of the more important ideas on which the work is based. This presentation is in English. A somewhat more extensive account has appeared previously in Norwegian in Article No. 54.

The present article is supposed to appear as a chapter in the volume "Statistical Problems in Population Research", edited by Nathan Keyfitz and to be published by the East-West Population Institute, The East-West Center, Honolulu, Hawaii. The Central Bureau of Statistics gratefully acknowledges the permission of the East-West Population Institute to publish the chapter in anticipation of the future publication of the book.

One of the more interesting early contributions to the theory of population forecasts was Leo Törnqvist's section of the publication "Calculations concernant la population de la Finlande, sa reproduction et son développement futur", authored by Jorma Hyppölä et al., and published by the Central Bureau of Statistics of Finland in 1949. The publication appeared in Finnish and Swedish, with a summary in French, and it well deserves to be better known than it is. In order to make Törnqvist's ideas accessible to the readers of this article, a translation into English of central parts of his contribution to the publication appears as an Appendix here. The Central Bureau of Statistics of Norway is grateful to Leo Törnqvist for his authorization of the translation and his permission to publish it.

Central Bureau of Statistics, Oslo, 1 November 1973

Petter Jakob Bjerve

FORORD

Statistisk Sentralbyrå arbeider for tiden med å forbedre metodene til å forhåndsberegne framtidige folketall. Denne virksomheten foregår i Sosiodemografisk forskningsgruppe. I denne artikkelen framstilles noen av de viktigste idéer som ligger til grunn for arbeidet. En noe fyldigere versjon er tidligere utgitt på norsk som Byråets artikkel nr. 54.

Artikkelen skal senere bli publisert som et kapittel i boken "Statistical Problems in Population Research", som er redigert av Nathan Keyfitz og som skal utgis av East-West Population Institute, The East-West Center, Honolulu, Hawaii. Statistisk Sentralbyrå vil takke instituttet for samtykke til opptrykk.

Et av de mest interessante bidrag til teorien for befolkningsprognoser er Leo Törnqvists kapittel i heftet "Beräkningar rörande Finlands befolkning, dess reproduktion och framtida utveckling", utgitt av Statistiska Centralbyrån i Helsingfors i 1949 med Jorma Hyppölä et al. som forfattere. Heftet kom ut på finsk og svensk, med et sammendrag på fransk, og innholdet fortjener å bli bedre kjent enn det er. For å gjøre Törnqvists idéer tilgjengelige for leserne av denne artikkelen, gjengis en engelsk oversettelse av de sentrale deler av hans bidrag til heftet som et appendiks her. Statistisk Sentralbyrå vil takke Leo Törnqvist for at han har autorisert oversettelsen og gitt tillatelse til publisering.

Statistisk Sentralbyrå, Oslo, 1. november 1973

Petter Jakob Bjerve

CONTENTS

INNHOLD

"Our discipline [demography] has been greatly influenced by demand for prediction of future population and by efforts on the part of our members to satisfy that demand. The public expects predictions. When I meet someone who asks me what I do, and I reply that I work at the mathematics of population, the reaction is something like: "Great! Then you're the man who can tell me what the population of the world will be in the year 2000," or even: "When, according to your figures, will there be too many people in the United States?" To protest that I and others practising the profession of demography are interested in facts by no means escapes the question; the most interesting facts are those relating to the future.

If the questioner is experienced enough to know that prediction is difficult, he at least supposes, when he hears about demographic models, that the accuracy of prediction is a test of the models. The demographer, he might say, has every right to look into many aspects of past and present population, but he validates this right, and distinguishes himself from the ordinary run of men, by now and again doing a prediction. Like the Pharisees in the Bible who asked Jesus for a miracle, our public wants a sign from heaven that we are authorized, that we are not mere guessers like themselves. Our calculations may be esoteric, but everyone can understand a prediction that the census of 1980 will count 230,000,000 persons in the United States, and any Pharisee can check the calculation when the census total is released about August 1980. By doing this small miracle of prediction we not only help people in their practical affairs between now and 1980, but when that date comes round we will have proved our competence by an unmistakable sign. The desire to be useful, and even more the desire to prove competence, are strong. We do not reply as Jesus did to the Pharisees, "There shall be no sign given to this generation."

Rather we attempt a miracle, and the miracle fails miserably."

Nathan Keyfitz, Presidential Address to the Annual Meeting of the Population Association of America, Washington, D.C., April 23, 1971. Demography 8 (4):571-580.

LEVELS OF ERROR IN POPULATION FORECASTS

By JAN M. HOEM*

1. Introduction and Main Conclusions

1.A. The chief purpose of making a population forecast, in the view of this author, is to contribute to improved planning and to better decisions.[1] Accuracy of the prediction, in the sense that it turns out to agree well with subsequent actual population trends, cannot be the only or even the main goal towards which the makers of forecasting models should aim. Forecasts have well-known "publication effects" which influence their accuracy. On the one hand, people will be apt to behave as if the prediction will largely come true, thereby creating an effect in the direction of accuracy, while it may be desirable to plan for a different development. On the other hand, the inherent tendency of self-defeat of a forecast which really brings ill bodings to a population, does not in itself make it any less valuable as a planning instrument. (The fact that these effects work in conflicting directions does not mean that they neutralize each other.)

Yet it is important to have a reasonable degree of accuracy. Grossly unreliable forecasts are of little value for planning. One must, therefore, deplore the fact that the purely demographic population forecasting methods which are in current use, largely give unreliable "predictions" of future population trends. An improvement on this front cannot apparently be brought about solely through the development of more sophisticated demographic models. A large experience shows that more refined demographic methods have not given more accurate results than simple procedures.[2] Hope for real improvement appears to be linked with the inclusion of non-demographic factors in the models. In Norway, this will probably come about through

*Acknowledgements.

I wish to thank Morris L. Eaton, Monica Fong, Svein Longva, Thorvald Moe, Tore Schweder, and Bjørn Tønnesen for reading the present article in manuscript and for giving many helpful comments. I am also grateful to Lars Widén for giving the explanation on which the first part of Section 9.D is based.

the incorporation of the population forecasting model in a macroeconomic planning model. One central motivation for this procedure is the desire to include the influence of population policy explicitly in the calculations.

If one's primary concern is forecasting inaccuracy, one should not be too optimistic, however, as to the possible gains resulting from model development. The most important improvement in forecasting procedures will probably not be their increased reliability but their greater relevance to economic and social planning. One must also bear in mind that such a development cannot be brought about overnight; it will only come as the result of a concentrated effort over an extended period.

1.B. Population forecasts will always contain elements of uncertainty. When a forecaster publishes the results of his calculations, it is part of his task to firmly remind the users of this. To-day, this is usually done primarily through the presentation of several alternative forecasting series. Forecast producers have frequently (and in part rightfully) been criticized for being much too vague as to the reliability of the forecast, even when there are alternative series. Several authors have insisted that forecasters should bring this custom to an end, and that they should change over to specifying probability distributions for future population numbers. Much more precise statements about forecasting uncertainty would then be possible. -Less demanding authors have requested the publication of standard errors along with the forecast of future numbers.

1.C. In principle, this type of approach evidently is a goal towards which forecasters should strive and to which one may possibly find a reasonably accurate and operational solution some time in the future. The timeliness of forwarding such a request at the present time is, however, seriously open to questioning. Much too little is known about central population processes to merit the publication of probabilistic measures like standard errors of this type, to say nothing of entire probability distributions. The publication of such measures would give a quite unwarranted impression of the precision of current knowledge. It is difficult to understand what gain there could be in introducing an additional source of uncertainty, such as this would be.[3]

1.D. It is a main standpoint of this author that in countries with good data the most decisive part of such forecasting error as there is any hope to control, consists in erroneous model specification. Present-day models delete factors which seriously influence population trends, or such factors are inadequately represented. Thus, the most important immediate task facing the forecaster is to get a more adequate forecasting model.

This stand is reinforced by the fact that the derivation of the probability distribution of future population numbers on the basis of a given forecasting model is no walk-over. Formulas for the first few moments can be ground out by standard methods, but a drive to establish the entire probability distribution quickly runs into serious technical problems. In a situation where available resources are limited, this represents a further argument for giving low priority to such a line of attack at present.

1.E. Under these circumstances, it may be useful to list the kinds of elements which contribute to making population forecasts unreliable. In the present Chapter, the sources of forecasting error are grouped into three types, viz.,

I. estimation and registration errors;

II. errors due to random fluctuations;

III. erroneous trends in future mean vital rates.

These types are subdivided again into six levels altogether, the key-words for the levels being estimation and registration errors, pure randomness, random vital rates, unincorporated gradual changes in mean vital rates, gross shifts, and serious model misspecification.[4]

This Chapter describes these levels of error and discusses their consequences, thus sketching the main reasons for taking the position outlined above.

2. Preliminaries

2.A. As is well known, the advance calculation of future population figures can appear in many modes. The present paper will address itself solely to the situation where the figures are positively aimed at saying something about the future of a real population. Thus, such things as counter-predictions and calculations made for analytical purposes only will be left aside, and the analysis will concentrate on predictions/forecasts, as already indicated by the language used above.

