Statistics Netherlands
Department of population

PO Box 4000
2270 JM Voorburg
The Netherlands

June 2000

Dealing with uncertainty in population forecasting

Joop de Beer

Abstract

The uncertainty of forecasts of the size and age structure of the population at the national level depends on the uncertainty of forecasts of fertility, mortality, and international migration. The uncertainty of forecasts of these three demographic components depends on the question in what way their future development may be different from the past or different in another way than expected. The answer depends on the way we make our forecasts. Appropriate demographic parameters to be forecasted are selected on the basis of demographic analyses aimed to identify parameters that are stable through time or change in a systematic way or have a stable relationship with selected explanatory variables. Forecasts are uncertain to the extent that it is uncertain whether these parameters will remain stable in the future. The selected parameters are forecasted on the basis of extrapolation, explanation and/or expectation. If forecasts are based on extrapolations, we can say that the future is uncertain because the direction of trends may change. If forecasts are based on explanatory models, the future is uncertain because the form of the relationships may change or the exogenous variables may develop in a different direction than expected, or variables not included in the model may have an unexpected impact in the future. Since there is no unambiguous objective way of assessing the degree of uncertainty due to these sources the assessment of the uncertainty of forecasts of fertility, mortality, and migration depends heavily on judgment. The degree of uncertainty of forecasts of population size and structure can be specified by determining forecast variants of fertility, mortality, and migration, e.g. low and high variants or scenarios. On the basis of these variants deterministic variants of population size and structure can be calculated. One problem in using deterministic variants is that they do not provide an indication of the probability that the interval between the variants or scenarios will cover the true value. If assumptions about the probability of forecast intervals of fertility, mortality, and migration are specified, the degree of uncertainty of forecasts of the size and age structure of the population can be assessed on the basis of a stochastic approach.

1. Introduction

In most industrialised countries population forecasts are published by the national statistical institutes. Because of the uncertainty of forecasts, many statistical agencies call their forecasts projections (Keyfitz, 1982). Population projections are calculations of the future development of the population size and structure based on certain assumptions about the future course of population change (Keilman, 1990). Thus projections are *conditional*, showing the consequences of the assumptions that are made (e.g. Ahlburg and Land, 1992). The assumptions can be arbitrary (e.g. the assumption that fertility will reach replacement level) or normative (e.g. the assumption that net migration will be zero); the assumptions can be predictions deduced from a verified theory under the *ceteris paribus* clause or scenarios describing - internally consistent - alternative futures. Or the assumptions can be considered to describe the most probable future development. In the latter case the projection is called a forecast. Forecasts are *unconditional*: based on current scientific insights, the forecaster gives his best guess of the future development of population (Keilman, 1990). If statistical agencies publish projections rather than forecasts, it is left to the user to judge the plausibility of the underlying assumptions. Keilman argues that national statistical agencies should publish population forecasts rather than projections since their professionals making the projections should be better equipped to assess the plausibility of assumptions than the average user. In fact, users act as if the statistical agencies produce forecasts, assuming that the agencies base their calculations on plausible assumptions (Keyfitz, 1982).

The main reason for labelling a population forecast a projection is uncertainty. Uncertainty implies that forecast errors may occur. Forecast errors should, however, not be mistaken for mistakes. If no relevant, available information is ignored and if state-of-the-art methodology is used in a proper way, the forecaster is not to be blamed for making forecast errors. Forecast errors are unavoidable. Therefore it does not make sense to demand that a forecast will come true. Zero errors are a matter of good luck rather than proof of the quality of the forecast. This does not imply that the size of forecast errors is irrelevant. One would expect adequate forecasts to come up with smaller errors than inferior forecasts. But at the moment a forecast is made, it will always be uncertain whether the errors of that forecast will be smaller than the errors of any other forecast, no matter how carefully the forecast is prepared. Acknowledging that forecast errors are inevitable, uncertainty is no reason not to publish a forecast nor to label a forecast ‘projection’. In spite of uncertainty - or rather because of uncertainty - users need forecasts. “Projections intended for planning purposes of any kind should be labelled ‘forecast’” (Pittenger, 1977). Instead of labelling a forecast a projection, the forecaster should give an indication of the degree of uncertainty of his forecast. Clearly most users want one population forecast describing the most likely future. It is the task of demographers to attempt to produce forecasts that are as accurate as possible. It is, however, also the task of demographers to provide information about the uncertainty of their forecasts. “Demographers have a responsibility for communicating their knowledge about uncertainty” (Lutz, Goldstein and Prinz, 1996).

This paper discusses the uncertainty of population forecasts at the national level. As the uncertainty of population forecasts depends on the uncertainty of the components of population change, viz. fertility, mortality, and migration, the main part of this paper focuses on the uncertainty of forecasts of these three demographic components. Section 2 argues that the uncertainty of forecasts of fertility, mortality, and migration depends on the way these forecasts are made. First, demographic parameters to be forecasted are selected on the basis of demographic analyses aimed to identify parameters that are either stable through time or change in a stable way, or have stable relationship with explanatory variables (this will be discussed in section 3). Subsequently the selected parameters are forecasted on the basis of extrapolation, explanation, or expectation. Sections 4-6 discuss the sources of uncertainty of forecasts of fertility, mortality, and migration based on these approaches. The overview of sources of forecast errors leads to the conclusion that there is no unambiguous objective way of assessing the degree of uncertainty of forecasts of fertility, mortality, and migration. The assessment of uncertainty of population forecasts is based on assumptions of which the validity is uncertain. This implies that the assessment of uncertainty depends heavily on judgment. The degree of uncertainty of forecasts of fertility, mortality, and migration can be specified by determining forecast variants. Section 7 distinguishes two types of deterministic variants: scenarios on the one hand and low and high variants on the other. One problem in using deterministic variants is that they do not provide an indication of the probability of the forecast interval. Hence it is not clear for the user how seriously he should take the uncertainty into account. Section 8 discusses methods for assessing the probability of forecast intervals for fertility, mortality, and migration. Obviously the ‘true’ probability that a forecast will come true is not known. But that does not imply that the forecaster should not aim at giving an indication of the probability of his forecast. On the basis of assumptions about the width of the forecast intervals of fertility, mortality, and migration, the uncertainty of forecasts of the size and age structure of the population can be assessed. Section 9 discusses both deterministic and statistical methods that can be used for this purpose. Finally section 10 gives the main conclusions.

2. Sources of uncertainty

In order to assess the uncertainty of population forecasts it is necessary to identify the possible sources of forecast errors. Population forecasts are usually based on the cohort-component model. This implies that forecasts of population size and age structure are based on an assessment of the jump-off population and on assumptions about the future course of fertility, mortality, and migration. The uncertainty of population forecasts depends on the uncertainty of the validity of the assumptions underlying the use of the cohort-component model. Keilman (1990) distinguishes three types of uncertainty:

1. Randomness of population processes. The time when individuals experience a demographic event is random. If the probability distribution is known, the degree of uncertainty of the forecast depends on the variance.
2. Ignorance of the processes: the probability distribution is not known. More research may lead to a more accurate assessment of the uncertainty of the forecast.
3. Inherent unpredictability. Individuals have a choice of possible actions. This makes their behaviour essentially unpredictable.

In the practice of population forecasting the latter two types of uncertainty are relevant: either we do not know the probability distribution or we cannot know the probability distribution. The question which of both interpretations is correct may be interesting from a philosophical point of view, but is not very relevant for the actual practice of population forecasting. The fact of the matter is that population forecasts are uncertain and that the 'true' probability distribution - if there is one - is not known. This does not imply that we should not forecast the population nor that we should not aim at assessing the probability that the forecast will come true. Without forecasts and without taking into account the degree of uncertainty of the forecasts we cannot make rational decisions. Thus in making forecasts demographers should not only aim at producing as accurate forecasts as possible, but also at assessing the degree of uncertainty of their forecasts as accurately as possible. The first step in this process is to identify the possible sources of forecast errors.

Sources of forecast errors in population forecasting can be classified according to various criteria. Hoem (1973) distinguishes six sources of forecast errors. Keilman (1990) extends the classification to seven sources of errors:

1. Errors in observed trends. Due to differences between the *de jure* and the *de facto* population, data derived from population registers may not accurately reflect actual demographic behaviour. Moreover, if population registers are absent or incomplete, data from sample surveys have to be used, which will create sampling and non-response errors.
2. Errors in the jump-off population. As population statistics are not perfectly accurate, the population numbers estimated on the basis of population statistics will not correspond perfectly with the population register.
3. Rounding errors. These errors are unimportant unless the forecast involves very small populations at the local scale.
4. Inaccurate model specification. The population projection model might be not entirely valid. For example, if international migration is not included in the model, forecast errors will arise.
5. Randomness in parameters. Observed rates are composed of a trend plus a random fluctuation. Population forecasts are based on a projection of the trend. Hence randomness will lead to forecast errors.
6. Errors in values of exogenous variables. Keilman considers birth and death rates as exogenous variables in cohort-component models. Obviously errors in forecasting births and deaths as well as migration numbers will lead to errors in the forecasts of population size and structure.
7. Sudden shifts in parameters. War, disasters, severe economic depression, and sudden changes in legislation or technology may affect the course of demographic parameters.

The importance of these seven sources of forecast errors for the uncertainty of population forecasts differs strongly. In countries with adequate population registers errors 1 and 2 are small, particularly if the population forecasts are aimed to forecast the *de jure* rather than the *de facto* population. Errors 3 and 4 can be ignored in almost all countries. Error 5 will generally be small in comparison with errors in forecasts of the trends (unless random fluctuations are assumed to affect the trend, e.g. if a non-stationary stochastic time-series model is used to project fertility, mortality, or migration; in that case, however, error 5 becomes a special case of error 6). So errors 6 and 7 are the main sources of uncertainty of population forecasts. Note that error 7 is a special case of error 6: sudden shifts in parameters cause errors in the forecasts of births, deaths, and migration.

Alho (1997) notices that Keilman's classification emphasises the data analytic aspects of error. He argues that, when the emphasis is on *ex ante* errors, it is useful to categorise sources of errors from the point of view of statistical modelling. The categories of errors distinguished by Alho underlie Keilman's sources 5, 6, and 7. Alho (1990) distinguishes four categories of errors:

1. Model misspecification. The assumed parametric model may not be correct.
2. Errors in parameter estimates. Even if the assumed parametric model would be correct, its parameter estimates will be subject to error.
3. Random variation. Even if the parameters could be specified without error, there is residual error because any mathematical model is only an approximation.
4. Errors in expert judgment. In applying statistical models choices have to be made which may create bias in the forecast. Moreover, it is common use in population forecasting to set target values, which usually are largely based on expert judgment.

One benefit of Alho's classification is that he distinguishes between model-based and judgmental forecasts. He does not, however, make a distinction between different types of statistical models, e.g. between univariate extrapolation models and multivariate explanatory models. Since the uncertainty of forecasts depends on the type of method that is used, it is useful to specify sources of forecast errors by the type of forecasting method. Therefore the present paper is based on a more detailed analysis of the sources of uncertainty of population forecasts.

No matter whether causal explanations or observed regularities are used, population forecasts are based on the assumption that certain parameters are constant through time or on an assumption about the way the parameters may change - and on an assumption about the impact of future changes in the economic, cultural, political and technological context.
The parameters refer either to relationships between the demographic variables to be forecasted and explanatory variables or to changes in demographic variables through time, i.e. to *explanation* or *extrapolation* respectively. The uncertainty of forecasts depends on the validity of the assumption that parameters will remain constant in the future or on the justification of *expectations* about the way the parameters will change and on the validity of *expectations* about societal changes which will affect demographic variables.
This also applies for forecasts based on a qualitative approach. True, if a qualitative approach is followed, no explicit assumptions about parameter values are made. In order to arrive at quantitative forecasts, however, some implicit assumptions are made, either referring to changes through time or to the impact of explanatory variables. For example, if it is assumed that fertility trends are determined by cultural changes, such as individualisation, it can be forecasted that cohort fertility will continue to decline, because the individualisation process is assumed to continue, without specifying a quantitative model describing this relationship. Other factors, however, may affect fertility in a different direction. For example, an improvement of child care facilities and paternal leave, or a rise in family allowances may have a positive effect on fertility. This implies that an (implicit) assessment of the relative weights of the impact of the separate factors has to be made. Thus a qualitative approach requires the same type of assumptions as an application of quantitative methods, the main difference being that in a qualitative approach underlying assumptions remain implicit.

