Arbeidsnotater

STATISTISK SENTRALBYRÅ

Dronningensgt. 16, Oslo-Dep., Oslo 1. Tlf. 41 38 20, 41 36 60

IO 72/3

Oslo, 11 February 1972

LEVELS OF ERROR IN POPULATION FORECASTS.

Comments upon Nathan Keyfitz's paper[x)]

"On future population"

by Jan M. Hoem

CONTENTS

x) Presented to the IASPS Satellite Symposium on Statistical Problems in Population Research at the East-West Population Institute, Honolulu, Hawaii, August 2-6, 1971.

1. INTRODUCTION

Keyfitz's paper to the Honolulu Symposium (1971) gave rise to a lively debate in which most of the people present took part. There was considerable agreement on many of the points which he raised, while we got a definite division of opinion on other issues. The present author was among those who endorsed many of Keyfitz's standpoints, but who disagreed with one of his main conclusions. Much of what follows can be seen as a commentary to his paper and the ensuing discussion, written after the event[1]. It may be as well, then, to start by stating my stand on the three main issues of Keyfitz's paper, summarised in his concluding section.

(i) We can easily dispose of the first conclusion, because I agree completely with Keyfitz's point of view. The choice of a forecasting model is seriously problematic because current prediction methods simply are not good enough.

(ii) Nor do I have any quarrel with Keyfitz's demand that the forecasting assumptions should be fully and clearly stated. I would contend that some forecasters do a good job in this respect, however. I also feel that Keyfitz goes quite a bit too far if he wants the general reader to be able to reproduce all forecasting calculations on his own. For one thing, I cannot see what interest the general public would take in this possibility. Secondly, modern population forecasts utilize such a large number of individual vital rates that the cost of presentation and publication will often be prohibitive. Information about details of this kind represents a service which we in Norway are willing to extend to individual research and planning institutions on an ad hoc basis (and usually only in the form of computer print-outs), but not to the general public.

(iii) His contention that official producers of statistics should provide the probability distribution of the population at each future date, calls for a more involved comment. I agree that this is, in principle, a goal towards which we should strive, and to which we may possibly find a reasonably accurate and operational solution some time in the future. I seriously question the timeliness of forwarding this request just now, however, and I do this precisely because I agree so fully with the first of his issues.

During the discussion in Honolulu, Keyfitz formulated his stand through the following question: When forecasters can publish mean future population figures, then why can they not also present corresponding standard errors? After all, future population size is a random variable.

In my opinion, this represents an incorrect interpretation of the forecaster's situation. Firstly, forecasters like the U.S. Census Bureau and the Central Bureaux of Statistics of a lot of countries do not produce one series of future population figures, but several alternative series. None of these figures can be regarded as anything so precise as the expected value of the corresponding random variable. We do not know enough about population processes to attempt anything that hazardous.

Secondly, we know even less about the numerical size of expected future unreliability of forecasts. We would, therefore, deceive ourselves if we tried to calculate (and publish) such standard errors, because this would give an impression of better knowledge than what we actually have. I do not understand what the purpose would be of bringing in yet another type of uncertainty into the forecasts, such as this one would be. As I see it, we cannot get anywhere in this direction before we know more about the underlying structure which generates future populations.

A large part of the present paper will go into sketching my main reasons for taking this stand.

2. PRELIMINARIES

2.A. The chief purpose of making a population forecast, as I see it, is to contribute to improved planning and to better decisions[2)]. 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 accuracy <u>is</u> important. Grossly unreliable forecasts are of little value for planning. The negative feed-back effect is one reason why forecasts do not get home often. There are many other such reasons. It is a purpose of this paper to review them.

We shall group the reasons why population forecasts are inaccurate into sources on various levels, thus pursuing an idea introduced by Paul Meier during the discussions at the Honolulu Symposium[3].

2.B. As Keyfitz explains (1971, p. 31-32), the advance calculation of future population figures can appear in many modes. In the present paper, I shall address myself solely to the situation where the figures are positively aimed at saying something about the future of a real population. Thus, I will leave aside such things as counter-predictions and calculations made for analytical purposes only, and I will concentrate on predictions/forecasts, as already indicated by the language used above.

With the planning purpose of the forecasts in mind, let us give some further consideration to forecasting modes. 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 cash in 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 2.A above). This idea has been extensively used in indicative economic planning in France, but not, to my 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 tied together by the model[4]:

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.C. Most producers of future population figures of the kind which we take into account in this paper (including the Central Bureau of Statistics of Norway) will insist upon calling their commodities population <u>projections</u>. I quite agree with Keyfitz when he argues[5] that most users will take interest in the figures produced mainly insofar as they can be regarded as <u>predictions</u>, however, and the majority will treat them as that anyway.