Bearing the planning purpose of forecasts in mind, it is important to note that even predictions can appear in several modes.[5]

Following Leif Johansen (1970), we shall distinguish between a pure forecast, an indicative forecast, and a forecast explicitly incorporated in a decision-making process.

A pure forecast represents an attempt at predicting more or less unconditionally "expected" or "most probable" future trends.

An indicative forecast tries to capitalize on the tendency of becoming a self-fulfilling prophecy which is inherent in a forecast (unless it causes alarm and leads to action in the opposite direction, as indicated in Section 1.A above). This idea has been extensively used in indicative economic planning in France, but not, to our knowledge, intentionally in population prediction in spite of its obvious importance at least for regional forecasts.

In forecasting models for decision-making processes, one will distinguish between four types of elements, which are linked by the model:[6]

Firstly, there is a group consisting of the variables which are beyond the control of the decision maker, and which are not much influenced by his actions. This group of variables is amenable to pure forecasts.

Secondly, the decision maker has a set of instruments (decision variables, policy variables) which can be used to influence future trends.

Thirdly, a set of targets will be specified, and a purpose of the forecasting exercise is to show how these targets may be attained. Thus, the forecast will be partly normative.

Finally, there will be an additional category of variables whose values will be determined at least partly by decisions made, but which are not deemed sufficiently important to be included among the target variables.

Depending on the structure of the model specified, the actual forecasting may be carried out by means of an iterative procedure. It may also be put into effect sequentially through periodic revisions. What results from such a procedure would be neither a pure forecast nor fully a normative plan. Although it would have elements of both, they would be so tightly interwoven that an attempt at classification in either category would be of little value.

Evidently, the self-defeating and self-fulfilling tendencies of a forecast will come into play no matter which of these modes it is made in.

2.B. Most producers of future population figures of the kind which we take into account in this Chapter (including the Central Bureau of Statistics of Norway) will insist upon calling their commodities population projections. Their own behaviour seems to belie this contention, however. Such producers will typically make an effort to base their calculations on assumptions which are as "realistic" as at all possible, and when the actual population trends turn out to deviate from the "projection", they will produce new calculations based on updated assumptions. Although most of them will not accept this, it looks as if such forecasters really attempt to predict future population trends, but realize that they are not very good at it.

Moreover, as Keyfitz argues,[7] most users will take interest in the figures produced mainly insofar as they can be regarded as predictions, and the majority will treat them as that anyway.

It seems that most of the advance calculations actually made, can be placed somewhere in the area bordered by real projections on the one hand and pure forecasts on the other hand. Occasionally, attempts are made at taking the influence of policy variables implicitly into account through the choice of specifications of future vital rates, but in general policy-making effects are left out. This is one of the most important sources of forecasting inaccuracy as well as a major defect of forecast as a tool in planning, and we shall return to it on several occasions in later Sections.

2.C. Before going into the presentation proper of the Chapter, we conclude this preliminary section by introducing some conventions and notation. For convenience, we shall take the forecasting time unit to be a year. Time is reckoned from the beginning of the forecast, and year t is the year between time t - 1 and time t. Let us use the notation X(t) for the vector of the actual population at time t. The forecaster is required to produce a prediction $\tilde{X}(t)$ for X(t) for t = 1,2,...,T . Present forecasting models will typically calculate the forecast recursively, so that $\tilde{X}(1)$ is calculated first, the result is used to calculate $\tilde{X}(2)$, and so on. Forecasts of a national population will typically be produced through a linear model. If we disregard international migration, this means that one will use a forecasting relation of the form

$$\tilde{X}(t) = M(t)\,\tilde{X}(t-1)\ , \quad \text{for } t = 1,2,\ldots,T\ . \qquad (2.1)$$

Here, M(t) is a projection matrix. As is implied by the notation, it may depend on time. If international migration is taken into account, the net number of immigrants must be added to this expression.

Nobody seems to have found satisfactory methods of accounting for international migration, however.

Forecasting models based on cohort ideas can usually be written in the form given in (2.1). So can regional projection models based on (area-) component methods.[8] While much of the literature on population prediction in effect limits itself to national forecasts, the present discussion will cover prediction of the regional distribution of the population as well. The restriction to the national level sometimes covers up some of the complexities involved.

For simplicity and concreteness, most of the account below will be tied formally to linear forecasting models for closed populations. It is easy to see how much of it can be transferred to other situations.

3. Sources of Forecasting Inaccuracy

3.A. This Section contains an overview of the various sources of inaccuracy, grouped into three types, which are subdivided again into six levels altogether. The types are numbered I, II, and III. The levels are numbered consecutively from 1 to 6. Further consideration is given to them in later Sections.

Type I. Estimation and registration errors

Level 1: None of the parameters of the forecasting model are really known to the forecaster. Statistical estimates must be calculated from available data. This gives rise to a series of sources of error. We shall reckon with

(a) estimation variance (which is due to the fact that the data will be regarded as a sample);

(b) registration errors giving defective data for the parameter estimates;

(c) errors in size and composition of the initial population $X(0)$.

Such errors are propagated through the entire forecasting period. The effect is similar to that of

(d) rounding errors;

and these will be included on Level 1.

Type II. Errors due to random fluctuations

Level 2: Pure randomness. Even if the survival probabilities in force during a given forecasting year were known, the proportion of survivors in each population group would not be exactly equal to this probability. There would be some random variaton. The same is true for births, and so on.

Level 3: Random vital rates. Pollard (1968) and Sykes (1969) have criticized models assuming that the elements of the projection matrix M(t) are non-random parameters. Pollard mentions that mortality rates depend on weather conditions: a hard winter will cause increased mortality, in particular for the old and the very young, and, conversely, a mild winter will give rise to lower than normal mortality.[9] Sykes (1969, p. 118) asserts[10] that "natural and social phenomena such as droughts, epidemics, revolutions, and the like, would result in substantial departures from the mean performance in births and deaths." Thus, both of them seem to think in terms of a kind of mean development of the projection matrix, with superposed fluctuations. The mean development would then be represented by the expectation EM(t) of the projection matrix M(t) , while the fluctuations would be measured, among other things, by the covariance matrix $\Sigma(t)$ of the elements of M(t).

This seems to agree well with statements given in the discussion during the Honolulu Symposium, where Paul Meier, in particular, emphasized that one ought to study fluctuations in vital rates as they have occurred during periods for which one has observations, and that such fluctuations should be built into population projection models.

Type III. Erroneous trends in mean vital rates

Level 4: Unincorporated gradual changes. Society changes and there is a corresponding gradual change in mean fertility, mortality, and so on. The difficulty of predicting such changes with sufficient accuracy represents a further source of uncertainty in the forecasts.

Level 5: Gross shifts in mean vital rates. In connection with certain major events, such as wars, serious economic depressions, break-throughs in medical techniques, and major changes in population policies (such as abortion practices), vital rates may get a sudden shift to a new level. Whichever type of model is used, it can be difficult to foresee when such a shift may possibly come and how important it will be, even in the immediate future.

Level 6: Serious model misspecification. In present-day forecasting models, important factors get left out or are specified quite incorrectly. For example, the lack of explicit attention given to policy variables gives rise to errors on this Level.

3.B. Although we have suggested typical causes for some of the error levels above, the classification is essentially one of effects on vital rates. Except possibly for Level 1, the classification is intended to convey an impression of increasing seriousness of these effects, which constitute the sources of forecasting error.

Both Type II and Type III errors are really kinds of model errors, and one may have both unincorporated gradual changes, gross shifts, and misspecification at least of Level 3 errors. Therefore, no Type or Level is called "model errors."

All of these kinds of errors can occur no matter which forecasting mode one applies. Thus, this discussion is relevant to them all.

Eaton (1971) has classified forecasting errors somewhat differently. It appears that he would call Levels 1, 6, and possibly 4 initialization errors, while Levels 3, 5, and probably 2 would be future errors.

Keyfitz (1972) has suggested yet another classification. His component 1 corresponds to our Level 2. His component 2 corresponds to our Levels 1 and 4. His component 3 probably corresponds to our Level 3. Finally, his components 4 and 5 seem to correspond to our Level 6. He does not give separate attention to gross shifts (our Level 5).

4. Estimation and Registration Errors

4.A. The statistical estimation of model parameters will give rise to estimation error. As far as is known, nobody has actually carried out variance calculations to study the effect of this on the forecast, but it should be possible to do so by known methods. Haggstrom (1971) mentions that projections made for U.S. university enrolment are highly sensitive to small changes in parameter values when carried a large number of years into the future. Parameter variability has effects which accumulate as the projection period progresses.

4.B. Rounding errors surely build up in an entirely corresponding way, and may get a certain influence after a number of projection years. Goodman has given some consideration to the importance of rounding errors in one of his papers (1968), but otherwise this does not seem to have worried people who have written about forecasting. We may probably interpret this as signifying that such errors are much less problematic than other sources of inaccuracy.

4.C. The quality of the forecast depends on the quality of the data. A considerable part of the litterature on demographic methods is devoted to the question of what one should do when faced with defective data. Naturally, much of this material primarily considers problems concerning data for developing countries, but problems tormenting United States forecasters figured prominently during the discussions in Honolulu.