One cause of changes in parameters is heterogeneity of the population. Changes in the numbers of births, deaths, and migrants depend on changes in the underlying processes, fertility, mortality, and migration. The rate at which these processes take place differs between population categories, distinguished by age and gender and possibly other attributes (Willekens, 1990). If the composition of the population changes, this may lead to changes in aggregate rates and numbers. Thus disaggregation of demographic variables may provide more stable parameters. For example, if the level of educational attainment is assumed to be a major factor in explaining fertility, fertility rates distinguished by level of educational attainment can be expected to be more stable than aggregate fertility rates. If homogenous population categories can be identified which are characterised by stable parameters, changes in aggregate rates can be explained by changes in the relative size of population categories. On that assumption population forecasts can be based on forecasts of the future size of separate population categories. Thus if fertility is distinguished by level of educational attainment, forecasts of fertility can be based on forecasts of changes in the distribution of the population by educational level.
Demographic rates for homogenous population categories, however, may change in the future due to changes in the social context. When forecasting these changes by an extrapolation method or an explanatory model, the parameters of the model may turn out to change. There are no really universal constants in models of human behaviour. The values of parameters tend to vary between different periods because models provide a simplified picture of the real world. Models cannot include all variables that may affect - either directly or indirectly through other variables - the variables to be forecasted. It can be assumed that omitted variables lead to random fluctuations which can be modelled by the disturbance term. Omitted variables, however, may also cause a change of the future values of the structural parameters. Major shifts in variables that did not seem to influence past developments, may turn out to have a significant impact on future developments. Another possible cause of variation of parameter values is that the functional form of the model may be misspecified. Thus, when making forecasts, explanation and extrapolation have to be supplemented by expectations about the future values of the parameters. Obviously, one can assume that the parameters of the model will remain constant in the future, but it is important to note that this is an assumption, not a fact. Moreover, when using an explanatory model an assumption needs to be made about the expected values of the exogenous variables, i.e. the variables that are not themselves explained by the forecasting model.

In summary we can conclude that the uncertainty of population forecasts depends on the uncertainty of forecasts of fertility, mortality, and migration. The uncertainty of the forecasts of these demographic components depends on the question in what way their future development may be different from the past or different in another way than expected. The answer depends on the methods used for making forecasts. First, appropriate demographic parameters to be forecasted are selected on the basis of the criterion that they should either be stable through time or change in a systematic way or that they should have a stable relationship with explanatory variables (this is the subject of section 3). Subsequently the selected parameters are forecasted on the basis of extrapolation, explanation, and/or expectation. If forecasts are based on extrapolations, we can say that the future is uncertain because the direction of trends may change. If forecasts are based on explanatory models, we can say that the future is uncertain because the values of the structural parameters and/or the functional form of the relationships may change or because the exogenous variables may develop in a different direction than expected, or variables not included in the model may turn out to have an unexpected impact in the future. Sections 4-6 examine the uncertainty of forecasts of fertility, mortality, and migration based on explanation, extrapolation, and expectation respectively.

3. Demographic disaggregation

One key feature of population forecasting is disaggregation. Rather than directly forecasting total population growth, population forecasts are usually based on forecasts of the components of population growth, births, deaths, immigrants, and emigrants. The cohort-component model is used for this purpose. In this model changes in population size and structure depend on changes in births, deaths, and migration, distinguished by age and gender. Forecasts of the numbers of demographic events are usually based on forecasts of rates multiplied by the population at risk, except for forecasts of immigration or net migration as the population at risk cannot be defined in a meaningful way. True, for specific types of immigration information on some population at risk is available. For example, the size of immigration because of family reunion depends on the age/sex composition of the population of given nationalities in the receiving country. But this is not true for other types of immigration, such as labour migration and migration of asylum seekers. Therefore forecasts of total immigration are usually stated in absolute numbers rather than rates. Due to measurement errors another component may be added to fertility, mortality, and migration, viz. corrections in order to achieve consistency between data on population structure and population dynamics. For example, if population size and structure is measured on the basis of a census, the differences between the population at two points in time may not exactly be equal to the aggregation of data on births, deaths, and migration obtained from population registers and/or surveys in the period between the two measurements. Another reason for corrections may be that not all changes in population size and structure can be attributed to either births, deaths, or migration. Particularly migration data tend to be affected by measurement errors. For example, in the Netherlands a considerable part of emigration is not registered directly, as a lot of people moving abroad do not report this to the authorities. If the authorities discover that a person no longer lives at the address where he or she is registered, a so-called administrative correction is made. In the Dutch population forecasts these corrections are added to net migration. Inconsistency between data on population structure and population dynamics leads to measurement errors in age-specific fertility and mortality rates. If the number of people in a certain age group is underestimated, the age-specific rates for that age group will be overestimated. This may be an additional source of uncertainty of population forecasts. Since it is common practice in population forecasting to add corrections to net migration, in this paper corrections will not be discussed as a separate component.

Using the cohort-component model, forecasts of population growth are based on a distinction between its components, fertility, mortality, and migration. Instead population size can be forecasted directly on the basis of forecasts of population growth without disentangling its components. On the basis of data for England and the United States Leach (1981) claims that forecasts of population size based on a logistic curve would have led to considerably less revisions of successive population forecasts than the published official forecasts which reacted to changes in recent changes in the number of births. On the basis of Dutch data Van Alphen (1992) shows that the predictive power of the logistic curve can be improved by allowing the coefficient to vary through time. The parameters can be estimated by using the Kalman filter. Van Alphen concludes that the logistic would have produced more accurate forecasts of total population size than the official forecasts of Statistics Netherlands based on the cohort-component model. In comparing UN projections for the 1950s and 1960s with geometric extrapolations, Stoto (1983) concludes that on average the errors of extrapolations are no larger than those of the cohort-component model. Smith (1997) draws a similar conclusion from a survey of previous empirical studies. However, even though the accuracy of forecasts based on univariate extrapolation models may have been reasonable in a number of cases, this does not imply that they are a serious alternative for the cohort-component model. At the best they can play a useful role as a benchmark for judging the accuracy of forecasts of the cohort-component model. Univariate extrapolations of population size ignore the impact of the current age structure on future population growth and ignore the underlying changes in fertility, mortality, and migration. Moreover, population forecasts are not only used for assessing the future size, but also the age structure of the population.

Keilman (1990) distinguishes six stages in the production of a population forecast:

1. Identification of the population system: which demographic characteristics are to be distinguished?
2. Description of the population system: analysis of historical behaviour of the identified population categories.
3. Model construction: which interactions are to be included in the model?
4. Making assumptions about future values of parameters.
5. Execution and documentation.
6. Monitoring: analysis of forecast errors.

For assessing the uncertainty of population forecasts stages 1, 2, 4 and 6 are crucial.
The analysis of forecast errors has two aims. First, an analysis of errors may lead to the conclusion that a new forecast needs to be made on the basis of revised assumptions. Population forecasts can be updated with a fixed interval. For example, Statistics Netherlands updates the national Dutch population forecasts annually. The long-run assumptions on fertility, mortality, and migration underlying the forecasts, however, are revised after two years at the earliest. Keilman (1990) notes that one benefit of annual or bi-annual forecasts is that it reduces the risk of overreactions of the forecasters. Alternatively, population forecasts may be revised at the moment that the errors of the last forecasts turn out to become ‘too large’. Either way an analysis of errors of previous forecasts should be the first step in preparing new forecasts. The decisions made in one stage are related with subsequent stages. For example, the choice of demographic characteristics in step 1 is related with the analysis of trends in step 2 and the consequences for model construction in step 3. Steps 1 to 3 involve the choice of demographic characteristics which express the heterogeneity of the population. Willekens (1990) notes that “we do not fully

know the decision process of an individual, but we know that people with some attributes are more likely to show some behaviour than people with other attributes.” Therefore population forecasts should be based on a stratification of the population on the basis of personal attributes, e.g. age, gender, parity, marital status, or region of residence. If the population is distinguished by various attributes, population forecasts can be calculated by using a multistate cohort component model. In practice for national population forecasts the ‘simple’ cohort component model is used, whereas the multistate model is used for regional population forecasts or for forecasts of the population by marital status.

One alternative to the multistate model is the microsimulation model. The main advantage of the microsimulation model is that it can incorporate many variables. The main disadvantage is that it requires the estimation of a large number of parameters. Usually there are no sufficient data available. Consequently many assumptions have to be made about the parameters that cannot directly be based on observations. In this paper we focus on national population forecasts that are based on the ‘simple’ cohort component model, taking into account other variables than age and gender by examining their effect on age-specific fertility, mortality, and migration rather than including them in a multistate model.

Demographic analysis is aimed at finding stable parameters, i.e. parameters that either have a stable level or change in a systematic way, e.g. because they have a stable relationship with explanatory variables. Obviously stability of parameters is beneficial for making assumptions about future developments. As demographic rates differ between population categories, disaggregation may produce more stable parameters. Differences by age can be attributed to psychological and biological changes through the life cycle. Due to age differences in fertility and mortality rates the total numbers of births and deaths are affected by changes in the age structure of the population. Therefore the numbers of births and deaths are projected by multiplying age-specific rates by the population at risk. As changes in the age and gender-specific fertility and mortality rates, and migration numbers can be caused by either changes in the age pattern or changes in the intensity, forecasts of the rates and numbers usually are based on a distinction between these two sources of change. For example, changes in age-specific fertility rates can be caused by changes in the average ultimate number of children per woman and by changes in the timing of childbearing. It is important to distinguish both causes of change in making forecasts, because these types of changes have very different implications for the future fertility rates. It should be noted, however, that both types of change may not be independent of each other. For example, postponement of childbearing may lead to a decrease in ultimate family size.

Age differences cannot only be explained by the biological and psychological development of individuals, but may also be partly explained by the common experience of a group of people during their lifetime. For example, fertility behaviour of a cohort of women born in the same period may differ from that of other cohorts, due to circumstances in the period they grew up. Mortality rates at older ages may depend on the health conditions in younger ages. To the extent that demographic rates are more strongly related to the common history of a cohort than to the current social context, a distinction by cohort may lead to more stable parameters than a time series of period data. For each cohort the social and economic situation at each age differs from that for other cohorts at the same age, as a consequence of which demographic behaviour may differ between cohorts. "Each cohort has a distinctive composition and character reflecting the circumstances of its unique origination and history" (Ryder, 1965).