It seems that the majority 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 the general picture is that 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 Chapters.

2.D. Before going into the presentation proper of the paper, we conclude this preliminary Chapter by introducing some conventions and a notation which differs slightly from Keyfitz's. 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. I want to reserve the letter P for a probability distribution, so let us use the notation $\mathbf{X}(t)$ for the vector of the actual population at time t. The forecaster is required to produce a prediction $\tilde{\mathbf{X}}(t)$ for $\mathbf{X}(t)$ for $t = 1,2,\ldots,T$. Present forecasting models will typically calculate the forecast recursively, so that $\tilde{\mathbf{X}}(1)$ is calculated first, the result is used to calculate $\tilde{\mathbf{X}}(2)$, and so on. Forecasts of a national population will typically be produced through a linear model. It we disregard international migration, this means that one will use a forecasting relation of the form

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

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) with a time-dependent projection matrix $\underset{\sim}{M}(t)$. So can regional projection models based on (area-) component methods[6, 7], which are more sophisticated versions of the Markov chain model which Keyfitz mentions (1971, p. 25-26). While Keyfitz in effect limits himself to national forecasts, my discussion is intended to cover prediction of the regional distribution of the population as well. I feel that the restriction to the national level covers up some of the complexities involved in the nitty-gritty work in this field.

For simplicity and concreteness, most of what I shall have to say, will be tied formally to linear forecasting models for closed populations. It is easy to see how a lot of it can be transferred to other situations as well.

3. SOURCES OF FORECASTING INACCURACY

3.A. This Chapter 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 Chapters.

Type I. Estimation and registration errors

Level 1. One does not really know any of the parameters of the forecasting model. 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, and

(c) errors in size and composition of the initial population $\underset{\sim}{X}(0)$.

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

(d) rounding errors,

and we will include these as well on Level 1.

Type II. Errors due to random fluctuations

Level 2: Pure randomness. Even if we knew the survival probabilities in force during a given forecasting year, the proportion of survivors in each population group would not be exactly equal to this probability. There will be some random variation. Correspondingly for births, etc.

Level 3: Random vital rates. Pollard (1968) and Sykes (1969) have criticized models assuming that the elements of the projection matrix $\underset{\sim}{M}(t)$ 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[8]. Sykes (1969, p. 118) asserts[9] 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 $E\underset{\sim}{M}(t)$ of the projection matrix $\underset{\sim}{M}(t)$, while the fluctuations would be measured, among other things, by the covariance matrix $\underset{\sim}{\Sigma}(t)$ of the elements of $\underset{\sim}{M}(t)$.

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 one uses, 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. I have therefore avoided calling any Type or Level "model errors".

All of these kinds of errors can occur no matter which forecasting mode one applies. Thus our 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.

4. ESTIMATION AND REGISTRATION ERRORS

4.A. The statistical estimation of model parameters will give rise to estimation error. As far as I know, 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 variation 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 literature 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 developing countries, but trouble tormenting United States forecasters figured prominently during the discussions in Honolulu.

Even if we certainly are not relieved of such problems in Scandinavia either, the Norwegian population registration system, for one, is reasonably good, and the data errors which we do have are surely of a much smaller size order than those of many other countries.

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 $\mathbf{X}(0)$ is regarded as given, and so is the projection matrix $\mathbf{M}$, which is taken to be independent of time. $\mathbf{X}(t)$ becomes a random vector with some covariance matrix $\mathbf{C}(t)$. As a forecast for $\mathbf{X}(t)$, one uses

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

which gives

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

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 $X(0)$ till make $X(t)$ approximately normally distributed, and that

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

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[10]. (This is essentially the same idea as that due to Smith and Wilkinson[11], 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 $EM(\cdot)$ as known, the forecasting calculations can be carried out recursively, viz. through

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

and we still get (5.2) because

$$(6.2) \qquad E \prod_{s=1}^{t} \underset{\sim}{M}(s) = \prod_{s=1}^{t} E\underset{\sim}{M}(s).$$

6.B. Even though Sykes also gives some formulas for the general case, where $E\underset{\sim}{M}(t)$ and $\underset{\sim}{\Sigma}(t)$ 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.,

$$(6.3) \qquad E\underset{\sim}{M}(t) \equiv \underset{\sim}{M}$$

for a suitable matrix $\underset{\sim}{M}$, and $\underset{\sim}{\Sigma}(t)$ 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 $\underset{\sim}{M}$ and $\underset{\sim}{\Sigma}$ were based on Norwegian data from the years 1953-1968. Our main conclusions were as follows:

(i) Mortality fluctuations cannot 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 one regards the variability in this model 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[12]. 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. I tend to feel that the latter of the alternatives under point (iii) above is the more plausible. There is no reason to believe that the fertility level in Norway has really been constant over the years from 1953 to 1968, and that the variation which did occur, was due only to random variation of the type on which Pollard and Sykes seem to base their argument for random projection matrices. Surely, there has been a genuine development in $E\underset{\sim}{M}(t)$ 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 $E\underset{\sim}{M}(\cdot)$ around this mean to the fluctuations of $\underset{\sim}{M}(\cdot)$ around $E\underset{\sim}{M}(\cdot)$. 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 $\underset{\sim}{M}(\cdot)$.

It is not really adequate to use a constant projection matrix throughout the forecasting period either. [This means using (6.3) in (6.1). Whether one calculates $\underset{\sim}{M}$ from the observations in a straightforward manner, like we did, or one arrives at it in some other way, is a different matter.] When the forecast is made, one may have information supporting a certain time-trend in $E\underset{\sim}{M}(t)$. (Compare our comment on cohort methods at the end of Section 1.E.) 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 $E\underset{\sim}{M}(t)$ in a non-arbitrary way, and this is one of the more important sources of inaccuracy of birth forecasts. At this point, however, we have encounted problems which really belong on higher levels of error than the present Level 3, so we shall postpone discussing them to later Chapters.

6.D. The assumption 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[13], 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 $\hat{\underset{\sim}{X}}(t)$ recursively and still obtain (5.2) if the matrices are correlated. [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.]

If one prescribes some probabilistic model for the $\mathbf{M}(t)$-s, it is, of course, theoretically possible to make the forecast through simulation. The practical problems involved seem considerable, however. Since our knowledge of realistic stochastic models for these matrices is very vague, extensive simulation experiments will probably be necessary both to establish a model and later to make the forecast, and I believe that the corresponding resources could be put to better use in other work with forecasts.

6.E. Let me close this Chapter by mentioning that it has not been possible for us to extend the elegant results mentioned at the end of Chapter 5 to the case where the projection matrices are random. The problem is that it is too difficult to derive the probability distribution of $\pi_{s=1}^{t} \mathbf{M}(s)$.

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 good enough. Usually, it will be quite impossible to account for such matters when a forecast is made, however. Since we all agree 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 $E\mathbf{M}(t)$, and possibly in other characteristics also, such as in the covariance matrix $\mathbf{\Sigma}(t)$ of $\mathbf{M}(t)$.

7.B. (Level 4.) On the other hand, one expects the forecaster to make allowance in his calculations for a gradual change in $E\mathbf{M}(t)$ (and possibly also other characteristics), and he is expected to do so better than others. Even though it is a gradual development that he is supposed to foresee, he is faced with a spectre 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. One tries to make them represent reasonable and realistic future trends, and at the same time provide some impression of the range within which the trends reasonably can be expected to lie.

(ii) The other line of attack consists in further developing the ideas described in Chapter 6[14]: One may specify some mean trend in $E\underset{\sim}{M}(t)$ and take into consideration the uncertainty of forecasts of births and similar factors through a suitable specification of the probabilistic mechanism generating the random matrices $\underset{\sim}{M}(t)$. 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 $\underset{\sim}{\Sigma}(t)$ increase with t, since these elements represent the variances of the elements of $\underset{\sim}{M}(t)$.

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 procedure (i) above will typically be presented with relatively vague statements concerning the prospective accuracy of the alternative series[15]. Considering common experience with forecasting accuracy, this is understandable. On the other hand, one may perhaps criticize forecasters because they have not taken greater interest in producing measures of the reliability of their figures.

In principle, the second procedure above enables one to produce statements on a quite different level of preciseness. As a minimum, one can calculate standard errors for all figures forecast, for example for the predicted number of births in each forecasting year[16]. If the probabilistic model for the $\underset{\sim}{M}(t)$-s is sufficiently specified and if one can solve the mathematical problems, one may also find the probability distribution of $\underset{\sim}{X}(\cdot)$. In this case, this probability distribution will be the real forecast. On its basis, one may for instance calculate prediction regions, and one may put statistical decision theory to good use[17]. In theory, therefore, procedure (ii) above gives results which one should decidedly aim at. My account below will show why I am rather sceptical as to our possibilities of carrying out such a program today, however.