5. Level 2 Errors (Pure Randomness)

5.A. Assume that the survival probabilities, birth probabilities, and so on, which really are in force in a given year, were specified at the beginning of the year. Seen through the eyes of the probabilist, deaths, births, and so on, registered during this year, represent the outcome of a series of random "experiments". The variability of such an experiment can be measured, and this variability gives rise to Level 2 unreliability.

Pollard (1966), Sykes (1969), and Schweder (1971, 1972) have studied this phenomenon by means of branching process theory, and they all agree that Level 2 errors account for only a small part of the total unreliability of population forecasts. Before we go on to consider errors on higher levels, we shall mention one of Schweder's results, however. To be sure, it is not so useful in its present form because it is bound to the Level 2 errors, which have small importance, but it will be of considerable help if extension to models incorporating errors on higher levels proves possible.

5.B. Like Pollard and Sykes before him, Schweder bases his reasoning on the classical Leslie model of population dynamics and then adds "binomial" random variation. The initial population $X(0)$ is regarded as given, and so is the projection matrix M, which is taken to be independent of time. $X(t)$ becomes a random vector with some covariance matrix $C(t)$. As a forecast for $X(t)$, one uses

(5.1) $$\tilde{X}(t) = M^t X(0) ,$$

which gives

(5.2) $$\tilde{X}(t) = EX(t) .$$

This suggests that $\tilde{X}(t)$ is an appropriate forecast. One wishes to make statements concerning the discrepancy $X(t) - \tilde{X}(t)$ which must be expected between the actual population vector and the forecast. Let β be an arbitrary confidence level, $0 < \beta < 1$, and let δ be the β percentage point in the χ^2 -distribution with a number of degrees of freedom equal to the number of elements in X(t) . Schweder shows, among other things, that a large X(0) will make X(t) approximately normally distributed, and that

$$P\{\bigcap_i [\tilde{X}_i(t) - (\delta C_{ii}(t))^{\frac{1}{2}} \leq X_i(t) \leq \tilde{X}_i(t) + (\delta C_{ii}(t))^{\frac{1}{2}}]\} \geq \beta ,$$

which means that there is an approximate probability of at least β that the number of persons $X_i(t)$ in each population group i at time t will lie between the bounds

$$\tilde{X}_i(t) \pm [\delta C_{ii}(t)]^{\frac{1}{2}} ,$$

simultaneously for all population groups. A choice of β equal to some number like 0.8, 0.9, or 0.95, would make the upper bound here a high forecasting value and the lower bound a low forecast for $X_i(t)$.

6. Level 3 Errors (Random Vital Rates)

6.A. As mentioned above, Pollard (1968) and Sykes (1969) have suggested that one should regard the population projection matrix as random. The matrices M(1) , M(2) , ... are regarded as a sequence of random matrices determined by natural and social mechanisms, and the projection matrices then act upon the population one after another, thus producing the transition from the beginning of a year to the beginning of the next one.[11] (This is essentially the same idea as that due to Smith and Wilkinson,[12] who study a one-dimensional Markov chain which they call a branching process in a random environment.)

In these papers, both Pollard (1968) and Sykes (1969) restrict themselves to a situation where the matrices are stochastically independent. If we regard EM(·) as known, the forecasting calculations can be carried out recursively, viz., through

$$\tilde{X}(t) = EM(t)\ \tilde{X}(t-1) \quad \text{for } t = 1,2,\ldots,T ; \tag{6.1}$$

and we still get (5.2) because

$$E \prod_{s=1}^{t} M(s) = \prod_{s=1}^{t} EM(s). \tag{6.2}$$

6.B. Even though Sykes also gives some formulas for the general case, where EM(t) and Σ(t) may depend on t , most of the attention has been devoted to a situation in which there is a stationary level for the matrices as well as for their variability, i.e.,

(6.3) $$EM(t) \equiv M$$

for a suitable matrix M , and $\Sigma(t)$ is independent of t . Schweder and Hoem (1972) have also studied this case, and have carried out calculations to get an impression of numerical consequences for this "doubly stationary" model. The values for M and Σ were based on Norwegian data from the years 1953-1968. The main conclusions were as follows:

(i) Mortality fluctuations at ages below 50 do not greatly influence accuracy of population forecasts in a country like Norway, and the model reflects this in a satisfactory way.

(ii) As a consequence of the stationarity built into the model, the unreliability of the forecast of the number of births, as measured by the model, will stay on approximately the same level through the first sixty forecasting years (which was as far as our calculations went). This is reasonable if the variability in this model is regarded as the contribution to the total unreliability of the birth forecasts which is due to random fluctuation in the projection matrix around a given level.

(iii) On the other hand, the model turned out to imply a very high level of unreliability of the forecasts of births in the first forecasting years.[13] The level is so high that it either indicates that such birth forecasts are seriously unreliable, or else this model has overestimated the forecasting inaccuracy.

6.C. The latter of the alternatives under point (iii) above appears to be the more plausible. The fertility level in Norway really was not constant over the years from 1953 to 1968, and the changes which did occur, were not due only to random variation of the type on which Pollard and Sykes base their argument for random projection matrices. There has been a genuine development in EM(t) over these years. When the covariance matrix is estimated from the variation in the vital rates around a mean level for the observational years, one will have added the variation of EM(·) around this mean to the fluctuations of M(·) around EM(·). Thus, our measure of variability will be too large, consequently inflating our estimate for the part of the unreliability of the birth forecasts which is due to random variation in M(·) .

It is not really adequate to use a constant projection matrix throughout the forecasting period either.[14] When the forecast is made, one may have information supporting a certain time-trend in EM(t) . (Compare our comment on cohort methods below (2.1).) The use of (6.3) means that one omits utilizing such information.

It is easy to change the model to take these factors into account. One is still left with the necessity of determining the long-time trend in EM(t) in a non-arbitrary way, and this is one of the more important sources of inaccuracy of birth forecasts. At this point, however, the problems really belong on higher levels of error than the present Level 3, so their discussion will be postponed to later Sections.

6.D. The assupmtion that the projection matrices are uncorrelated, may be another weakness in the above models. In a later paper, Pollard (1970) has suggested that one may make them stochastically dependent,[15] and he introduces a simple second-order autoregressive model where the rate $\delta(t)$ of growth of the total population is described as a particular linear combination of $\delta(t-1)$ and $\delta(t-2)$. The model accounts for the total population only, and Pollard writes (p. 209) that the analysis becomes quite complicated if one attempts to introduce an age structure.

Evidently, any serviceable population forecasting model must contain some age structure. In addition to the mathematical complications which result from this, one will also get practical calculation problems if one makes the projection matrices stochastically dependent. It is not easy to see how one might calculate a forecast $\tilde{X}(t)$ recursively and still obtain (5.2) if the matrices are correlated.[16]

If some probabilistic model is prescribed for the M(t)-s , it is, of course, theoretically possible to make the forecast through simulation. The practical problems involved seem considerable, however. Since existing knowledge of realistic stochastic models for these matrices is very vague, extensive simulation experiments will most likely be necessary both to establish a model and later to make the forecast. Corresponding resources could probably be put to better use at other points of forecasting theory.

6.E. It has not been possible so far to extend the elegant results demonstrated by Schweder[17] to the case where the projection matrices are random. The problem is that it is too difficult to derive the probability distribution of $\pi_{s=1}^{t} M(s)$.

7. Type III Errors (Erroneous Trends in Mean Vital Rates)

7.A. (Level 5.) Occasionally, one may have a reasonable possibility of foreseeing that a gross shift in vital rates may occur in the near future, and one may then try to take this into consideration when forecasting if one's technical preparedness is sufficiently high. Usually, it will be quite impossible to account for such matters when a forecast is made, however. Since it is generally agreed that the forecaster does not have occult powers, one should not expect him to predict events of this kind, to say nothing of what effects they will have on population trends. Unfortunately, it happens that critics forget this in hectic moments.

In the formal theory, as we have described it above, such sudden events are reflected in shifts in $EM(t)$, and possibly in other characteristics also, such as in the covariance matrix $\Sigma(t)$ of $M(t)$.

7.B. (Level 4.) On the other hand, the forecaster is expected to make allowance in his calculations for a gradual change in $EM(t)$ (and possibly also other characteristics), and he is expected to do so better than others. Even though it is a gradual developmen that he is supposed to foresee, he is faced with a whole spectrum of possibilities in the exact specification of future trends, and trial calculations with alternative specifications, all of which seem reasonable and realistic, will usually result in population trends which are noticeably different from another. There is, therefore, considerable forecasting uncertainty on this Level as well.

There are at least two ways of explicitly allowing for this source of inaccuracy:

(i) The classical procedure is to produce and publish several alternative forecasting series. An attempt is made to make them represent reasonable and realistic future trends, and at the same time provide some impression of the range within which the figures reasonably can be expected to lie.

(ii) The other line of attack consists in further developing the ideas described in Section 6:[18] Some mean trend in $EM(t)$ may be specified. The uncertainty of forecasts of births and similar factors may be taken into consideration through a suitable specification of the probabilistic mechanism generating the random matrices M(t). The unreliability of the forecast will increase as one progresses into the forecasting period. One may account for this by letting at least the diagonal elements in $\Sigma(t)$ increase with t , since these elements represent the variances of the elements of M(t) .

Both of these procedures require the forecaster to decide upon medium future trends in population components, and in this respect they are similar to each other. They differ in their treatment of forecasting uncertainty.