In the cohort approach the average number of children a cohort of women will bear in their lifetime plays a central role. Assumptions are made about the completed fertility of cohorts of women who have not yet completed the child-bearing ages. One advantage of using cohort data for forecasting fertility is that completed family size tends to vary less than annual fertility does. If a birth is postponed during unfavourable economic or social circumstances, it can be expected that the child will be born later when circumstances have improved. Hence a fall in the number of births in some period can be made up in subsequent years.
The most widely used measure of fertility is the total fertility rate (TFR), the sum of the age-specific fertility rates. One important problem in using the TFR for forecasting purposes is that changes in the TFR may be caused by changes in the timing of fertility rather than by changes in the ultimate number of children per woman. Thus projecting changes of the TFR may lead to poor forecasts in the long run. This is, however, an argument against the use of the TFR for projecting fertility rather than against the period approach. If separate age-specific fertility rates are projected, changes in the timing of fertility can be taken into account. In projecting period age-specific fertility rates, however, one should take into account that changes in timing cause age-specific fertility rates at older ages to start to change in later years than fertility rates at young ages. If births are being postponed by young cohorts, in the first stage of the process fertility rates at young ages decline, but the fertility rates at older ages do not yet increase. Only several years later the fertility rates at older ages start to increase. Thus in the early stage of the process changes in the timing of fertility may lead to incorrect projections of period fertility rates. The cohort approach can be preferred if one assumes that fertility is primarily determined by aspirations of cohorts that are rather stable during the reproductive years of the life cycle of the cohort. Differences between successive generations, e.g. due to a rising level of educational attainment, may lead to differences in the timing of fertility and the ultimate number of children per woman.
To the extent that differences between cohorts can be expected to be stable, the cohort approach seems particularly useful for long-run forecasting. However, cohort differences are not always stable, whereas period effects are not necessarily temporary. As fecundity tends to decline with age, a postponement of births does not always imply that the desired child will be born later. Thus changes in the timing of fertility may have permanent

effects. Furthermore, the desired number of children may change during the life cycle due to changing social and economic circumstances. Consequently cohort effects may change. On the other hand period effects may have an permanent impact. For example, changes in policies on child care or parental leave which are implemented in some calendar years and hence can be regarded as period effects, may have an impact on the ultimate family size of cohorts. Hence the improvement of forecast accuracy caused by using cohort rather than period figures may be limited. Keilman (1990) concludes that the change from a period approach to a cohort approach by Statistics Netherlands in 1970 in preparing the national population forecasts did not result in an improvement of fertility projections.

Cohort effects can also be taken into account in forecasting mortality. There are two contradictory hypotheses about cohort effects on old-age mortality, based on different mechanisms (e.g. Van Hoorn and De Beer, 1997). The first hypothesis assumes that there is a selection mechanism. If mortality falls, more weak individuals survive longer. As they are subject to higher mortality rates later in life, mortality rates for the oldest age groups will increase. One alternative explanation of an increase of old-age mortality is that one disease can act as a risk factor for a second disease (e.g. diabetes mellitus increases the risk of cardiovascular diseases). Reductions in mortality in one chronic disease may increase the susceptibility to mortality due to another disease (e.g. Nusselder, 1998). An additional explanation of a negative development of old-age mortality might be the impact of wars: wars can lead to a selection mechanism due to excess mortality of young healthy men. The latter mechanism, however, may only have played an important role in selected countries.
The second mechanism is based on the assumption that experience of morbidity early in life will have negative consequences of mortality later in life, because damage from illness or injury increases the susceptibility for disease and mortality in the future. As a result a reduction of the incidence of morbidity will lead to a decrease of old-age mortality.
Clearly as long as it is not known which interpretation of the mechanism affecting old-age mortality is valid, it is questionable whether the cohort approach of projecting mortality will reduce the uncertainty of forecasts compared with the period approach.

It can be concluded that the uncertainty of forecasts of fertility and mortality depends on three questions:
1. are cohort effects more important than period effects?
2. are cohort effects more stable than period effects?
3. can cohort effects be identified in an early stage?

If these three questions can be answered positively, the cohort approach can be expected to reduce the uncertainty of forecasts in comparison with the period approach. However, these questions cannot be answered simply by empirical analyses. Unfortunately one cannot give a straightforward objective answer on the basis of statistical criteria.
Usually cohorts are distinguished by year of birth, but other types of cohorts can be identified, e.g. marriage cohorts or immigration cohorts. Here demographic rates are distinguished by duration of marriage or duration of residence respectively rather than by age, or in addition to age. As fertility occurs in sequential steps, it may be useful to employ fertility rates in the form of probabilities, such as the probability that a woman who has married in a certain period will have a child in a subsequent period and the probability that a woman who has given birth to a first child will have a second child, etc. If the probability that a demographic event follows another event within a given period is changing, this may be the result of changes in the duration between the intervals rather than of changes in the ultimate probability that the second event follows the first event. Hence it may be useful to analyse duration-specific rates.
If demographic events occur in a fixed sequence, the occurrence of an event may be useful for forecasting subsequent events. Recent observations on the former variables may improve the short-run forecast accuracy of the latter variables. For example, if the interval between first and subsequent births is known, recent changes in the number of first births may be used for forecasting the number of higher parity births in the short run. The number of existing children may be as important as age in forecasting fertility, especially when marriage patterns and timing patterns of fertility by age and birth interval are rapidly changing. For example, if births are being postponed by young cohorts, age-specific fertility rates will decline whereas the progression from marriage to the first birth for successive cohorts may remain constant (e.g. Al-Osh, 1986). Thus one might expect that the use of parity progression rates will improve forecast accuracy, at least in the short run, since the forecast of the number of women having their second child in the near future can be based on the number of women having had one child in recent years, etc. However, age-specific parity progression rates tend to fluctuate much more strongly than total fertility rates. The reason is that the numbers of births refer to relatively small numbers of women as they are distinguished by both age and parity. In the long run the use of parity progression rates implies that the accuracy of forecasts of women having two children depends on that of women having one child. However, it may be more difficult to forecast the percentage of women remaining childless than the percentage of women having two children. Thus the use of parity progression rates does not necessarily lead to a reduction of uncertainty on future numbers of births. Moreover, as recent observations on duration are censored, it is not known whether changes in duration-specific rates are the result of changes in the timing pattern or of changes in the ultimate probability that the second event follows the first. Furthermore, for many countries sufficient data for analysing changes in parity progression are lacking. Even if data on births by age and birth order can be obtained from birth statistics, data on the numbers of women by age and the number of children ever born are usually not available on a yearly basis.
Duration-specific data can also be used for projecting emigration. On the basis of data on emigration by duration since immigration, a projection can be made of the percentage of immigrants that tends to stay. This is useful for

projecting future emigration, given an assumption on future immigration. However, this may only lead to a limited reduction of uncertainty of forecasts on migration, as the uncertainty on the future size of net migration is mainly due to large fluctuations in immigration rather than to fluctuations in emigration.

In addition to cohorts additional sources of demographic heterogeneity can be taken into account. For example, forecasts of fertility and mortality can be distinguished by gender, marital status, ethnic origin or level of educational attainment. In addition forecasts of fertility can be distinguished by parity, mortality forecasts can be distinguished by cause of death, and migration forecasts can be based on a distinction by migration motives.

As population forecasts project the population by gender the assumptions on fertility, mortality, and migration need to be specified by gender. In addition a distinction by gender is useful in assumption making as there are strong differences in demographic rates between men and women. Particularly mortality rates differ strongly between males and females. As trends in life expectancy of males and females may be different, an explicit assumption should be made on future changes in the gender difference of mortality.
For different types of migration the sex ratio is different. For example, the share of men in labour migration is larger than that in family reunification. Thus future trends in migration for men and women may be different.
In cohort-component models fertility rates are related to females only. Usually the sex ratio of births is assumed to remain constant.
Uncertainty in projecting gender-specific data involves the question to what extent gender differences will change in the future. The difference in life expectancy between women and men has not been constant through time. For example, in the Netherlands the difference was much smaller around 1950 than nowadays. Around 1950 the gender difference was only about 2.5 years, which is less than half of the current difference of almost 6 years. Thus the biological effect seems to be limited. The main cause of the widening gap was an increasing trend for men in vascular diseases, lung cancer and other lung diseases. Cigarette smoking is the major common risk factor. Peto et al. (1994) estimate that almost 35% of male mortality in the Netherlands in 1990 is caused by smoking compared with only 5% of female mortality. Valkonen and Van Poppel (1997) estimate that about two-thirds of the gender difference in mortality is due to smoking. Women started to smoke on a large scale much later than men. Consequently the gender difference of mortality has narrowed in recent years. Another cause of the gender difference is the fact that, in contrast with men, for women a lower socio-economic status only leads to slightly higher death rates. One probable explanation is the impact of working conditions on the mortality of men. The current decrease of the gender difference in life expectancy may be caused by more similar lifestyles of the sexes, but apart from smoking behaviour, there is only little evidence yet. Thus it is rather uncertain to what extent the gender difference will continue to decline (Van Hoorn and De Beer, 1997).

Besides age and gender other demographic attributes can be useful for explaining changes in fertility, mortality, and migration. If the majority of couples decide to have their children within wedlock, marital status may be a useful variable for fertility projections. However, as increasingly couples decide to cohabit instead of or before marrying and moreover as cohabiting couples do not decide to marry before they expect to have a child, the usefulness of marital status for forecasting fertility is decreasing. Marital status will depend on fertility rather than the other way around. Moreover, the number of extramarital births has increased strongly. Thus it is questionable whether a distinction of the population by marital status leads to a reduction of uncertainty of fertility forecasts.
For migration marital status can provide information on the potential future size of family reunification and family formation. The future size of family reunification depends on the number of married foreign born persons living alone, whereas the future development of family formation migration can be expected to depend on the number of unmarried foreign-born persons. However, one major uncertain factor is the question to what extent the number of mixed marriages will increase.
Mortality rates differ strongly by marital status. Mortality rates of married people are considerably lower than those of divorced, widowed or never-married persons. A change in the future distribution of the population by marital status can be expected to affect the future development of mortality. However, because of the strong increase of unmarried cohabitation it is uncertain whether the marital status-specific differences in mortality in the future will be the same as in the past. Moreover, the direction of the causal relationship is uncertain. 'Selection' in marriage may be one cause (people with handicaps are likely to marry less than healthy people). 'Protection' through marriage may be another (partners care for each other). Thus it is uncertain what the impact of future changes in marriage behaviour on future mortality will be.
Because of the strong increase of unmarried cohabitation, the number of persons living with a partner seems to be relevant for forecasting fertility, mortality, and migration rather than the percentage being married. Thus forecasts of fertility, mortality, and migration might be based on forecasts of the population by household position. For example forecasts can be made of the percentage of persons living with their partner, persons living alone, persons being a lone parent, children living with their parents and people living in other household situations. However, one major problem in making these types of forecasts is that there are considerably less reliable data on (changes in) household positions than on fertility and mortality (De Beer, 1994). Moreover, changes in household positions may be less stable than changes in fertility and mortality. Hence it is questionable whether the uncertainty of forecasts of fertility and mortality can be reduced by founding them on forecasts of household positions.

In forecasting changes in fertility a distinction by parity can be made. Analyses of separate time series of fertility rates for first, second, and following births provide a better insight than total fertility rates. When taking parity into account, a distinction can be made between fertility rates by birth order and parity progression rates. Age-specific

fertility rates by birth order refer to all women in separate age categories, whereas parity progression rates refer only to women with a given parity. In projecting fertility rates by birth order one can take into account whether a decline of the average number of children per woman is caused by a decline of the percentage of women having large families or by an increase of the percentage of women remaining childless. This distinction is important for making forecasts as one may assume that both types of change are caused by different social and economic processes.
Using parity progression rates, the percentage of women having two children is based on the percentage of women having one child, the percentage having three children on that having two children, etc. Although from a theoretical point of view this type of analysis seems to be preferable to using more aggregate data, in practice this does not necessarily lead to an improvement of forecast accuracy, as was argued above.

Demographic behaviour differs between ethnic groups. For example, the level of fertility of Turkish and Moroccan women residing in the Netherlands is considerably higher than the fertility level of the Dutch-born population. Thus forecasts of the total future level of fertility depend on forecasts of the future size of ethnic groups and on forecasts of changes in fertility level of individual groups. However, both types of forecasts are rather uncertain. First, the future size of ethnic groups depends partly on future immigration, which is difficult to forecast. Second, the fertility of several ethnic groups changes more strongly than that of the total population. Future fertility of ethnic groups do not only depend on societal trends that affect total fertility, but in addition on the way the process of integration evolves. Furthermore there are less reliable data on fertility by ethnic origin than on total fertility. Thus it is questionable whether separate forecasts of fertility for individual ethnic groups add up to a more certain forecast of total fertility than a direct forecast of total fertility.