7.C. Any forecast will be based on assumptions 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 $\boldsymbol{\theta}$. A specification of trends in population components during the forecasting period then actually consists in specifying a parametric matrix function $\mathbf{m}(t;\boldsymbol{\theta})$ and letting[18]

$$E\mathbf{M}(t) = \mathbf{m}(t;\boldsymbol{\theta}).$$

Level 4 errors result from a somewhat erroneous value of $\boldsymbol{\theta}$, or from a bit of deviation of $\mathbf{m}$ from what it should really have been. Level 5 unreliability is due to gross shifts in $\boldsymbol{\theta}$ or to substantial changes in $\mathbf{m}$. 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 one substitutes a more sophisticated purely demographic model for a simpler one. On the other hand, an extention of the forecasting model may entail something much more radical, such as the introduction of non-demographic forecasting variables and of policy variables, as sketched in Section 2.B above.

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[19].

Similarly, present-day knowledge does not permit us to take the publication effect (mentioned in Sections 2.A and 2.B) explicitly into account in the forecasting model[20]. (Of course, the publication effect is in part a consequence of the way in which the forecast is presented, which we shall discuss in Chapter 9 below.)

7.E. A main viewpoint of this paper 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 nothing we 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 working at it full steam. 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 Chapter, we shall contend 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 (1971, p. 29-30) 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

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

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. The operative forecast corresponds to Keyfitz's fifth mode of future population.) 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.

Evidently, Muhsam has pointed out a useful way of applying forecasts. In many connections one 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.

8.B. This type of example represents an over-simplified situation, however, where many problems do not surface. When such an example is presented without an extensive discussion of its limitations, one gets an overly optimistic 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 $\mathbf{X}(t)$ for all t, and on the operative forecast $\tilde{\mathbf{X}}(t)$ for all t. The probability distribution which should be specified, is the simultaneous distribution P for $\mathbf{X}(1), \ldots, \mathbf{X}(T)$. The risk to be minimized is

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

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, and

(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 which one is up against when one wants to establish the probability distribution P, even if Type III errors are disregarded. In his paper (1971), Keyfitz strongly recommends using (a simpler version of) P, but he seems to take the difficulties involved too lightly. He explains (and this was made even clearer during the discussion) that his prediction intervals and his normality assumptions should be regarded as statements of personal probability. It is still important that the reader be informed of how he arrived at his conclusions. I quite agree with Eaton (1971) when he puts Keyfitz on the spot here.

There is a lot of leeway for subjective considerations of future population trends in the theory which we sketched in previous Chapters. For example, there is much scope for such elements in the specification of the probabilistic mechanism generating the random matrices $\underset{\sim}{M}(t)$, which we touched upon in Section 7.B. In principle, the distribution of the population vectors $\underset{\sim}{X}(1)$, $\underset{\sim}{X}(2)$, ... can be derived from this specification. The way I see it, 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 $\underset{\sim}{X}$ has been generated.

8.D. Both the specification of such mechanisms and the mathematical derivation of the probability distribution of the population vectors is 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 to me exactly what Törnqvist really intended this 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[21)]. (The probability of 0.8 may also have been meant to cover other events.)

In 1965, Törnqvist reported on the outcome so far (Törnqvist, 1967). It turned out that the forecast from 1949 had been highly 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. I thus maintain that there are serious problems connected with establishing the probability distribution P. Similar problems arise when one wants to find the loss function L. Public population forecasts are wanted by a highly heterogeneous 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, we 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 Muhsam'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 which we would be up against if we tried this, is that we 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. As far as I know, attempts in this direction have not been singularly successful. Nevertheless, the concept of a welfare function has been highly fruitful in economics, and it is possible that the loss function could play a similar role in forecasting theory. While its direct applicability in practical forecasting work should not be overrated, I thus believe that it could be quite useful in forecasting theory.

9. THE PRESENTATION OF POPULATION FORECASTS

9.A. When population forecasts are so unreliable, the presentation to the public is highly important. One will expect the forecaster to have a better understanding than others of future population trends, and the users will rightly expect 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, first-line 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 carried out in Norway.

People who are able to understand something 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 $E\underset{\sim}{X}(\cdot)$ 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 it turns out that we are able to carry out a program of this type, it would probably never be accessible to the great majority of users. For one thing, they have no need for such a complex offer; for another, they would really rather need 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 continuously remind the user 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 which is 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[22]. I have some difficulty in understanding this fear, and wonder whether it may not be 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[23], and then other deliberations than the fear of dissimilar expectations about the future must be decisive for the choice between publishing several forecasting series or 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 behind 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 I would endorse.

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 curiousity 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[24].