A forecast made according to produce (i) above will typically be presented with relatively vague statements concerning the prospective accuracy of the alternative series.[19] Considering common experience with forecasting accuracy, this attitude is understandable. In principle, the alternative assumptions on which the series are based, are selected in the light of what expert opinion considers reasonable and probable. The deliberations involved are a long distance away from what would be needed to warrant an interpretation in probabilistic or confidence type terms, however. To give such an interpretation would be to overly dignify our notions of the possible span of future trends through suggesting that we can make precise, quantitative statements far beyond what present-day methods permit. On the other hand, many forecasters may certainly be criticized for not having made greater efforts to explain what their series should be taken to mean.

In principle, the second procedure indicated above makes statements possible on a quite different level of preciseness. As a minimum, standard errors for all figures forecast can be calculated, for example for the predicted number of births in each forecasting year.[20] If the probabilistic model for the M(t)-s is sufficiently specified and if the mathematical problems can be solved, one may also find the probability distribution of $X(\cdot)$. In this case, <u>this probability distribution will be the real forecast</u>. On its basis, one may for instance calculate prediction regions, and one may put statistical decision theory to good use. In theory, therefore, procedure (ii) above gives results which one should decidedly aim at. The account below will bring out the basis for scepticism as to the possibility of carrying out such a program today, however.

7.C. Any forecast will be based on the assumption that something is kept constant, whether this is the projection matrices themselves, their rates of time-change, or something else. Let us symbolize this "something" which is kept constant by a parameter vector Θ. A specification of trends in population components during the forecasting period then actually consists in specifying a parametric matrix function $m(t;\Theta)$ and letting[21]

$$EM(t) = m(t;\Theta) \ .$$

Level 4 errors result from a somewhat erroneous value of Θ, or from a bit of deviation of m from what it should really have been. (If the discrepancy in Θ is due to the observations, such as if it is an estimation error, it belongs on Level 1.) Level 5 unreliability is due to gross shifts in Θ or to substantial changes in m . Level 6 errors arise because important factors are entirely absent in this specification, or because they have been included in a quite incorrect manner. To take account of missing factors may mean to further partition the population into subgroups, in a way similar to what will usually happen when a more sophisticated purely demographic model is substituted for a simpler one. On the other hand, an extension of the forecasting model may entail something much more radical, such as the introduction of non-demographic forecasting variables and of policy variables.

7.D. (Level 6.) Inter-regional migration and international migration are prime examples of demographic phenomena which are inadequately treated in present forecasting models. Common model relations for internal migration accord badly with current knowledge, and they make no explicit allowance for policy implications. Level 6 deficiencies are strikingly evident. To the extent that international migration is accounted for at all, assumptions commonly seem unreasonably arbitrary and often have the character of calculation examples more than anything else. Computations made in different countries are not harmonized in any way. Thus, the number of out-migrants each year for all countries taken together does not equal the number of in-migrants, so one has not even been able to secure elementary consistency.[22]

Similarly, present-day knowledge does not permit us to take the publication effect explicitly into account in the forecasting model.[23] (Of course, the publication effect is in part a consequence of the way in which the forecast is presented, as will be discussed in Section 9 below.)

7.E. A main viewpoint of this Chapter is that Level 6 errors are so important for population forecasts that time and additional resources in future work in this area should primarily be concentrated here. There is not much the forecaster can do about gross shifts.

Type II fluctuations have less influence, comparatively speaking. Estimation and registration errors come into the same category in a country like Norway; and in other countries, where they represent a serious problem, people are devoting much time to it. Finally, much of Level 4 unreliability is really due to unsolved problems on Level 6.

Even if one takes this stand, it may have some interest to take a look at what could be achieved if Type III unreliability were non-existent. Turning to this question in the next Section, we shall argue that a drive in such a direction, desirable as a solution clearly is, is apt to quickly run into serious technical problems of its own, something which represents a further argument for letting this line of attack rest for the time being.

8. A Further Discussion of the Probability Distribution of the Population Vector.

8.A. Let us now reason as if Type III inaccuracy were non-existent. As Keyfitz (1972) explains, Muhsam (1956) has suggested how knowledge of the probability distribution of the future population can help economic and social planning through application of the user's loss function. (Muhsam repeated his ideas in a later paper (1967) with a different but similar example.) Muhsam's suggestion actually amounts to utilizing the probability distribution

$$F(y,t_0) = P\{Y(t_0) \leq y\}$$

of the total population $Y(t_0)$ at some future time t_0 to develop some operative forecast $\tilde{Y}(t_0)$ for planning purposes. (Note that the function $F(\cdot,t_0)$ is now the real forecast submitted by the statistician.) If L denotes the loss function as usual, let us use the name risk-minimizing operative forecast for the value $\tilde{y}_0$ which makes

$$R(\tilde{y}) = \int_0^\infty L(\tilde{y},y)\, F(dy,t_0)$$

a minimum. $\tilde{y}_0$ will depend on L as well as on F, and therefore will be specific for the individual user.

8.B. This is clearly a useful way of applying forecasts. In many cases, the forecaster may be interested in the number of persons in a single population group (without splitting it into subgroups). Of course, this need not be the total population, as in

Muhsam's example; it may be births or something else.

This type of example represents an overly simplistic situation, however, where many problems do not surface. When such an example is presented without an extensive discussion of its limitations, one gets too optimistic a view of the possibility of directly applying these ideas in practice. Let me mention some commonplace points.

Firstly, water reservoirs, schools, and other items of the infrastructure which a society builds, should not only be suitable for the population at a given future time t_0; they must do service over a large number of years, frequently much longer than until the end of the forecasting period. Instead of concentrating on the population at a given time t_0 , the planner must take into consideration the population over an extended period.

Secondly, plans frequently cannot be based upon total population or the size of a single population group. One will need to know how this population is distributed over several subgroups, such as sex, age, and residence. (Just think of school locationing, the building of correctional institutions, social service planning, and so on.)

To cover all these possibilities, the loss function should depend on the entire vector $X(t)$ for all t, and on the operative forecast $\tilde{X}(t)$ for all t . The probability distribution which should be specified, is the simultaneous distribution P for $X(1)$, ... , $X(T)$. The risk to be minimized is

$$R(\tilde{X}) = \int L(X,\tilde{X})\ P(dX) .$$

Practical application of this theory is possible only if three problems are solved, viz.,

(i) establishing the probability distribution P ;

(ii) finding the loss function L ;

(iii) minimizing the risk R.

The third of these problems does not require any understanding of factual matters. "Nothing more" than knowledge of mathematics and numerical analysis is necessary, so the job seems suitable for contracting out to consultants.

Problems (i) and (ii) deserve some further consideration.

8.C. It is important to understand the dimensions of the problem one is up against when one wants to establish the probability distribution P , even if Type III errors are disregarded. Some authors who advocate the use of P , seem to take this too lightly.[24,25]

There is a lot of leeway for subjective considerations of future population trends in the theory which we sketched in previous Sections. For example, there is much scope for such elements in the specification of the probabilistic mechanism generating the random matrices M(t) , which we touched upon in Section 7.B. In principle, the distribution of the population vectors X(1) , X(2) , ... can be derived from this specification. One must proceed in such an orderly manner to get an acceptable result. It is quite unsatisfactory to make a straight jump and arrive at P without explicitly taking into account the way in which X has been generated, the way Agnew (1972) does, for instance.

8.D. Both the specification of such mechanisms and the mathematical derivation of the probability distribution of the population vectors are quite problematic, however, and it seems doubtful whether it is possible with present-day knowledge to arrive at a P in an operational form which can be used in practical planning.

It may be of some interest to note that an attempt in this direction has been made by Leo Törnqvist. Through his contribution, the 1949 Finnish population predictions (Hyppölä et al., 1949) contained a "high," a "medium," and a "low" series, with a view to the "high" and the "low" series constituting an 80 per cent prediction interval for the actual population growth. It is not quite clear exactly what this was really intended to mean, but it looks as if there was supposed to be a probability of 0.8 that the actual number of births in all forecasting years would fall between the "high" and the "low" prediction bounds.[26] (The probability of 0.8 may also have been meant to cover other events.)

Reporting on the outcome up to 1965, Törnqvist (1967) showed that the forecast from 1949 had been quite inaccurate, and that both total population and births had been above the upper prediction bound each year since 1950. It is worth noting that the groundwork for this forecast seems to have been much more thorough than in many other cases.

The prediction coefficient of 0.8 given by Törnqvist should be regarded as a subjective estimate. He made extensive studies of the variation in population parameters, but he gave up specifying detailed probabilistic mechanisms and deriving consequences for future population numbers because a staggering amount of work would have been required.

8.E. Problems similar to these arise when one wants to find the loss function L. Public population forecasts are wanted by a highly heterogenous group of users. Only a very small number of them can be expected to know what a composite probability distribution like P is, or indeed, what a loss function is. For the major part, one cannot even expect to be able to explain these concepts. There will be, at best, a small circle of important users who will be able to cooperate on this level, perhaps mainly planning divisions in the ministries and a few, centrally placed, independent consultants. Typically, however, such users will not only have as their task to give dimensions to single projects of the type one finds in Musham's examples. Quite on the contrary, they will usually apply the forecasts in a complex planning context during assessment of future development of large sectors of society. It is hardly probable that such users will be able to reduce even the most important of their considerations to a loss function with a form so operational that it can be applied for the minimizing of numerical risk. One of the problems facing those who try this, is that they would need to have public preferences made precise in a degree far beyond what public governments have been willing and able to do so far.