Changes in the total size of immigration are the result of different trends in the size of various immigration categories. Different types of migration may be affected by different factors. For example, changes in labour migration can be explained by the situation on the labour market both at home and abroad, family reunion and family formation are affected by the size, age structure and household situation of the resident foreign population, and changes in the numbers of asylum seekers and refugees depend on the occurrence of armed conflicts and the violation of human rights and on immigration policy in the receiving countries. One main problem in making assumptions on future migration on the basis of migration categories distinguished by motive is that there are severe data problems. For example, in the Netherlands migration statistics are based on population registers. In principle, non-native people intending to stay in the Netherlands for at least six months are recorded in the register and counted as immigrant. The migration statistics only provide information on demographic variables, such as age, sex and marital status, and not on the motive of migration. As to asylum seekers, the Ministry of Justice provides monthly information on the number of asylum requests. However, only part of the asylum seekers is registered as immigrant and hence counted in the migration statistics. Asylum seekers stay in special centres during the first months after their arrival in the Netherlands. Only when they move to a more definite address, they are registered as immigrant. As yet there are only rather crude estimates of the numbers of asylum seekers that are included in the migration statistics. Although the migration statistics do not contain information on migration motives, on the basis of an analysis of demographic variables such as age, sex, marital status and marriage duration that are available from migration statistics, the size of family reunion and family formation can be estimated (De Beer et al., 1993). Family reunion and family formation migration can be distinguished on the basis of assumptions on the demographic characteristics of people belonging to both categories. For example, it can be assumed that women arriving with children but without spouse, their children, most of the minors arriving as individual and individually arriving women who have been married for a long time can be considered as family reunists. Individually arriving men and women who have recently married are likely to be family formation migrants. People who marry shortly after their arrival or people who enter into a consensual union are also included in the category of family formation.

In projecting mortality causes of death can be taken into account. In theory one might expect that a distinction by cause of death would improve forecasts of mortality. However, in practice the benefits have not yet been made clear. For instance Alho (1992) concludes that using causes of death did not improve the USA-forecasts. Especially when non-linear models are used, implausible figures for rapidly changing causes of death may result. Other authors arrived at the same conclusion (Murphy, 1990, and McNown and Rogers, 1992).
An important problem is the extrapolation of causes of death which have changed rapidly in recent years. Within a few years the projections tend to become rather extreme. Causes of death that are declining rapidly tend to disappear completely, whereas causes of death that are increasing tend to become predominant. When projecting mortality by cause of death the results depend strongly on the choice of the extrapolation method (linear or non-linear) and on the choice of the base period. For example, in studies that project death causes for several European countries, it appears that in many cases projected trends for different countries cross each other (Caselli, 1993 and Tabeau et al., 1997). This is caused by the fact that the projected mortality rates by a certain cause of death for countries which started at a high level of mortality and which have witnessed a downward trend in the recent past, may become lower than the projected mortality rates of countries with constant low mortality or than that of countries in which mortality was very low and recently showed a moderate upward trend. For short-term forecasts which are usually mainly based on an extrapolation of recent trends, it is doubtful whether the use of causes of death leads to quite different results than extrapolating total mortality (Van Hoorn and De Beer, 1997). If trends of the important causes of death and of total mortality are rather regular and more or less linear and for all causes of death and overall mortality the same extrapolation method is used, one can expect that rather similar projections of life expectancy for the near future will result. One major problem in

projecting causes of death, is the substitution between different causes of death, i.e. the impact of a decline in one cause of death on the occurrence of other causes of death. Up till now, no satisfactory solution was found for this phenomenon. Usually independence of mortality from different causes of death is assumed, but that is probably not a correct way of dealing with this problem. There are indications that positive as well as negative relationships exist which can also correlate with specific risk factors and even with medical treatment. Probably the underlying general health plays an important role. This may cause a strong interdependency between causes of death (Van Hoorn and De Beer, 1997).

Obviously combinations of characteristics are possible. For example, one might distinguish fertility by cohort as well as parity, marital status and ethnic origin. The choice of the appropriate level of aggregation depends on three criteria (Willekens, 1990):

1. stability: a regular pattern of change enhances predictability;
2. explanation: a stable relationship with explanatory variables is helpful for forecasting;
3. data availability: if limited data are available, no detailed disaggregation is possible.

A fourth criterion may be added: user's requirements (Keilman, 1990). However, the present paper is restricted to national population forecasts by age and gender. In this context other variables are only relevant to the extent that they are useful for improving the forecasts of the population by age and gender (the second criterion). Generally heterogeneity may have an adverse effect on finding an explanation. Heterogeneity, however, does not necessarily lead to less stable parameters. Changes in aggregate demographic rates can be attributed to either changes in the size of separate population categories if these categories have different rates or to changes in the rates for these categories. If the rates for the separate categories do not change, aggregate rates vary due to changes in the relative size of population categories. In that case heterogeneity leads to less stable rates. However, if the rates for the separate categories change, the aggregate rates may well change less than the rates for some categories. For example, if there is convergence between population categories, and the rates of separate groups move into the direction of an 'innovative' group, the aggregate rate may well change less than the rates of some of those groups. Also if different population categories are changing their demographic behaviour in opposite directions, aggregate rates may change less than detailed rates (Willekens, 1990). Generally, if the changes in demographic rates for each category are relatively large compared with changes in the relative size of the categories, aggregate rates may be more stable than disaggregated rates. Moreover, disaggregated rates are more strongly subject to random fluctuations due to smaller numbers of events. Smith (1997) argues that disaggregation does not necessarily produce more accurate forecasts, as it may distract from long-term historical trends. Lee et al. (1995) claim that disaggregation may conceal dynamic regularities. The conclusion must be that the selection of an appropriate level of disaggregation is not an obvious, unambiguous, straightforward choice. The degree of uncertainty of demographic rates at the aggregate level cannot not simply be assessed by adding up the uncertainty of all underlying causes of change.

Once demographic characteristics are selected, assumptions have to be made about the future values of the identified demographic parameters. Keilman (1990) distinguishes four types of assumptions:

1. disclaimers: the forecasts are not valid in the event of war, natural disaster or major economic crisis;
2. general assumptions about the social, economic, cultural, and political context;
3. assumptions about the future values of summarising indicators, such as the total fertility rate, life expectancy and net migration;
4. detailed assumptions about age and gender-specific fertility, mortality, and migration rates or numbers.

The general assumptions are based on explanations of demographic variables by non-demographic variables. Usually these assumptions are qualitative. On the basis of these general assumptions, quantitative assumptions about the future values of summarising indicators are specified. These assumptions may be based on quantitative explanatory models, but in practice they are usually based on extrapolations. Detailed assumptions are almost always based on extrapolations. Both forecasts based on explanations and extrapolations involve expectations. Forecasts based on explanations involve expectations about the future development of the exogenous variables. Extrapolation is based on expectations about the length of the period for which extrapolation of observed trends is justified. The uncertainty of forecasts based on explanation, extrapolation, and expectation is discussed in the next three sections.

4. Explanation

In making population forecasts one has to decide whether one assumes that observed changes in fertility, mortality, and migration will continue in the future. Preferably such a decision should be based on an understanding of the underlying causal mechanisms that can explain the observed developments. Willekens (1990) even states that "understanding is the precondition for effective forecasting (…) Demographic forecasting should be rooted in an understanding of the causal processes at work." This implies that the uncertainty of population forecasts depends on the uncertainty about the validity of theories that are assumed to explain developments in fertility, mortality, and migration.

As Blaug (1980) argues, "no social science can boast of the universal laws of modern chemistry, the numerical constants of particle physics, and the predictive accuracy of Newtonian mechanics. The comparison between social and natural science looks a little better in terms of biology, geology, physiology, and meteorology, but even there is still a long gap between our knowledge of human behaviour and our knowledge of natural phenomena." For at least two reasons social science cannot be the basis for population forecasts the way natural science can predict natural phenomena:

1. human actors differ from the physical objects of natural science;
2. demographic developments cannot be modelled by a closed system.

Ad 1. Human behaviour depends on motives and intentions rather than physical causes. Some philosophers, particularly German philosophers in the tradition of Heidegger, drew the conclusion that human behaviour cannot be explained by causal theories, but that one should aim at 'understanding' (usually the German term 'Verstehen' is used) or interpreting human behaviour. They emphasise that social science must be based on first-person knowledge by means of empathy instead of third-person knowledge by means of observation (Blaug, 1980). Clearly this would exclude the possibility of objective forecasts. This view, however, ignores that social science has developed a number of theories that have been successfully applied in widely different situations. For example, in economics the law of demand and supply, the law of diminishing returns and the law of diminishing marginal utility are universal relations that have been corroborated extensively by observations. In demography the concept of the demographic transition may be regarded as an example ("one of the best-documented generalizations in the social sciences," Kirk, 1996), even though there is some discussion about the underlying theoretical explanation (e.g. Johansson, 1997). The other extreme position is taken by behaviourists in the tradition of Skinner and bio-sociologists like Wilson who argue that human behaviour can be explained by cause and effect. They claim that human behaviour is perfectly predictable if we know the causal mechanism and the occurrence of a cause. However, as Blommestein and Van Vught (1988) argue, "it is inconceivable that we will ever be able to make an inventory of all relevant initial conditions or to formulate all the relevant laws. Our theories remain simplifications of and conjectures about an endlessly complex reality (...). There will always be initial conditions of which we had not thought about in advance." Similarly Boudon (1986) argues that the great number of factors that affect human behaviour makes it unpredictable. Moreover, interaction between individuals has to be taken into account. Individual behaviour depends on the social context. Thus social change is not simply an addition of individual actions. In addition, chaos theory argues that even if behaviour is deterministic and if we know the true theory, future developments may still be unpredictable (e.g. Hall, 1992; Gleick, 1987). If developments are determined by non-linear dynamic processes, very small differences in the initial conditions can lead to very large differences in the long run. Thus both the view that human behaviour is absolutely unpredictable and the view that - in theory - human behaviour is perfectly predictable seem untenable. Instead of taking an extreme position it is more useful to attempt to assess the degree of predictability of human behaviour.

Ad 2. Developments in fertility, mortality, and migration are influenced by economic, social, cultural, and political developments. Thus the predictability of demographic changes depends on the predictability of changes in the social context. Even if demographic changes could perfectly be explained, forecasts would still be uncertain as future social developments are uncertain.

Each of the three components of population growth, fertility, mortality, and migration, can be explained by various theories. Even though each theory can be supported by empirical evidence (otherwise a theory would not survive), there is no theory capable of explaining all changes or differences. Some theories supplement each other, while others are mutually exclusive. Thus forecasts cannot be based in a straightforward, unambiguous way on explanations. The forecaster has to make choices which cannot be based on purely objective criteria.

In explaining changes in fertility two economic theories have received much attention during the last three decades.