The U.S. Bureau of the Census has chosen a different way out. 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. Pursuing this line of thought, I must take issue with demands that the forecaster should enable the user to decide which single set of alternative assumptions he finds most relevant[25], or even that he should "express, no matter how tentatively, his own judgment as to which series to choose" (Schmitt, 1971, p. 8). 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 is an expression of our opinion of how population trends seem to become. 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.

10. CONCLUSIONS

In conclusion, let me summarize the stand taken in this paper as follows:

(i) The unreliability of population forecasts has many sources. Apart from data problems, the most important of these are errors of model specification and gross shifts in vital rates. Gross shifts are due to events which can hardly be foreseen, so there is not much one can do about it. While specification errors are endemic in this type of work, greater resources should be devoted to alleviating the present acute situation and to making population forecasts into a better planning instrument.

(ii) Calculating and publishing standard errors for predicted variables does not make much sense until better forecasting models are available. The idea of establishing the probability distribution of the prediction meets with the same counter-argument, and, in addition, seems to be up against unsolved mathematical problems.

(iii) The present policy of publishing several alternative forecasting series is basically sound and should be maintained. Users should be continuously reminded of forecasting unreliability, and the publication of alternative series is one way of achieving this. Users should not be encouraged to select a single series on which they may base all their planning.

ACKNOWLEDGEMENTS

I wish to thank Svein Longva, Thorvald Moe, Tore Schweder, and Bjørn Tønnesen for reading the present paper in manuscript and for giving many comments which were of great use. I am also grateful to Lars Widén for giving the explanation on which the first part of Section 9.D is based.

While working with this paper, I was considerably encouraged by discovering that Gus Haggstrom (1971) has taken essentially the same position in his work with models for American university enrolment as I have done here.

REFERENCES

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

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

[3] Cox, P.R. (1966): "Projections of population size and age-sex structure. Statement by the Rapporteur." Pp. 253-258 in Volume 1 of [39].

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

[5] Eaton, Morris L. (1971): "Error and variance in the population prediction process." University of Copenhagen, Institute of Mathematical Statistics, Preprint No. 9.

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

[7] 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.

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

[9] Hoem, Jan M. (1972): "Usikkerhet ved befolkningsprognoser." Central Bureau of Statistics of Norway, Working Paper, in press.

[10] Hyppölä, J.; A. Tunkelo; and L. Törnqvist (1949): "Beräkningar rörande Finlands befolkning, dess reproduktion och framtida utveckling." (In Finnish and Swedish, with a Summary in French. A translation into English of central parts of Törnqvist's contribution can be obtained from the present author.) Statistiska Centralbyrån, Helsingfors. Statistiska Meddelanden, Nr. 38.

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

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

[13] 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.

[14] Keyfitz, Nathan (1971): "On future population." Paper presented to the Honolulu Symposium.

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

[16] 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 [39].

[17] Muhsam, H.V. (1967): "The use of cost functions in making assumptions for population forecasts." Pp. 23-26 in Volume 3 of [39].

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

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

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

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

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

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

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

[25] 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, in press.

[26] 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.

[27] 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 [39].

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

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

[30] 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.

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

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

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

[34] Törnqvist, Leo (1967): "The post-War population development of Finland compared with predictions made after the War." Pp. 44-46 in Volume 3 of [39].

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

[36] 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.

[37] 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.

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

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

FOOTNOTES

1) A more extensive review of these matters, written in Norwegian for domestic consumption, will appear elsewhere (Hoem, 1972). A more extensive list of references will then be given.

2) Compare, e.g., the discussions by Cox (1966), pp. 256-257, and by Siegel (1972).

3) 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, I have found it convenient to have six altogether.

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

5) Keyfitz (1971), pp. 20-21, and particularly in (1970), where, among other things, he says (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).

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

7) 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.

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

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

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

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

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

13) Compare Smith and Wilkinson (1970).

14) This type of idea figures prominently in Eaton's contribution (1971), and it was strongly advocated by Norman Ryder during the discussion in Honolulu.

15) 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."

16) Formulas for the case of independent $\underset{\sim}{M}(t)$ have been given by Sykes (1969) and in a different form by Schweder and Hoem (1972).

17) Compare Chapter 8 below.

18) It is really the entire distribution of $\underset{\sim}{X}(\cdot)$ which should be specified and the uncertainty of this specification which should be discussed. For simplicity we consider only $E\underset{\sim}{M}(t)$, however.

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

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

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

22) Compare, e.g. footnote 15.

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

24) Personal communication from Lars Widén, December 29, 1971. A comparison with Section 10.C and footnote 15 brings out the contrast between the publication policies which the three neighbouring Scandinavian countries have chosen.

25) Compare, e.g., Keyfitz (1971), pages 21, 22, 31.