This problem is similar to the one facing economists wishing to establish an operational welfare function for Society. Attempts in this direction have not been singularly successful. Nevertheless, the concept of a welfare function has been highly fruitful in economics. It is possible that the loss function could play a similar role in forecasting theory, but its direct applicability in practical forecasting work should not be overrated.

9. The Presentation of Population Forecasts

9.A. When population forecasts are so unreliable, the presentation to the public is highly important. The forecaster is expected to have a better understanding than others of future population trends, and the users will rightly want him to give sufficient advice to enable them to use his product. This is made difficult by the heterogeneity of the user group. Preferably, the forecaster should present his results differently to different categories of users. The form of the presentation will depend also on how far model work has progressed and on resources available. In an ideal situation where the most important model errors have been eradicated and where one has established the probability distribution P for the population vectors with satisfactory approximation, the really competent users might be serviced by direct interaction with the forecaster and the use of a system of computer programs, much in the way short term economic planning is currently carried out in Norway.

People who are familiar with concepts as esoteric as the difference between simultaneous prediction regions and corresponding marginal prediction intervals, surely will have no difficulty in keeping forecasting unreliability clearly in mind even if they are presented with a single operative forecast, whether this is a straightforward expected population EX(•) or some more sophisticated risk-minimizing operative forecast. Nevertheless, they would find a listing of some percentiles of the distribution P helpful. Part of the interpretation of the prediction would therefore consist in making tables of particularly interesting percentage points of P . These may come from the one-dimensional marginal distributions generated by P , or, preferably, they may be simultaneous percentage points similar to those of Section 5.B.

9.B. Suggesting these possibilities is looking quite some time ahead. Even if a program of this type could be carried out, it would probably never be accessible to the great majority of users. For one thing, they have little need for such a complex offer; for another, they would really rather have other elements than what this system can give (e.g., a finer regional specification); and finally, most users would not be able to utilize the system. Therefore, there will always be a need for a presentation in a more conventional form. Since it is so important to remind the user continuously of the unreliability of forecasts, such a presentation for the general public probably ought to have the form of several forecasting series, much in the way commonly accepted today. There will be the advantage that the series will be based on a better theory than today's, and they will represent something much more precise than present-day series. (For example, they may represent series of percentage points of P .) Still, there should be several series, not a single one.

9.C. It is sometimes contended that economic and social planning in a country or a region ought to be based on a single forecasting series because one would otherwise risk having different institutions base their plans on dissimilar expectations of population trends.[27] This fear is perhaps due to a somewhat exaggerated opinion of the effectiveness of the planning process as well as of the reliability of population forecasts. The main point must surely be that planning should be flexible so that it can be adjusted to changing circumstances,[28] and then other deliberations than the fear of dissimilar expectations about the future must be decisive for the choice between publishing several forecasting series and a single one. If planning is so strongly tied to a single series that it is not flexible in this way, the very argument which leads to the advocacy of a single series should make one expect another set of problems. Just assume that the single alternative on which all planning has been based, turns out to underestimate future population growth. One would then expect capacity problems to occur simultaneously in a number of places, something which could have been avoided if some people made their plans according to a larger expected population growth than others.

To the extent that the alternative series give the planner an impression of the sensitivity of his target variables to population forecasting errors, he may try to compensate for such unreliability through his policy recommendations. The consistent publication of a single series may prevent him from taking advantage of this possibility, something which would constitute a disservice to serious planning work.

The use of a single forecasting series would also be contrary to the philosophy underlying Muhsam's proposal (Section 8.A), which, roughly speaking, is that the individual user should "play safe" in his assumptions concerning future population. This is a part of his suggestion which is very reasonable.

9.D. Any promotion of the idea that the user should be able to concentrate on a single series, is unfortunate. It distracts attention from forecasting unreliability, which is undesirable. In this connection, the experience of the Swedish National Central Bureau of Statistics has some curiosity value, and it can serve as a warning against trying one's hand with, say, three forecasting series, because this will induce the lay user to jump to the middle alternative.

The Swedes used to compute forecasting series with three alternative sets of fertility assumptions and to publish all series. They were then pestered by inquiries as to which alternative people were supposed to use, however, and they got involved in long discussions which "always" ended in the choice of the middle series. To spare themselves from this, they now publish the middle series only, although they still calculate more series.[29]

The U.S. Bureau of the Census has chosen a different solution. By 1953 it became customary for them to publish four series of population projections in each set, thereby avoiding any middle series

(Siegel,1972). This policy has later been adopted by others, including the Central Bureau of Statistics of Norway (1972).

9.E. According to this line of thought, any demand that the forecaster should enable the user to decide which single set of alternative assumptions he finds most relevant,[30] or even that he should "express, no matter how tentatively, his own judgment as to which series to choose" (Schmitt, 1971, p. 8) becomes seriously questionable. True enough, the population statistician ought to have a better grasp of emerging trends than the lay user, but this should also mean that the statistician can best appreciate the unreliability involved. He would, therefore, be shirking his ultimate responsibility.whenever he invited the user to disregard it.

When the Central Bureau of Statistics of Norway, for one, publishes a forecast with four alternative series, this means that in our opinion, future population trends seem to fall somewhere in the area mapped out by the alternatives presented. We could have published many other alternatives, but we do not do so, because we regard them as unrealistic. On the other hand, we are genuinely in doubt as to future trends, and we try to make the user take all four series into consideration simultaneously.

References

[1] Agnew, Robert A. (1972). "Crude confidence interval estimates for future U.S. population levels." Letter to the Editor, The American Statistician 26 (1):47-48.

[2] Athreya, Krishna B. and Samuel Karlin (1971). "On branching processes with random environments." Ann. Math. Statist. 42 (5):1499-1520 and (6):1843-1858.

[3] Brown, B.W. and I.R. Savage (1971). "Statistical studies in prediction of attendance for a university." Unpublished manuscript.

[4] Central Bureau of Statistics of Norway (1972). "Population Projections [for Norway] 1971-2000."

[5] Cox, P.R. (1966). "Projections of population size and age-sex structure. Statement by the Rapporteur." Pp. 253-258 in Volume 1 of [42].

[6] Danmarks Statistik (1970). "Befolkningsprognoser 1970." Statistiske Undersøgelser, Nr. 27.

[7] Eaton, Morris L. (1971). "Error and variance in the population prediction process." Preprint 1971, No. 9, Institute of Mathematical Statistics University of Copenhagen. To appear in This Volume.[31]

[8] Economic Commission for Europe (1971). "Meeting on Demographic Projections, Geneva, March 1971. Report." Document Conf. Eur. Stats./WG.35/5.

[9] Goodman, Leo A. (1968). "An elementary approach to the population projection-matrix, to the population reproductive value, and to related topics in the mathematical theory of population growth." Demography 5 (1):382-409.

[10] Haggstrom, Gus (1971). Letter of August 18, 1971, to William Kruskal.

[11] Hoem, Jan M. (1973). "Usikkerhet ved befolkningsprognoser." Central Bureau of Statistics of Norway, ARTIKLER , No. 54.

[12] Hyppölä, J.; A. Tunkelo; and L. Törnqvist (1949). "Beräkningar rörande Finlands befolkning, dess reproduktion och framtida utveckling." Statistiska Centralbyrån, Helsingfors. Statistiska Meddelanden, Nr. 38. (In Finnish and Swedish, with a Summary in French. A translation into English of central parts of Törnqvist's contribution appears as an Appendix to this article.)

[13] Johansen, Leif (1969). "Ragnar Frisch's contribution to Economics." Swedish J. Econ. 71:302-324.

[14] Johansen, Leif (1970). "Prinsipielle betraktninger om prognoser og langtidsplanlegging." University of Oslo, Institute of Economics, Memorandum dated May 12, 1970.

[15] Keyfitz, Nathan (1970). "Projection and prediction in demography: A review of the state of the art." Paper presented to the IUSSP Meeting in Mexico City, August 1971.

[16] Keyfitz, Nathan (1972). "Statisticial issues in projection and prediction." This Volume.[31]

[16a] Keyfitz, Nathan (1972 a): "On future population." J. Amer. Statist. Ass. 67:347-363.

[17] Muhsam, H.V. (1956). "The utilization of alternative population forecasts in planning." Bull. Res. Council Israel 5 (2-3): 133-146.

[18] Muhsam, H.V. (1966). "Projections of urban and rural population, economically active population, households and families. Statement by the Moderator." Pp. 272-279 in Volume 1 of [42].

[19] Muhsam, H.V. (1967). "The use of cost functions in making assumptions for population forecasts." Pp. 23-26 in Volume 3 of [42].

[20] Pollard, A.H. (1970). "Random probability fluctuations and the binomial hypothesis." J. Inst. Act. 96 (2):251-264.

[21] Pollard, J.H. (1966). "On the use of the direct matrix product in analysing certain stochastic population models." Biometrika 53 (3-4):397-415.