1. The new home economics theory developed by Becker (1960) emphasises that raising children is more time-intensive than other activities. Starting point is the assumption of maximisation of utility by the household. Besides the usual income restriction, Becker added time as a restriction. As raising children costs relatively much time, the costs of children are determined to an important extent by the price of time. Since women tend to spend more time on the raising of their children than men, the income that a woman could earn if she participated in the labour force has an impact on fertility. An increase of the wage rate of women has a negative impact on fertility. Moreover, the income effect of the men may not be positive either. If the household income increases, the desired 'quality' of children may increase, i.e. the expenditure per child may increase. If the income elasticity of the 'quality' of children is larger than that of the total expenditure to children, an increase of income will lead to a decline of fertility.
2. The relative income hypothesis developed by Easterlin emphasises the role of income in relation to aspirations (e.g. Easterlin, 1976; Lee, 1976). If people perceive themselves as having a relatively high income, the wife works less and fertility is high. The aspiration level depends on wealth during the youth in the parental home. Easterlin assumes that the relative income of people belonging to a certain cohort depends on the size of that cohort in relation to other cohorts. The economic situation for relatively large cohorts is less favourable than for smaller cohorts. If it is assumed that in the demand of labour there is no perfect substitution between different age categories, the labour market situation for people belonging to a large cohort is less favourable and hence the perspective of income is less favourable. Thus the Easterlin hypothesis implies a negative relationship between the level of fertility and the size of the cohort. This theory can explain a cyclical movement of births. The small generations of the 1930s produced many children in the 1950s, who had relatively low fertility in the 1970s. One problem in assessing the validity of the Easterlin hypothesis is that it is confirmed by one cycle only. Empirical research on the micro level does not indicate that relative income (defined as the income compared with the income of the parents) is an important explanatory variable (e.g. Jencks, 1981). Butz and Ward (1979) claim that relative income does not play an important independent role in explaining the development of fertility through time on the macro level either, whereas Devaney (1983) concludes that even though relative income has a significant impact on fertility, the level of income of women is the dominant factor in explaining the development of fertility in the United States:

"Hence projections of U.S. fertility depend crucially on the future path of female wages" (Devaney, 1983). However, she adds, "unfortunately, it is extremely difficult to predict the path of female wage rates." Keyfitz (1982) points out that even though the Easterlin effect may exist, it may not be manifest in the future development of births due to the occurrence of other effects, such as the entry of women into the labour market. De Beer (1991b) finds that although the Easterlin hypothesis is not rejected by the data, a declining long-term trend in average family size is predominant in explaining changes in births after World War II in the Netherlands.

One objection against both the new home economics model and the Easterlin hypothesis is that the emphasis is on economic factors, regarding reproductive behaviour primarily as rational behaviour, based on weighing costs and benefits of having children. Leibenstein (1980) questions to what extent fertility behaviour is rational: "Some people may be almost universal maximizers, but most people are not. At most, most people are maximizers only part of the time."

Another problem in following a purely economic approach is that social and cultural changes have been strong during the last decades. Even though an economic theory of fertility may be correct, the conditions have changed that much that the forecasting power of a purely economic model seems to be limited. For example, the decline of fertility in the 1960s and 1970s can be explained on the basis of the new home economics model by assuming that the negative relationship between fertility and the income level of women has become stronger due to changing attitudes toward the labour force participation of women. However, the latter development is not explained by the economic model of fertility. In explaining changes in fertility one should take into account social and cultural changes, such as secularisation, urbanisation, emancipation, modernisation, and individualisation.

In explaining changes in mortality three types of explanatory factors can be distinguished.

1. Evolutionary theories aim to provide a basic scientific theory of the age pattern of mortality. Tuljapurkar and Boe (1997) conclude that various studies strongly argue against the existence of an upper limit to the lifespan of all individuals of any species. These theories focus on the 'intrinsic' character of the ageing process, as determined by biochemical, genetic, cellular or physiological processes in individuals. Biomedical studies focus on the relationship between morbidity and mortality. Illness plays an important role in survival. Mortality rises with increasing functional and cognitive disability. Survival to advanced ages may depend on avoidance of severe illness. Illness in childhood may have effects on mortality late in life.
2. Changes in mortality can be explained by changes in the economic, social and physical environment. The decline of mortality rates that has started more than a century ago can be explained by improvements in public health, advances in medical treatment, improved hygiene and rising standards of living (e.g. Santow, 1997). Before the second World War most progress in reducing mortality occurred at younger ages. This was largely attributable to public health improvements and rising standards of living rather than to medical care. Post-war improvements in old-age mortality may have been linked with biomedical advances (such as antibiotics and treatment for heart diseases) and other developments in the provision of health care. There is only little direct 'proof' that the quality of the environment affects mortality (WHO, 1995). The most important environmental factor appears to be air pollution caused by SPM (suspended particulate matter; smog and aerosols). According to the WHO study the high mortality in Eastern Europe is mainly caused by socio-economic factors rather than by the poor quality of the environment.
3. Changes in mortality through time or differences in mortality between individuals can be explained by changes or differences in behaviour. Smoking, obesity, alcohol abuse, and diet are important determinants of longevity. Differences in life styles are strongly related with education, occupation, income, and wealth. Mackenbach (1988) claims that life style changes are at least as important as medical care and the introduction of new medicines.

One major problem in assessing to what extent mortality is affected by these factors is that there are multiple causal pathways. For example, the gender difference is partly caused by biological factors, but another part of the differences can be attributed to behaviour. Another example: the high death rates of the lower social classes may be partly due to individual behaviour (much smoking, unhealthy eating habits), but also to bad working and living conditions, social stress and even childhood circumstances, for instance negative factors during mother's pregnancy. However, also selectivity by health may play a role. Unhealthy people have more difficulties in obtaining a good socio-economic position.

Massey et al. (1993) give an overview of theories of international migration. They distinguish five theories explaining why international migration begins and three theories explaining the perpetuation of migration:

1. The neoclassical economic macro theory assumes that geographical differences in the supply and demand of labour are the major factors causing migration.
2. The neoclassical micro theory assumes that decisions to migrate are based on a cost-benefit calculation by rational individuals. Differences in earnings and employment rates are key variables.
3. The new economics of migration views migration as a family strategy to diversify sources of income. In contrast to the neoclassical theory, wage differentials are not seen as a necessary condition for international migration. Economic development in areas of origin will not necessarily reduce pressure for migration.
4. According to the dual labour market theory demand for low-level workers in developed countries is the main factor causing international migration. Employers seek low-wage migrant workers in order to avoid raising wages of native workers. In this model international migration is demand-based and initiated by recruitment of employers or governments in destination countries. Wage differentials between origin and destination countries are neither necessary nor sufficient conditions for migration.

5. World systems theory focuses on the structure of the world market economy. Globalisation facilitates international migration. International migration is affected less by wage or employment differences between countries than by the international flow of capital and goods.
6. Network theory emphasises that migrant networks reduce the costs and risks of international migration and this increase the likelihood of migration. The development of such networks depends on government policies toward family reunification. Once started, migrant networks can make international migration relatively insensitive to policy interventions.
7. Once international migration has begun, institutional theory points to the fact that private organisations develop to support and sustain the movement of migrants by providing transport, labour contracting, housing, and legal services.
8. According to the cumulative causation theory the establishment of international migration flows creates feedbacks that make additional migration more likely. For example, low-skill jobs may be labelled as 'immigrant jobs'.

These theories are concerned largely with migration flows that occur more or less voluntarily in response to economic and social conditions. They do not address migration that results from political factors (violence, human rights violation, repression) (Russell, ...). On the basis of empirical research Edmonston (1992) finds that refugee migration depends on threats of violence, authoritarian government structures, and worsening socioeconomic conditions. However, he concludes that there is no clear, single cause of refugee migration.
Macura (1994) identifies three major causes of international migration: income differentials (contributing to South-North migration pressures), political developments (e.g. the end of the Cold War, causing East-West migration, and armed conflicts, giving rise to large flows of refugees) and rapid population growth in the developing countries (increasing emigration pressure from Third World countries). It is likely that these causes will lead to a continuing migration pressure on developed countries in the future. However, many receiving countries have responded to the increase of immigration by tightening immigration policies. It depends on the effectiveness of these policies whether the increasing migration pressure will lead to increasing immigration.
Massey et al. (1993) conclude that alternative theories of international migration are not contradictory or mutually exclusive. For example, the cost-benefit model assuming that migration is an aggregate outcome of individual decisions that are based on a rational evaluation of costs and benefits ignores social and economic structures on national and international levels that affect migration also (Massey, 1990). Thus the micro and macro theory supplement each other. In a review of empirical studies of international migration in North America, Massey et al. (1994) conclude that each of the surveyed theoretical models received at least some empirical support: "What is unclear is how well the various models perform against each other, and how much of an independent contribution to explanatory power each model might retain in a simultaneous examination of theoretical propositions."

If forecasts of fertility, mortality, and migration are based on explanations, one major problem is that there is no unambiguous relationship between theories and explanatory variables. On the one hand several theories contain the same variables. Hence the interpretation of an empirical relationship is not always obvious. For example, the level of educational attainment serves as explanatory variable in many theories explaining fertility and mortality. The educational level may appear as an indicator of human capital, but this variable may also represent the impact of attitudes. On the other hand if one theory is operationalised, alternative explanatory variables may be used to represent unobserved theoretical concepts.

Once explanatory variables are identified, they can be used for population forecasts in three ways:
1. disaggregation;
2. qualitative arguments;
3. quantitative explanatory model.

Ad 1. Disaggregation can be based on an identification of explanatory variables. For example, fertility rates can be distinguished by the level of educational attainment. Assuming that the fertility rates of the separate education categories remain constant, changes in fertility can be explained by changes in the distribution of the population by level of educational attainment. By the same token, changes in mortality rates can be explained by changes in the size of population categories characterised by different life styles. For example, a distinction can be made between smokers and non-smokers (the latter category should be subdivided into those who have never smoked and those who stopped smoking). Because of data problems, disaggregation may be based on proxies rather than on the explanatory variables directly. For example, instead of identifying categories on the basis of several variables, fertility rates can be distinguished by cohorts which may be assumed to reflect the impact of the social context on fertility behaviour (Willekens, 1990). Similarly mortality rates may be distinguished by cause of death as an alternative to a distinction of the population by risk factors.
In general, there are three sources of uncertainty of forecasts based on disaggregation. First, it is uncertain to what extent the demographic rates of the separate population categories will change. Second, the future size of the different population categories is uncertain. Third, the interdependency between the distinct demographic rates is uncertain. For example, in forecasting mortality distinguished by cause of death, one has to take into account the mutual dependency of changes in the separate causes of death. A decline of mortality by one cause of death may lead to an increase in deaths by another cause.

Ad 2. Assumptions underlying population forecasts can be based on qualitative considerations. Forecasts of fertility, mortality, and migration can be based on qualitative assumptions about the future direction of economic, cultural, political, and technological developments. In order to obtain quantitative forecasts of fertility, mortality, and migration additional assumptions have to be made. These assumptions can be based on extrapolations which are consistent with the qualitative explanations. For example, on the basis of an evaluation of positive and negative factors determining changes in mortality, one may conclude that the increasing trend in life expectancy will continue. Alternatively the assumptions can be based on experts' expectations. Usually experts' opinions refer to levels of fertility, mortality, and migration to be reached in a so-called target year. The values of the three components in the period up to the target year can be obtained by means of an interpolation method between the current and the target values. Uncertainty of forecasts based on qualitative considerations arise because both the direction of the future development of the non-demographic variables and because the size of their effect on fertility, mortality, and migration are uncertain.

Ad 3. Forecasts of fertility, mortality, and migration can be based on quantitative explanatory models. Three types of explanatory models can be distinguished:

a) time-series models explaining changes through time;
b) cross-section models explaining differences between individuals, population categories, regions or countries at the same point in time;
c) longitudinal models, explaining differences between observational units at several points in time.

Ad a. Annual changes in fertility, mortality, and migration can be related to changes in other variables. For example, various empirical studies of fertility examine the impact of income variables, such as average real wage level of men and women or estimated permanent income, and the labour force participation rate of (married) women (see e.g. Gregory et al., 1972; Venieris et al., 1973; Ward and Butz, 1980; Ermisch, 1979; Wilkinson, 1973, Madduri and Gupta, 1974; Brooks, Sams and Williams, 1982). Changes in international migration can be related to differences in unemployment and income between countries of origin and destination (e.g. De Jong and Visser, 1997). One main problem in specifying multivariate time-series models is that there is a considerable risk of estimating spurious relationships. For that reason, 'error-correction' models can be applied. These models take into account both the relationship between the level of demographic variables and explanatory variables and the relationship between the changes. These models have not yet been widely used for explaining demographic variables. One notable exception is the model developed by Ermisch (1992) to explain fertility rates.