[22] Pollard, J.H. (1968). "A note on multi-type Galton-Watson processes with random branching probabilities." Biometrika 55 (3):589-590.

[23] Pollard, J.H. (1970). "On simple approximate calculations appropriate to populations with random growth rates." Theor. Popul. Biol. 1 (2):208-218.

[24] Savage, I.R. and B.W. Brown (1960). "Methodological Studies in Educational Attendance Prediction." University of Minnesota. Department of Statistics.

[25] Schmitt, Robert C. (1971). Invited contribution to the discussion in the Honolulu Symposium.

[26] Schweder, Tore (1970). "Composable Markov processes." J. Appl. Prob. 7 (2):400-410.

[27] Schweder, Tore (1971). "The precision of population projections studied by multiple prediction methods." Demography 8 (4): 441-450.

[28] Schweder, Tore and Jan M. Hoem (1972). "The unreliability of population forecasts; numerical illustrations based on Norwegian data." Central Bureau of Statistics of Norway, Working Paper IO 72/1.

[29] Siegel, Jacob S. (1955). "Some aspects of the methodology of population forecasts for geographic subdivisions of countries." Pp. 113-133 in Volume 3 of "Proceedings of the World Population Conference (Rome) 1954." U.N. Publication, Sales No. 1955. XIII. 8.

[30] Siegel, Jacob S. (1967). "Some principles and methods of projections of urban-rural population by age and sex." Pp. 91-96 in Volume 3 of [42].

[31] Siegel, Jacob S. (1972). "Development and accuracy of projections of population and households in the United States." Demography 9 (1).

[32] Smith, W.L. and W.E. Wilkinson (1969). "On branching processes in random environments." Ann. Math. Statist. 40 (3): 814-827.

[33] Smith, W.R. and W.E. Wilkinson (1970). "Branching processes in Markovian environments." The Consolidated University of North Carolina, Institute of Statistics, Mimeo Series No. 657.

[34] Sykes, Zenas (1969). "Some stochastic versions of the matrix model for population dynamics." J. Am. Statist. Ass. 64: 111-130.

[35] Thompson, Jean (1970). "The growth of population to the end of the century." Social Trends 1:21-32.

[36] Tinbergen, Jan (1966). "On the Theory of Economic Policy. (Fourth Printing.)" North-Holland Publ. Co., Amsterdam.

[37] Törnqvist, Leo (1967). "The post-war population development of Finland compared with predictions made after the War." Pp. 44-46 in Volume 3 of [42].

[38] U.N. Manual III (1956). "Methods for Population Projections by Sex and Age." U.N. Publication, Sales No. 1956. XIII. 3.

[39] U.N. Population Division (1971). "Population projections by sex and age and the need for their co-ordination." Paper presented to the ECE Meeting on Demographic Projections, Geneva, March 1971. Document Conf.Eur.Stats./WG. 35/2.

[40] U.S. Bureau of the Census, Current Population Reports, Series P-25, No. 448, "Projections of the population of the United States, by age and sex (interim revisions): 1970-2020." U.S. Gov't Printing Office.

[41] Working Group on Social Demography (1970). Report from its Fourth Meeting, Strasbourg, June 1970. SOA/ESDP/1970/1.

[42] "Proceedings of the World Population Conference 1965." Volumes 1 to 3. U.N. Publication, Sales Nos. 66. XIII. 5-7.

Added in proof:
A significant contribution to the general theory of population forecasting has been given in the following two review papers, which cover many aspects not discussed in the present article.

Johnston, Denis F. (1967): "Long-range projections of labor force." Mimeographed. U.S. Bureau of Labor Statistics.

Johnston, Denis F. (1970): "Forecasting methods in the social sciences." Technological Forecasting and Social Change 2 : 173 - 187.

Footnotes

1 A similar view is held by Cox (1966), pp. 256-257, and by Siegel (1972).

2 An extensive review of these matters has appeared elsewhere (Hoem, 1972).

3 While working with this paper, the author was considerably encouraged by discovering that Gus Haggstrom (1971) has taken essentially the same position in his work with models for American university enrolment. So have Savage and Brown (1960), who write as follows (pp. 37-38):

> "No estimates of standard errors or confidence have been computed for the prediction of college attendance in this section or in Section 9, for although technically possible the computations did not seem warranted. The important point to grasp is that different methods of prediction, although giving generally similar patterns, have substantial variability for each year. This suggests that the choice of a model is probably a source of variability at least as great as the inherent statistical errors in the problems. Since there are no overwhelming reasons for preferring one model to another it would appear that the computation of standard errors of predictions can lead only to a false sense of confidence in the estimates' precision."

Savage and Brown (1960) do give prediction intervals for other forecasts, presumably because they feel more confident about the choice of model in those cases. Brown and Savage (1971) have later reported on their experiences over the period 1960-1970.

4 Paul Meier lucidly explained the concept of levels of error in population prediction in a contribution to the discussion at the Honolulu Symposium. The idea itself is of older date. It is inherent in previous work, e.g., by Pollard (1966,1968), Sykes (1969), Schweder (1971), and in Keyfitz's previous paper (1970, p. 16) on this subject. While Meier suggested using four such levels, it seems more convenient to have six altogether.

5 This idea is prominent in the economic literature, but many authors writing about <u>population</u> prediction seem unaware of it.

6 These ideas have roots back to work done by Ragnar Frisch in the 1940s. Compare the foreword in Tinbergen (1966) and Section 3.2 in Johansen (1969).

7 Keyfitz (1970, p. 21):

> "The distinction between projection and prediction cannot always be maintained. A projection is of interest in the degree in which its assumptions are realistic which is to say in the degree in which it is a prediction."

Compare also, e.g., Siegel (1955), p. 113, and Siegel (1972).

8 Working Group on Social Demography (1970), p. 44. Siegel (1967), Section 28.

There are other methods for regional population forecasting, such as apportionment methods, which do not fit into this scheme. Compare Siegel 1955, p. 116-117; Working Group on Social Demography (1970), p. 44; U.N. Manual III, Section 361.

9 This idea has been taken up by A.H. Pollard (1970), who has given mortality some further study along these lines.

10 Actually, the statement is made in a slightly different connection, but it seems quite appropriate to use it here.

11 In Schweder's terminology (1970), the stochastic process $\{M(t) : t=1,2,\ldots\}$ is a predecessor of $\{X(t) : t=1,2,\ldots\}$.

12 Smith and Wilkinson (1969, 1970). Cfr. also Athreya and Karlin (1971).

13 As explained under point (ii) above, our measure of unreliability stayed on this level throughout the forecasting period.

14 This means using (6.3) in (6.1). Whether one calculates M in a straight forward manner, like we did, or one arrives at it in some other way, is a different matter.

15 Compare Smith and Wilkinson (1970).

16 For example, we can use (6.1) no longer, because if we do that, formula (5.2) will not hold, since (6.2) is false in general when the matrices are correlated.

17 See Subsection 5.B above.

18 This type of idea figures prominently in Eaton's contribution (1972), and it was strongly advocated by Norman Ryder during the discussion at the Honolulu Symposium.

19 The following statement from the preface of a recent Danish population projection (Danmarks Statistik, 1970) with five forecasting series, is probably uncommonly precise:

> "While there may be certain grounds for assuming that series 2, 3, and 4 give an impression of coming population trends which is more probable than series 1 and 5, it is at present very difficult to assess

which one out of series 2, 3, or 4 should be regarded as the most relevant expression of population trends. When, in spite of this, Danmarks Statistik recommends that series 3 is to be used in connection with public planning, etc., the reason is the wish that a dissimilar assessment of the results presented should not give public planners in different institutions cause to base their work on dissimilar expectations about population trends."

20 Formulas for the case of independent M(t) have been given by Sykes (1969) and in a different form by Schweder and Hoem (1972).

21 It is really the entire distribution of X(·) which should be specified and the uncertainty of this specification which should be discussed. For simplicity we consider only EM(t) , however.

22 Cfr. U.N. Population Division (1971), Sections 5, 9, 21, 22, and Economic Commission for Europe (1971), Section 16.

23 Muhsam (1966, p. 277) has given this matter some particular consideration. It has also figured in the economic literature. See, e.g., Johansen (1970).

24 Agnew (1972) represents a particularly bad case. Keyfitz (1972a, pp. 358 and 360) makes a couple of bad splashes.

25 Muhsam (1956,p. 145; 1967, Section 7) admits that demographic theory does not yet offer a basis for the derivation of P. He contends, however, that it is sufficient for his purpose to make vague guesses as to the form of the distribution. One must be permitted some scepticism as to the practicability and usefulness of such an attitude.

26 Törnqvist was worried (1949, p. 75) that he had overestimated this probability, and that it was really less than 0.8.

27 Compare, e.g., footnote 19.

28 The British, who publish a single forecasting series, strongly stress this point (Thompson, 1970).

29 Personal communication from Lars Widén, December 29, 1971. A comparison with footnote 19 and the following paragraph brings out the contrast between the publication policies which the three neighbouring Scandinavian countries have chosen.

30 The idea that the user should settle for a single series, is widespread. For example, it is quite evident in Schmitt (1971) and in Keyfitz (1972) and again in the quotation of footnote 19.