Ad b. Since population forecasts aim to project changes through time on an aggregate level, the use of macro models based on time series may seem a logical choice. The processes that determine changes in fertility, mortality, and migration, however, are active at the micro level. "Changes in the socio-economic, cultural and legal context are mediated by micro-level processes"(Willekens, 1990). Analyses of individual data are needed to identify the behavioural processes that underlie observed demographic changes. An advantage of micro models compared with aggregate models is that the coefficients in micro models correspond more closely to the theoretical concepts. One main problem in using micro models, however, is lack of data. Particularly for mortality hardly individual data on explanatory factors are available. At the most individual data for some specific causes of death are available. Therefore cross-sectional analyses of mortality are usually limited to analyses of differences between regions or countries. Another important problem in analysing micro data is that usually only very limited data on changes through time are available. Moreover, even if individual behaviour is explained, the consequences on the macro level are not known. The macro level is not simply the sum of individual behaviour because of the effect of the macro context at the micro level. To some extent changes in fertility have occurred for different groups simultaneously: the 'mood of the times' (Jencks, 1981) is an important factor. Jencks emphasises that, as it is difficult for couples to decide beforehand how having children will be, their decision depends heavily on their social environment.

Analyses of micro data can provide information on determinants of fertility, mortality and migration be means of examining differences between individuals. For forecasting purposes the question is how differences between individuals at one point in time can be translated into changes through time (Keyfitz, 1982). If several social categories are distinguished by level of fertility, forecasts of the fertility level at the macro level can be based on a forecast of the relative size of the separate social groups if it is assumed that the differences in fertility between the social groups remain constant. For example, a latent class model can be used to identify different types of life styles. If an assumption is made on the relative size of the groups with different life styles for a young birth cohort, the level of fertility of that cohort can be forecasted (Vermunt, 1991a). If subgroups are identified, no separate projections of all explanatory variables are needed. It suffices to project the relative size of the subgroups. However it is questionable whether changes in the level of fertility are completely due to changes in the relative size of different groups assuming the level of fertility of each group to remain constant. Alternatively, if some social category is considered as a 'forerunner' and it is assumed that the other categories will show the same behaviour in the future, forecasts of the macro level of fertility can be based on forecasts of the rapidity of the convergence of the fertility rate. On the basis of a diffusion model it can be assumed that 'new' types of behaviour are first adopted by 'forerunners', e.g. highly educated, urban, non-religious people, followed later by other groups (De Feijter, 1991). To the extent that assumptions on future fertility can be based on behaviour that is already actually shown by some (groups of) individuals uncertainty of forecasts can be reduced. However, relationships between demographic variables and explanatory variables may not be constant. One possible cause is that cross-sectional differences may be affected by cohort effects. It is difficult to disentangle age patterns and cohort effects. In a cross section behaviour at different ages is observed for different cohorts. It is uncertain whether the relationship between demographic and explanatory variables at each age is the same for each cohort. If a survey provides information on the values of the explanatory variables at the moment of the survey only, a static model can be constructed, explaining differences in the demographic variables, e.g. the number of children ever born, by differences in the explanatory variables. To some extent dynamics can be taken into account by means of including interaction between age and the explanatory variables (Calhoun and De Beer, 1991). In both cases one should also take into account general fertility trends that are valid for all social categories simultaneously. Usually micro models can only explain a small fraction of the variance. Only if changes in the non-explained differences of behaviour cancel each other out, these models can produce reliable forecasts on the macro level. It is, however, questionable whether these changes are not correlated between individuals. General social trends can be expected to affect the behaviour of many individuals in a similar way. One source of uncertainty of forecasts based on analyses at the micro level is that usually only limited time-series data are available. One alternative may be an analysis at the meso level, e.g. an analysis of regional differences. If on the basis of time series analyses of regional data vanguard regions can be identified, information on those regions can be used for forecasting the other regions. Uncertainty comes down to the question to what extent other regions will follow the same pattern. Obviously this is similar to looking at international differences. However, one may expect cultural differences between regions within the same country to be smaller than differences between countries.

Ad c. If a survey provides longitudinal data, event history analysis can be applied to estimate the probability of a particular event during a given period, e.g. childbearing within a year. For example, De Beer et al. (1991) construct a hazard rate model that explains the probability of a first birth within a year. This model makes it possible to calculate the rate of permanent childlessness on the basis of assumptions of the future values of the explanatory variables. When using hazard rate models the effect of changes in the explanatory variables can be taken into account, whereas on the basis of cross-sectional data the static situation rather than the underlying dynamic processes is analysed. On the basis of estimated transition probabilities that depend on individual attributes a microsimulation model can be applied for forecasting. In such a model a Monte Carlo mechanism determines which individuals belonging to a 'model' population experience an event in a given year. Then the population is updated and the process is continued. Using microsimulation models makes it possible to take into account the effects of different variables simultaneously. However, one problem in using microsimulation for forecasting is that it requires assumptions to be made about a large number of parameters which often has to be based on few observations.
If longitudinal observations are available, cohort effects can be separated from the age pattern. However, if used for forecasting purposes, it is questionable whether the age and cohort effects can be assumed to remain constant. Particularly for young cohorts the estimates are rather uncertain. Moreover, if there is interaction between age and cohort effects, assumptions have to be made on the future values of the interaction effects. Furthermore event history analysis is mainly based on data collected in retrospective surveys. The quality of the data for the explanatory variables often is doubtful due to errors of omission. Moreover, it is difficult, if not impossible, to collect reliable retrospective information on attitudes (Keilman, 1991).

In identifying and applying a quantitative explanatory model for forecasting at least six sources of uncertainty can be distinguished:
1. uncertainty about the choice of the explanatory variables;
2. uncertainty about the distinction between endogenous and exogenous variables;
3. uncertainty about the operationalisation of unobserved variables
4. uncertainty about the functional form of the relationships;
5. uncertainty about the values of the structural parameters;
6. uncertainty about the future values of the exogenous variables;
7. uncertainty due to random errors.

Ad1. The choice of the explanatory variables.
The choice of the explanatory variables can be based on theoretical considerations and empirical tests. In addition availability of data is an important restriction. As there is no single theory explaining fertility, mortality, or migration, the forecaster should either choose between alternative theoretical approaches implying different explanatory variables or use variables from different theories. One major problem in using theories for forecasting is that we will never know for certain whether a theory is true. Empirical tests will not be conclusive. If forecasts based on some theory do not correspond with observations this can be explained by claiming that the *ceteris paribus* condition is not fulfilled. “[...] it is always a scientific theory *together* with this [ceteris paribus] clause which may be refuted. But such a refutation is inconsequential for the specific theory under test because by replacing the *ceteris paribus* clause by a different one the specific theory can always be retained whatever the tests say” (Lakatos, 1970). If forecasts conflict with observations it can be concluded that apparently other factors (“possibly hidden in some distant and unspecified spatio-temporal corner of the universe”, Lakatos, 1970) have a disturbing influence. Subsequently one may search for an explanation of those disturbances. For that explanation the *ceteris paribus* clause applies also, etc. In general, theories need not be rejected if not all forecasts are correct: “a principle of falsification that removes theories because they do not fit the facts would have to remove the whole of science (or it would have to admit that large parts of the whole of science are irrefutable)” (Feyerabend, 1975). One example is the discussion by Keyfitz (1982) of the Easterlin theory. Keyfitz argues that even though the Easterlin theory (fertility depends on the family’s economic status relative to that of their parents and the economic status is negatively correlated with cohort size, which leads to a negative relationship between cohort size and its level of fertility) may be true, its effect is likely to be swamped by other relations, e.g. the emancipation of women. In other words, according to Keyfitz, if the predicted relationship between cohort size

and fertility does not actually occur, this does not necessarily imply that the Easterlin theory is not true. Thus a forecaster cannot decide on purely objective grounds which theory should be used for forecasting. “Debates over theory-choice cannot be cast in a form that fully resembles logical or mathematical proof” (Kuhn, 1970a). This does not imply that theory choice is irrational. There are “good reasons for theory choice. These are reasons of exactly the kind standard in philosophy of science: accuracy, scope, simplicity, fruitfulness, and the like” (Kuhn, 1970b). But Kuhn emphasises that these are values and not hard criteria. They can be applied in different ways by different researchers.

The fit of a model to a given set of observations does not give an unambiguous answer to the question which model should be used for forecasting for at least two reasons. First, as Keyfitz (1972) points out, “if the data come from the past, and the inference or prediction concerns the future, we can never assume that data and inference apply to exactly the same population.” Keyfitz (1982) notes that much research is based on cross-sectional comparisons. This allows conditional conclusions based on comparative statics, whereas a forecast requires dynamic and unconditional conclusions. Secondly, as Theil (1966) remarks, “[there is] no fundamental law stating the disturbances are small or ‘as small as possible’; it is far from sure whether the small residuals are due to the characteristics of the population or the investigator’s tenacity.”

The fundamental problem in looking for explanations in order to use them for forecasting is that correlations between variables can be observed but not causality. One can only speak of a cause against the background of a theory. An empirical relationship between variables is no sufficient condition for causality. Granger (1969) proposes a practical solution by defining causality in terms of conditional probabilities. The variable Y is no cause of X if $P(X_{t+1}|I_t) = P(X_{t+1}|I_t\text{-}Y_t)$, where I_t is all information available at time t and I_t-Y_t is all information excluding Y_t. The basic assumption underlying this definition is that the future cannot cause the past. Obviously this definition requires additional assumptions about the underlying distribution and the information. The choice of the information is determined by the assumed theory and the availability of data. Since usually the distribution is not known, one may use conditional expectations, which make it possible to define causality in terms of predictions. When it is assumed that the prediction is a linear function of the elements of I_t, the occurrence of causality can be tested on the basis of multivariate time-series models (Granger and Newbold, 1977). Even if this practical approach is followed, however, tests of causality depend on the choice of variables and the functional form of the model. For example, causality may appear due to the omission of a variable that is correlated with both Y and X. Thus empirical tests of causality depend on assumptions of which the validity is uncertain. This does not imply that looking for causality is useless. When making forecasts, the search is for regularities that do not appear in one specific historic setting only, but that have predictive power for other situations, i.e. regularities that have a law-like behaviour. “The belief in causality is metaphysical. It is nothing but a typical metaphysical hypostatisation of a well justified methodological rule - the scientist’s decision never to abandon his search for laws. The metaphysical belief in causality seems thus more fertile [...] than any indeterminist metaphysics” (Popper, 1959). The conclusion must be that, as it will never be certain whether a theory is true, forecasts based on any theory are bound to be uncertain.

Ad 2. Distinction between endogenous and exogenous variables.

Endogenous and exogenous variables cannot be distinguished on strict logical grounds. In the real world there is no such distinction. The choice is based on theoretical or - more often - practical considerations. For example, in explaining changes in fertility on the basis of changes in female labour force participation, one may question whether changes in the labour market should be taken as exogenous, for changes in fertility may affect labour force participation of women. Moreover, immigration has an effect on labour supply and may thus indirectly affect female labour force participation. Particularly in the long run the distinction between endogenous and exogenous variables becomes less clear-cut. For example, changes in population size due to fertility and migration may affect economic growth (e.g. due to an increase of consumption) which will have an impact on the labour market which in its turn may affect fertility and migration. Thus one may question whether economic variables can be regarded as exogenous variables in forecasting demographic changes in the long run.

Ad 3. Unobserved variables.

Theories often contain concepts that cannot be observed directly. In specifying a quantitative model the theoretical concepts have to be replaced by observed variables. For example, various theories of mortality are based on the concept of ‘frailty’. As frailty of individuals is not observed, these theories can only be tested indirectly.

The problem of unobserved variables does not only apply to the explanatory variables. For example, in most countries there are hardly observations on migration by motive. Different types of migration can be explained by different variables. For example, labour migration depends on the labour market in the countries of origin and destination, while family reunification depends on the resident foreign population in the country of destination. Due to the lack of detailed migration data, it is difficult to assess the impact of explanatory variables.

As observed variables are only proxies of the theoretical concepts, it is uncertain whether empirical relationships really correspond with the theoretical relationships. Consequently it is uncertain whether empirical relationships are stable.

Ad 4 Functional form of relationships.