31 The reference to "This Volume" in the list of references alludes to the volume "Statistical Problems in Population Research", in which the present article is supposed to appear as a chapter.

THE CONSIDERATIONS WHICH DETERMINED THE CHOICE OF THE PRIMARY ASSUMPTIONS FOR THE FINNISH POPULATION FORECAST OF 1949[1)]By Leo Törnqvist[2)]

Population forecasts[3)] could be produced in a large number of different ways depending upon which series of vital statistics are chosen as the primary forecasting series. In the choice of primary forecasting series, it is appropriate to select those series whose development has been more regular than the series which could be calculated as soon as the primary forecasting series are given. If one can choose for one's primary series one which has so small variations during the observational period that one can replace the series by a constant without loosing any features which are essential for the understanding of the progress of events, a natural forecasting assumption will be that this time series will continue to be approximately constant during the forecasting period. Such approximately constant time series do not necessarily have to be available in the system of time series which form the initial material of the forecasts. If one can derive new series from these original time series through some calculating rule, and these new series have been approximately constant (invariant) during the period of observation, and furthermore if these new series conversely uniquely determine the series which one wishes to calculate, these derived, approximately invariant time series could be chosen as primary forecasting series. Thus, the first stage in the forecast production consists in seeking out time series which could be regarded as invariant except for minor, random deviations from the mean value. If one finds more approximately constant series than the minimum number which is necessary for the forecast, it is practical to choose such series for primary forecasting series as would minimize the calculations necessary to derive the time series one is really interested in from the primary forecasted series.

---

1) Pp. 69 ff in Jorma Hyppölä; Aarre Tunkelo; and Leo Törnqvist (1949): "Calculations concerning the population of Finland, its reproduction and future development" (in Finnish and Swedish, with French summary and translations of table and diagram headings), Statistiska Meddelanden, No. 38, utgivna av Statistiska Centralbyrån, Helsinki.
2) Translation from the Swedish text by Jan M. Hoem, authorized by Leo Törnqvist.
3) "Forecast" is a translation of the Swedish work "prognos", which may also be interpreted as "projection", depending on the intentions of the author. Törnqvist intended "prognos" to mean "forecast".

In establishing forecasts for the primary series, it is necessary to take into consideration the fact that these series are not completely invariant, but must be regarded as statistical variables which only give a certain amount of information, depending on the length of the observational period relative to the forecasting period, concerning the distribution of the figures which one would get if the series were known even for the forecasting period.

The value taken by the primary series at a given future forecasting time, can be regarded as a sample from this distribution. Which value the sample will represent cannot be foreseen with much precision. On the basis of the variations in the primary series during the observational period, it is possible, however, to get a reasonably good understanding of the approximate position of the percentiles which have particular interest, such as the first, fifth, and ninth decile of the probability distribution which describes the hypothetical mother population from which the observed or "ex post" observable figures could be regarded as samples. While the work which could be put into the task of making precise the notions about the future which one can get through a study of the observations in hand must be seriously limited in practice, it seems permissible to be content with making more or less subjective primary forecasting assumptions concerning the position of these deciles. In this connection it seems to us to be appropriate to make three assumptions concerning the future development. The first one, which we will call the optimistic one, gives the forecasts which are so high that the producer of the forecast estimates that there is a probability of approximately 1/10 that the forecasted figure will be exceeded by the corresponding figure calculated "ex post" for the time period or time in question. The second and most interesting assumption aims at figures with the property that there is a probability of one half that the corresponding figures calculated "ex post" will be larger and one half to be smaller than the forecasted (median) figure. We shall speak of this assumption as the most probable one. The third assumption, which we shall call the pessimistic one here, again aims at the series of figures with the property that there is a probability of approximately 1/10 that the figure calculated "ex post" shall prove to be smaller than the corresponding figure according to the pessimistic assumption. As there are several primary series, these three assumptions may be combined in many different ways to get the series derived from the assumptions. Since a complete analysis of all combinations would require a large amount of work anyway, one must be content with the combinations which seem most interesting. The most interesting forecast is, of course, the one which is based upon the most probable assumptions. When the various optimistic primary assumptions are combined, one gets a derived forecast which as a rule is even more optimistic than the primary optimistic forecasts, and conversely for the derived forecasts based upon the pessimistic assumptions.

(Törnqvist then goes on to describe the technique used to arrive at his three forecasting alternatives. We leave out approximately a page and a half of his text.)

We have finally developed primary forecasting assumptions No. 3, 4, and 5 concerning the development of fertility after relatively free experimentation, keeping in mind the trend in the specific fertility rates described above, corrected with respect to the variations in the marriage rates. As we said already, the assumptions concerning fertility constitute, relatively speaking, the weakest points in the population forecast. The optimistic and the pessimistic forecasting assumptions concerning the development of nativity are rather different from each other. The possibilities of establishing good approximations for the corresponding theoretical forecasting series based on the observational material in hand is considerably less than for the assumptions concerning mortality. It is, therefore, possible that the probability that the number of births during one or more years after 1947 will fall outside of the interval between the pessimistic and the optimistic forecast, must be estimated to exceed the risk of approximately 2/10 which we have tried to make our guidelines in our attempt at choosing the basic assumptions concerning the probable development of nativity.

In the establishment of the forecast of the survival probabilities it has not been possible to take into consideration the increase in these probabilities which can be expected to follow from the great medical progress lately (penicilin, streptomyecin, aeromycin, and so on). It therefore seems more probable that the survival probabilities at least within the next decades will become larger than according to assumption No. 4, than that they would become smaller than according to this assumption. By comparing the development "ex post" with the forecasting assumption No. 4, one can get an evaluation concerning to what extent these and other similar medical developments, which so far have not been able to make any noticeable impact upon our observations, have influenced the development of mortality.

When these basic assumptions concerning mortality and nativity are combined with the assumption that there is no out- or in-migration during the forecasting period which would considerably influence population development, it is possible, solely on a basis of these assumptions and known statistical data for the year 1945, to develop statistical time series describing the development and the composition according to age and sex of the population until the year 2000. For deriving these time series from the given assumptions, we have employed well known computational methods.

Utkommet i serien ART

Issued in the series Artikler fra Statistisk Sentralbyrå (ART)

Nr. 1 Odd Aukrust: Investeringenes effekt på nasjonalproduktet *The Effects of Capital Formation on the National Product* 1957 28 s. Utsolgt

" 2 Arne Amundsen: Vekst og sammenhenger i den norske økonomi 1920 - 1955 *Growth and Interdependence in Norwegian Economy* 1957 40 s. Utsolgt

" 3 Statistisk Sentralbyrås forskningsavdeling: Skattelegging av personlige skattytere i årene 1947 - 1956 *Taxation of Personal Tax Payers* 1957 8 s. Utsolgt

" 4 Odd Aukrust og Juul Bjerke: Realkapital og økonomisk vekst 1900 - 1956 *Real Capital and Economic Growth* 1958 32 s. kr. 3,50

" 5 Paul Barca: Utviklingen av den norske jordbruksstatistikk *Development of the Norwegian Agricultural Statistics* 1958 23 s. kr. 2,00

" 6 Arne Amundsen: Metoder i analysen av forbruksdata *Methods in Family Budget Analyses* 1960 24 s. kr. 5,00

" 7 Arne Amundsen: Konsumelastisiteter og konsumprognoser bygd på nasjonalregnskapet *Consumer Demand Elasticities and Consumer Expenditure Projections Based on National Accounts Data* 1963 44 s. kr. 5,00

" 8 Arne Øien og Hallvard Borgenvik: Utviklingen i personlige inntektsskatter 1952 - 1964 *The Development of Personal Income Taxes* 1964 30 s. kr. 5,00

" 9 Hallvard Borgenvik: Personlige innteksskatter i sju vest-europeiske land *Personal Income Taxes in Seven Countries in Western Europe* 1964 16 s. kr. 5,00

" 10 Gerd Skoe Lettenstrøm og Gisle Skancke: De yrkesaktive i Norge 1875 - 1960 og prognoser for utviklingen fram til 1970 *The Economically Active Population in Norway and Forecasts up to 1970* 1964 56 s. kr. 6,00

" 11 Hallvard Borgenvik: Aktuelle skattetall 1965 *Current Tax Data* 1965 38 s. kr. 6,00 Utsolgt

" 12 Idar Møglestue: Kriminalitet, årskull og økonomisk vekst *Crimes, Generations and Economic Growth* 1956 63 s. kr. 7,00

" 13 Svein Nordbotten: Desisjonstabeller og generering av maskinprogrammer for granskning av statistisk primærmateriale *Decision Tables and Generation of Computer Programs for Editing of Statistical Data* 1965 11 s. kr. 4,00

" 14 Gerd Skoe Lettenstrøm: Ekteskap og barnetall - En analyse av fruktbarhetsutviklingen i Norge *Marriages and Number of Children - An Analysis of Fertility Trend in Norway* 1965 29 s. kr. 6,00

" 15 Odd Aukrust: Tjue års økonomisk politikk i Norge: Suksesser og mistak *Twenty Years of Norwegian Economic Policy: An Appraisal* 1965 38 s. kr. 6,00 Utsolgt