Usually the choice of the functional form of the relationships cannot be based on a theory. Generally theories can justify the choice of variables, but not the mathematical form of the model. Hence in many cases the choice of the functional form is based on practical considerations (e.g. assuming linearity for the sake of simplicity) or empirical tests. One problem in choosing an appropriate function is that there are usually several functions which give almost the same degree of fit (Klaassen and Pawlowski, 1981). Different functions, however, may lead to widely different forecasts, if one assumes that the values of the explanatory variables in the forecast period differ considerably from the values in the sample period.

Ad 5. Values of structural parameters.
The values of the structural parameters can be estimated on the basis of data in some observation period. In using an explanatory model for forecasting, the question is whether the estimated values will be valid during the forecast period also. Structural parameters may vary through time. If long time series of data are available, one may divide the sample period into sub-periods and test whether the parameter values vary between them. Assuming that the pattern of change of the parameter values is regular, one may extrapolate changes of the parameter values. In fact, this means that the number of parameters of the model is extended. It is uncertain whether the assumption about the pattern of change of the structural parameters is valid, particularly if there is no explanation of the change in the parameter values. One major cause of changes in structural parameters is that they may be affected by variables absent from the model used for forecasting. In the future omitted variables may develop in a different way than in the past, which may affect the social context and hence the structure of the model. For example, unpredicted political changes may lead to changes that affect demographic behaviour. International political developments can cause a change of the direction and size of migration flows. For example, policies on asylum seekers changed drastically in various Western European countries in the first half of the 1990s. Another obvious example is the collapse of the communist system in Eastern Europe around 1990 which had a huge impact on fertility, mortality, and migration in the 1990s. Thus future changes in parameter values may be rather uncertain.
Another source of changes in parameter values is heterogeneity. If the values of the structural coefficients differ between social groups, the relationship between the demographic variables to be forecast and the explanatory variables on the macro level is not equivalent to the corresponding micro relationship. The micro coefficients depend on the distribution of the explanatory variables on the micro level. If this distribution changes through time, the macro coefficients will change even though the micro coefficients may remain constant. Estimations of the macro model using least squares leads to aggregation bias. The expectation of the estimated coefficients does not equal the average of the corresponding micro coefficients (Theil, 1971).

Ad 6. Future values of exogenous variables.
When using an explanatory model for *ex ante* forecasting one has to make assumptions about the future values of the exogenous variables. Particularly for long-term forecasts this is an important source of uncertainty of forecasts. If changes in fertility, mortality, or migration follow changes in other variables with a certain time lag, forecasts of demographic variables can be based on observations rather than forecasts of explanatory variables. However, this is only useful for short-term forecasts. For example, on the basis of monthly data Macunovich and Easterlin (1988) find that unemployment leads realised fertility by 9-16 months. On the basis of Dutch data De Beer (1991a) concludes that indices of consumer confidence are a useful leading indicator for the total fertility rate with a lead time of about one and a half year.
Even if the observations rather than the actual development of demographic variables lags behind observations of certain non-demographic variables, the non-demographic variables may serve as leading indicators. For example, in the Netherlands the monthly number of deaths is related with temperature. If it is relatively cold or warm, significantly more people die. As information on temperature is available some months earlier than data on deaths, temperature can be used to project the number of deaths some months ahead. Since in using leading indicators, observations rather than forecasts of explanatory variables can be used for forecasting demographic variables, uncertainty of forecasts can be reduced. This is particularly useful if turning points can be forecasted on the basis of observed turning points in leading variables. However, the method can be used for the short run only. And even in the short run there is some uncertainty in using this method. If there have been few turning points in the past, it is rather uncertain whether a variable is useful as a leading indicator. Moreover, if the length of the lead time changes, a turning point may be forecasted for the wrong year.
For long-term forecasts exogenous variables need to be forecasted. Forecasts of the exogenous variables could be based on another model, possibly developed by other forecasters, but that other model would also include exogenous variables. So unless one has a closed model, i.e. a model in which all variables are endogenous, at the end of the day one has to formulate expectations about the future values of exogenous variables which cannot be based on an explanatory model. If the model contains two or more explanatory variables, an additional source of uncertainty is the question to what extent the explanatory variables are correlated. If the explanatory variables are positively correlated and if they affect the variable to be forecasted in the same direction, the uncertainty is greater than if the variables are not or negatively correlated or if they have opposite effects on the variable to be forecasted. The uncertainty of population forecasts based on expectations about the future development of non-demographic variables may be considerable. Keyfitz (1982) questions whether expectations about future developments of non-demographic variables can help in improving demographic forecasts: "however elusive future population may be, the future value of economic variables is even more so." One source of uncertainty in forecasting exogenous variables is the timing. Even if the direction of change is forecasted correctly, forecast accuracy may be poor if the timing of the tendencies is not forecasted correctly. "One can have a perfectly good theory with all the factors in place, but if their incidence can vary by a quarter century or so the theory is useless for forecasting to the horizons usually required" (Keyfitz, 1990b).
Migration does not depend on variables of one country only. Most theories are based on the differences of e.g. income and employment between countries. Hence forecasts of explanatory variables in various countries are needed. Another problem is that migration is affected by changes in government policy which may be very

difficult to forecast. Policy changes may, for example, depend on the outcomes of future elections. Not only the question which policies will be conducted is important, but also the effectiveness of policy should be assessed. It should be determined how strongly actual migration is affected by governmental policy. As a restrictive immigration policy may result in an increase of illegal or undocumented immigration, it makes an important difference whether the population forecast aims to describe the future *de jure* or the *de facto* population. Section 6 discusses the uncertainty of expectations about the future values of the exogenous variables.

Ad 7. Random errors.
Even if the explanatory model used for forecasting is correct and the future values of the explanatory variables are known with certainty, forecasts are still uncertain due to random errors. One source of disturbances is the impact of variables not included in the model. The degree of uncertainty of forecasts due to random errors depends on the validity of the assumptions about the statistical properties of the disturbances of the explanatory model. If the future values of the error term have zero mean and constant variance, and if they are serially independent (i.e. if the error is independent from one period to the next), the fit of the model to the sample data provides an indication of the degree of uncertainty of the forecasts due to random errors. If the explanatory variables explain a major part of the variance of the variable to be forecasted, random errors may only have a relatively minor effect on the uncertainty of the forecasts. If, however, the explanatory variables only explain a limited fraction of the variance, uncertainty of forecasts is large even if the explanatory model is valid. If errors are serially correlated, the uncertainty of forecasts may be large. First, autocorrelation of errors enlarges the risk of misspecification. If in estimating the explanatory model, autocorrelation is ignored, the danger of spurious correlation may be considerable, as the standard significance tests are misleading. The conventional *t* test will tend to reject the hypothesis when, in fact, there might be none. On the basis of Monte Carlo simulations Granger and Newbold (1974) suggest that one should use a critical *t* value of 11.2 rather than the normal value of 1.96 to assess the significance of an estimate of a coefficient. However, the general practice is still to consider coefficients with an estimated *t* value of 1.96 as 'significant'. Secondly, even if the occurrence of autocorrelation is taken into account, one source of uncertainty is that usually the form of the autocorrelation or the value of the involved parameters is unknown. Analogously, heteroscedasticity (i.e. the assumption that the variance of the disturbance term is not constant across observations) may have an adverse affect on the forecast variance. Whereas autocorrelation is a common problem of time-series analyses, heteroscedasticity usually arises in cross-section data where the explanatory power of the model tends to vary across observations.

In summary one can conclude that in using an explanatory model for forecasting fertility, mortality, and migration, there are two main sources of uncertainty.
1. Uncertainty about the structure and stability of the model. If a time-series macro model is used, it is uncertain whether the empirical relationships identified in the sample period will also be valid in the forecast period. Heterogeneity and omitted variables are two main causes of changes in the model structure and in the parameter values. If a cross-sectional micro model is used for forecasting, one main source of uncertainty is the question whether the specific economic and social context in the period of observation has a different impact than in the forecast period. For example, if an explanatory model of fertility is estimated on the basis of a survey during an economic boom, the results may differ from a survey that would have been held during an economic recession. Social trends affecting demographic behaviour of most individuals simultaneously may cause changes through time.
2. Uncertainty about the future values of the explanatory variables. In using a time-series explanatory model for forecasting, the accuracy of the forecast depends on the uncertainty of the future values of the exogenous variables. Thus the uncertainty of demographic forecasts depends on the uncertainty about the future economic and social context. If a micro model is used for forecasting, the uncertainty depends on the accuracy of forecasts of the future size of population categories. For example, if fertility depends on the level of educational attainment, uncertainty of forecasts of fertility depends on the accuracy of forecasts of the population by educational level.

5. Extrapolation

Because there are no scientific theories that yield unambiguous, straightforward demographic forecasts, population forecasts are in practice mainly based on extrapolations. Keyfitz (1982) argues that "observed regularities serve perfectly well for forecasting as long as they continue to hold." Willekens (1990) notes that extrapolation methods may accurately forecast demographic events in the short run: "The reason is inertia. Demographic behaviour is closely linked to attitudes and values that most people do not change easily. Most of the time the parameters of the demographic processes do not change rapidly." For the long run, however, extrapolations are only justified on the basis of an assumption about the underlying causal mechanism: "it is by being able to understand the causal processes that underlie trends that we are in a position to know how safely we can extrapolate a trend into the future" (Ryan, 1970). On the basis of a survey of empirical studies Smith (1997) concludes that there is no evidence that causal models produce more accurate forecasts than extrapolation methods.

A given time series can be extrapolated in quite different ways, dependent on the choice of the extrapolation method. Extrapolation methods can be distinguished by what they assume constant. Either the level or change in the demographic variables can be assumed to remain constant. Alternatively, the coefficients of some model can be assumed to be constant. The model can describe a deterministic or a stochastic trend. Another distinction is

between models projecting one demographic indicator (e.g. the total fertility rate or life expectancy at birth) and models projecting age-specific rates. Instead of choosing between projecting period and cohort variables one can construct a model describing the relationship between period and cohort changes. Finally, rather than projecting a time series on the basis of some model describing the past development of the time series, one may extrapolate a variable on the basis of a graph.
Thus six methods for extrapolating demographic time series can be distinguished:
1. constant rates, numbers or changes;
2. deterministic time-series models;
3. stochastic time-series models;
4. models for age-specific rates;
5. translation models;
6. graphical extrapolation;

Ad 1. Constant rates, numbers or changes.
In forecasting fertility the age-specific fertility rates may be kept constant. In forecasting mortality the rate of change of the age-specific mortality rates may be kept constant. In forecasting migration, net migration numbers may be kept constant. In some countries net migration is ignored in the population forecast. If it is not explicitly assumed that migration actually will be zero, the projections should not be presented as a forecast but as a scenario. Uncertainty on future migration should result in specifying variants rather than in avoiding a forecast. If zero migration is a political target, assuming zero migration can be regarded as a forecast only to the extent that the policy is expected to be effective. Assuming a constant level of future migration does not imply that migration actually is expected to be constant, as the assumptions may be aimed to describe the average size of migration rather than annual fluctuations during the forecast period. The constant value may be equal to the last observed value or to some other value, e.g. the mean value over some base period.
A distinction can be made between constant rates as forecasting method and projecting constant rates on the basis of a time-series model. In the former case constancy may be assumed irrespective of the time-series properties of the variable to be projected. In the latter case constancy can be an optimal forecast, given the past behaviour of the time series. For example, on the basis of Dutch data De Beer (1990) shows that the time-series model that is suitable to project net migration implies a constant future level. Another example is the projection of US fertility by Thompson et al.(1989). They employ a rather sophisticated time series model, but the optimal forecasts of the total fertility rate according to that model turn out to be about constant. The assumption of constant numbers may turn out to be a good forecast if births or migration move up and down without a rising or falling trend. In practice the forecasting performance of sophisticated models often turns out not to be superior to that of constant forecasts (Makridakis et al., 1982).
A constant level does not imply that there are no changes. For example, if the age pattern of fertility is changing, even though the total fertility rate may remain almost constant due to offsetting increases and decreases at different ages, assuming constant fertility may lead to large forecast errors. Or, another example, if total net migration is zero, net migration is not necessarily zero for all age groups. If the age patterns of immigration and emigration are not identical, net migration will be positive for some age groups and negative for other age groups.
The uncertainty of a forecast assuming constant levels or changes, depends on the question whether the value of the level or rate of change is determined correctly. If future rates or numbers are assumed to be equal to the present values, forecast errors may be large if the most recent observations are around a peak or trough level. If the future level is assumed to be equal to some other constant value, the forecast errors may be large if there is a rising or falling trend. If the future rate of change is assumed to be equal to the average value during recent years, forecast errors may be large if the base period is not chosen correctly.