" 16 Svein Nordbotten: Long-Range Planning, Progress- and Cost-Reporting in the Central Bureau of Statistics of Norway *Langtidsprogrammering, framdrifts- og kostnadsrapportering i Statistisk Sentralbyrå* 1966 17 s. kr. 4,00

" 17 Olav Bjerkholt: Økonomiske konsekvenser av nedrustning i Norge *Economic Consequences of Disarmament in Norway* 1966 25 s. kr. 4,00 Utsolgt

Nr. 18 Petter Jakob Bjerve: Teknisk revolusjon i økonomisk analyse og politikk? *Technical Revolution in Economic Analysis and Policy?* 1966 23 s. kr. 4,00

" 19 Harold W. Watts: An Analysis of the Effects of Transitory Income on Expenditure of Norwegian Households 1968 28 s. kr. 5,00

" 20 Thomas Schiøtz: The Use of Computers in the National Accounts of Norway *Bruk av elektronregnemaskiner i nasjonalregnskapsarbeidet i Norge* 1968 28 s. kr. 5,00

" 21 Petter Jakob Bjerve: Trends in Quantitative Economic Planning in Norway *Utviklingstendensar i den kvantitative økonomiske planlegginga i Norge* 1968 29 s. kr. 5,00 Utsolgt

" 22 Kari Karlsen og Helge Skaug: Statistisk Sentralbyrås sentrale registre *Registers in the Central Bureau of Statistics* 1968 24 s. kr. 3,50

" 23 Per Sevaldson: MODIS II A Macro-Economic Model for Short-Term Analysis and Planning *MODIS II En makroøkonomisk modell for korttidsanalyse og planlegging* 1968 40 s. kr. 4,50 Utsolgt

" 24 Olav Bjerkholt: A Precise Description of the System of Equations of the Economic Model MODIS III *Likningssystemet i den økonomiske modell MODIS III* 1968 30 s. kr. 4,50

" 25 Eivind Hoffmann: Prinsipielt om måling av samfunnets utdanningskapital og et forsøk på å måle utdanningskapitalen i Norge i 1960 *On the Measurement of the Stock of Educational Capital and an Attempt to Measure Norway's Stock of Educational Capital in 1960* 1968 60 s. kr. 5,00

" 26 Hallvard Borgenvik: Aktuelle skattetall 1968 *Current Tax Data* 1969 40 s. kr. 7,00

" 27 Hallvard Borgenvik: Inntekts- og formueskattlegging av norske kapitalplasseringer i utlandet *Income and Net Wealth Taxes of Norwegian Investment in Foreign Countries* 1969 40 s. kr. 7,00

" 28 Petter Jakob Bjerve og Svein Nordbotten: Automasjon i statistikkproduksjonen *Automation of the Production of Statistics* 1969 30 s. kr. 7,00

" 29 Tormod Andreassen: En analyse av industriens investeringsplaner *An Analysis of the Industries Investment Plans* 1969 26 s. kr. 5,00

" 30 Bela Balassa og Odd Aukrust: To artikler om norsk industri *Two Articles on Norwegian Manufacturing Industries* 1969 40 s. kr. 5,00

" 31 Hallvard Borgenvik og Hallvard Flø: Virkninger av skattereformen av 1969 *Effects of the Taxation Reform of 1969* 1969 35 s. kr. 7,00 Utsolgt

" 32 Per Sevaldson: The Stability of Input-Output Coefficients *Stabilitet i kryssløpskoeffisienter* 1969 40 s. kr. 7,00

" 33 Odd Aukrust og Hallvard Borgenvik: Inntektsfordelingsvirkninger av skattereformen av 1969 *Income Distribution Effects of the Taxation Reform of 1969* 1969 29 s. kr. 7,00

" 34 Odd Aukrust og Svein Nordbotten: Dataregistrering, dataarkiver og samfunnsforskning *Data Registration, Data Banks and Social Research* 1970 43 s. kr. 7,00

Nr. 35 Odd Aukrust: PRIM I A Model of the Price and Income Distribution Mechanism of an Open Economy *PRIM I En modell av pris- og inntektsfordelingsmekanismen i en åpen økonomi* 1970 61 s. kr. 7,00

" 36 Arne Amundsen: Konsumets og sparingens langsiktige utvikling *Consumption and Saving in the Process of Long-Term Growth* 1970 18 s. kr. 5,00

" 37 Steinar Tamsfoss: Om bruk av stikkprøver ved kontoret for intervjuundersøkelser, Statistisk Sentralbyrå *On the Use of Sampling Surveys by the Central Bureau of Statistics, Norway* 1970 46 s. kr. 7,00

" 38 Svein Nordbotten: Personmodeller, personregnskapssystemer og persondataarkiver *Population Models, Population Accounting Systems and Individual Data Banks* 1970 28 s. kr. 7,00

" 39 Julie E. Backer: Variasjoner i utviklingen hos nyfødte barn *Variations in the Maturity Level of New Born Infants* 1970 36 s. kr. 7,00

" 40 Svein Nordbotten: Two Articles on Statisticial Data Files and Their Utilization in Socio-Demographic Model Building *To artikler om statistiske dataarkiver og deres bruk i sosio-demografisk modellbygging* 1971 30 s. kr. 7,00

" 41 Per Sevaldson: Data Sources and User Operations of MODIS, a Macro-Economic Model for Short Term Planning *Datagrunnlag og brukermedvirkning ved MODIS, en makroøkonomisk modell for planlegging på kort sikt* 1971 31 s. kr. 7,00

" 42 Erik Biørn: Fordelingsvirkninger av indirekte skatter og subsidier *Distributive Effects of Indirect Taxes and Subsidies* 1971 42 s. kr. 5,00

" 43 Hallvard Borgenvik og Inger Gabrielsen: Aktuelle skattetall 1970 *Current Tax Data* 1971 53 s. kr. 7,00

" 44 Vidar Ringstad: PRIM II En revidert versjon av pris- og inntektsmodellen *PRIM II A Revised Version of the Price and Income Model* 1972 43 s. kr. 7,00

" 45 Jan M. Hoem: Purged and Partial Markov Chains *Lutrede og partielle Markovkjeder* 1972 16 s. kr. 5,00

" 46 Jan M. Hoem: Two Articles on the Interpretation of Vital Rates *To artikler om tolking av befolkningsrater* 1972 33 s. kr. 7,00

" 47 Inger Gabrielsen: Aktuelle skattetall 1972 *Current Tax Data* 1972 58 s. kr. 8,00

" 48 Vidar Ringstad: Om estimering av økonomiske relasjoner fra tverrsnitts-, tidsrekke- og kombinert tverrsnitts tidsrekke - data *On the Estimation of Economic Relations Using Cross Section-, Time Series - and Combined Cross Section - Time Series - Data* 1972 26 s. kr. 7,00

" 49 Jan M. Hoem: On the Statistical Theory of Analytic Graduation *Statistisk teori for analytisk glatting* 1972 41 s. kr. 7,00

" 50 Henry M. Peskin: National Accounting and the Environment *Nasjonalregnskap og miljøverdier* 1972 60 s. kr. 8,00

" 51 Eivind Gilje: Analytic Graduation of Age-Specific Fertility Rates *Analytisk glatting av aldersspesifikke fødselsrater* 1972 49 s. kr. 8,00

Nr. 52 Jan M. Hoem og Arne Rideng: Kommentarer til Statistisk Sentralbyrås framskriving av folkemengden i kommunene 1972-2000 *Comments to the Regional Population Projections for Norway* 1972 29 s. kr. 7,00

" 53 Juul Bjerke: Estimering av konsumfunksjoner på grunnlag av nasjonalregnskapsdata 1865-1968 *Estimating Consumption Functions from National Accounts Data* 1972 60 s. kr. 8,00

" 54 Jan M. Hoem: Usikkerhet ved befolkningsprognoser *Inaccuracy of Population Projections* 1973 63 s. kr. 8,00

" 55 Erik Biørn: Prognoser for de langsiktige endringer i sammensetningen av det private konsum *Long Term Forecasts for the Changes in the Composition of the Private Consumption* 1973 71 s. kr. 8,00

" 56 Jan M. Hoem: Inhomogeneous Semi-Markov Processes, Select Actuarial Tables, and Duration-Dependence in Demography *Inhomogene semimarkovprosesser, selekte aktuartabeller og varighetsavhengighet i demografi* 1973 54 s. kr. 8,00

" 57 Svein Brenna: Revisjon av indeksene for utenrikshandelen *Revision of Indices for Foreign Trade* 1973 47 s. kr. 7,00

" 58 Jan M. Hoem: Statistisk Sentralbyrås utvalgsundersøkelser: Elementer av det matematiske grunnlaget *The Sample Surveys of the Central Bureau of Statistics of Norway: Basic Mathematical Elements* 1973 59 s. kr. 8,00

" 59 Inger Gabrielsen: Aktuelle skattetall 1973 *Current Tax Data* 1973 63 s. kr. 8,00

" 60 Per Sevaldson: Om oppstilling og bruk av regionalt nasjonalregnskap *Construction and Use of Regional National Accounts* 1973 74 s. kr. 7,00

Publikasjonen utgis i kommisjon hos
H. Aschehoug & Co., Oslo, og er til salgs hos alle bokhandlere
Pris kr. 8,00

Omslag trykt hos Grøndahl & Søn, Oslo