Ad 2. Deterministic univariate time-series models.
Two broad categories of time-series models may be distinguished: deterministic and statistical or stochastic models. Deterministic models attempt to describe the global trend of a time series. Univariate time-series that are used for long run projections need to be capable of describing global trends. Thus deterministic models attempt to describe the complete history of a time series by some mathematical function. Trend curves can be described by various mathematical models. If a rise or fall to some constant level in the future is expected, a model describing a S-curve may be used, e.g. the logistic or the Gompertz model. If a cyclical pattern is expected a periodical function may be used. For example, according to the Easterlin hypothesis there is a cycle in births, caused by the assumed inverse relationship between the level of fertility of a birth cohort and its size (e.g. Ahlburg, 1983, 1986). Polynomials, which can be estimated very easily by means of linear regression against time, time squared, etc., can describe widely different patterns. Although polynomials of second or higher degree often fit rather well to observed time series, they are not very useful for forecasting as they generally yield extreme forecasts in the long run (De Beer, 1992). Alternatively, the long-run development can be described as a succession of periods with different linear changes by means of linear splines. The direction of the trend in the last years of the observation period can be used for extrapolation into the near future. De Beer (1989) applies this model to fertility. The linear spline model is easy to apply as it does not require much statistical expertise. The results are easy to interpret by users.
One problem in extrapolating trends for the long run is that it is difficult to assess whether the chosen trend curve is appropriate, as different functions may each fit reasonably well to historical time series while they produce widely different extrapolations (e.g. De Beer, 1992). Both the choice of a mathematical function and the estimates of the parameters depend heavily on the choice of the observation period.

There are three types of uncertainty when using a deterministic trend model for forecasting.
First, there is uncertainty due to random fluctuations. However, if the choice of the deterministic model is valid, these errors are only due to short run fluctuations which do not affect the long-run trend. In the long run these fluctuations should cancel out.
Second, the direction of the trend may change in the future, e.g. an increase is projected, whereas in reality there is a decrease or *vice versa.* If a time series has increased during a long period and if there are no reasons to assume that the underlying causes of this trend will cause the trend to be interrupted in the future, the forecast that the increase will continue may seem rather certain. However, if the observed time series is relatively short, it is rather uncertain whether the increase can be expected to continue in the long run. And even though the trend may have increased for a long time, it is not certain that the trend will continue endlessly.
Third, the direction of the trend (increase or decrease) may be forecasted correctly, but the projected size of the change may turn out to be too low or too high. Even though a variable has increased during a long period, the rate of increase may change. A small change of the rate of change may lead to great uncertainty in the long run.

Ad 3. Stochastic univariate time-series models.
The forecasts of stochastic (or statistical) time-series models are based on the estimation of a 'local' trend. Whereas in deterministic models random fluctuations are assumed to cause the series to deviate from the trend curve but not to affect the trend curve itself, in stochastic models the trend is affected by random fluctuations. Hence the trend does not necessarily follow the same direction through the years. This has important implications for forecasting, particularly in the long run. On the basis of a deterministic model, random fluctuations only have a temporary impact: the time series is assumed to move towards the deterministic trend. The use of a stochastic model, however, implies that random fluctuations may lead to permanent shocks. As a consequence the uncertainty of forecasts based on a stochastic model tends to be much larger than that based on a deterministic model. Obviously, this does not imply that the uncertainty is much larger *because of* the use of a stochastic model but rather that in using a deterministic model one *assumes* the uncertainty to be smaller because one assumes the trend to be stable.
A relatively simple method is exponential smoothing. The main idea is to adjust the old forecast by a fraction of the forecast error. The method is purely pragmatic. As it is not based on a theoretical statistical model there are no explicit statistical criteria for choosing between different models. Various trend specifications can be added to the model. Gardner (1985) gives an overview.
Since Box and Jenkins (1970) presented a procedure for identifying Autoregressive Integrated Moving Average (ARIMA) models, these models have become rather popular in economics and - although to a less extent - in demography. The ARIMA model projects a variable (or the change in a variable) on the basis of past values of that variable (or changes in the variable) and past values of an error term. The autoregressive part (i.e. the lagged variables) describes a gradually declining effect of disturbances on the future values of the series, the moving average part (i.e. the lagged error terms) describes a temporary effect. Identification is based on the pattern of the autocorrelation coefficients (i.e. the correlation between the value of the variable at time *t* and the value at time *t-i*). A gradually declining autocorrelation function suggests an autoregressive function, whereas a moving average model seems appropriate when the autocorrelations cut off after some interval. When the autocorrelations decrease very slowly, the series is likely to be non-stationary and therefore the series should be differenced. After a model is chosen and the parameters are estimated, the residuals are examined in order to establish whether the selected model is appropriate. If the residuals show some systematic pattern, the procedure is repeated. Some examples of applications of ARIMA-models to fertility are Lee (1974), Saboia (1977), McDonald (1979), De Beer (1992), and Bell (1992). For mortality ARIMA models are used by McNown and Rogers (1989) and Lee and Carter (1990). De Beer (1990) applies the model to net migration. For the identification of an appropriate ARIMA-model the time series needs to be of sufficient length. Box and Jenkins (1970) suggest at least 50 observations as a guideline, though this requirement obviously can depend on the particulars of the series involved (Bell, 1992). However, even though long time series may be available, it is questionable whether developments in the distant past can be used for identifying the model that is to be used for projecting the near future (Bell, 1992).
As the parameters of ARIMA-models are difficult to interpret, the method is to an important extent a black box for both the forecaster and the user. An alternative model of which the parameters are better interpretable is the structural time-series model. The projections of this model are based on separate estimates of the level and the trend of a time series (Harvey, 1984). De Beer (1988) applies this model to demographic data.
There are three sources of uncertainty in using a stochastic time-series model for forecasting.

- The use of a stochastic model implies that one expects random fluctuations to occur. Hence using a stochastic model one does not expect the projections exactly to come true. For example, using a random walk model, a constant level is projected. But that does not imply that the forecaster assumes that really no change will occur. Rather it is assumed that these changes are unpredictable (i.e. that there is no autocorrelation). If the assumption that the mean of the random fluctuations equals zero is correct, there is uncertainty about the future level in each separate year rather than about the average level in a number of successive years (however, the validity of this assumption is uncertain due to the next two sources of uncertainty).
- Parameter values may change. If parameters are estimated for some historic period, the values may not be valid for the future. Estimates of parameters depend on the choice of the estimation period. Although a long time series is preferable for accurate estimates, a long period implies that all kinds of societal changes may have taken place which may have affected the parameter values. If sufficient data are available it can be

tested whether parameter values have been constant in the past. But even if that is the case, it is uncertain whether that will be true in the future also.
- The model structure may change. In practice it is often difficult to distinguish stationary and non-stationary models, particularly if the length of the observed time series is limited. If a stationary model is chosen on the basis of an analysis of the observations during the sample period, whereas in the future the time series will turn out to be non-stationary, the uncertainty of the forecasts will be underestimated.

Ad 4. Models for age-specific rates.
Using the cohort-component model for population forecasting requires age-specific fertility, mortality and migration rates or numbers have to be projected. Two approaches for extrapolating age-specific figures can be distinguished:
a) separate extrapolations of the level and age pattern;
b) projection of age-specific rates.
Ad a. Indicators of the level and the age pattern of fertility, mortality and migration can be projected separately. For example, in projecting fertility the total fertility rate (TFR) or cohort completed fertility rate (CFR) can be projected. Subsequently some assumption on changes in the age pattern has to be made, e.g. by extrapolating the parameters of some model age schedule, such as the mean age at childbearing and the standard deviation. On the basis of these projections, the age-specific fertility rates can be calculated. One basic assumption underlying this approach is that the level and timing can be forecasted separately. On the basis of US fertility data Miller (1986) and Bell et al. (1988) conclude that the fertility age curve does not depend on the level of fertility. Thompson et al. (1989) fit gamma curves to age-specific fertility rates for each year in the period 1921-1984 and subsequently develop a multivariate time-series model for projecting the TFR and the 3 parameters of the gamma curve up to the year 2020. Alternatively, Bozik and Bell (1987) follow the principal components approach to provide a linear transformation of the data that approximates the fertility rates in each year. The time series of the TFR and the first four principal components were modelled multivariately, the remaining principal components were modelled separately.
A large number of mathematical functions have been fitted to observed age-specific fertility rates, e.g. the beta-distribution (Romaniuk, 1973; Mitra and Romaniuk, 1973), the gamma-distribution (Thompson et al., 1989), the Gompertz curve (Farid, 1973; Pitcher, 1978; Ross and Madhavan, 1981), the Hadwiger function (Hoem and Berge, 1975), spline functions (McNeil et al., 1977), the Coale-Trussell function (Coale and Trussell, 1974), the double exponential function (Rogers, 1986). For mortality the Gompertz curve, the Makeham curve (e.g. Keyfitz, 1990), the Heligman-Pollard curve (Heligman and Pollard, 1980) and the relational Brass method (Brass, 1974) have been applied. Rogers and Castro (1981) describe model age schedules for migration. If individual age-specific rates are projected separately, the resulting rates may not yield a plausible age pattern (De Beer, 1992). Projecting the parameters of a mathematical function results in smooth age patterns. Although one important benefit of fitting a mathematical function is that only a limited number of parameters has to be projected (e.g. Keyfitz, 1990), the other side of the picture is that if a model age schedule is used, only those changes in the age pattern that can be described by changes in the parameters can be projected, whereas it may be possible that the age pattern changes in such a way that a different function would be needed. Thus projecting the parameters of a fitted model age schedule does not necessarily lead to an improvement of forecast accuracy. One major cause of uncertainty in projecting age patterns is that usually model age schedules are fitted to period age-specific rates, as for cohorts complete age patterns are available for old cohorts only. However, the age pattern in a calendar year is the result of a cross-section of age patterns of successive cohorts. If the age patterns or intensities of successive cohorts are changing, the age pattern in a given calendar year may differ from that of all cohorts. Hence a projection of age schedules based on period data may result in a biased age pattern in the long run. It should be noted that one problem in projecting the TFR and life expectancy is that the levels of the TFR and life expectancy are affected by changes in the timing of fertility and mortality respectively. Thus it is not an optimal procedure first to project the TFR and life expectancy and then to project changes in the age pattern of fertility and mortality.
Ad b. The age-specific rates or numbers can be projected directly. One way of doing this is to project each of the separate age-specific fertility rates (see e.g. Passel, 1976). When the separate projections are taken together, they may yield an implausible pattern: “the long-term projections may show a distribution across age that does not make intuitive sense in terms of the fertility curve not having the same sort of smooth shape across age as historical data do.” (Bell, 1992). Alternatively, models can be used for projecting all age-specific fertility rates jointly. Some examples are the APC-ARIMA model suggested by Willekens and Baydar (1984) and the CARIMA-model proposed by De Beer (1985). The mist simple variant of the age-period-cohort (APC) model assumes that the level of the fertility or mortality rate for age x in year t can be explained by three main effects, viz. an age effect a_x, a period effect p_t and a cohort effect c_k. Using this model one can separate period and cohort effects on statistical grounds. Both types of effects can be projected separately. If the age effects are assumed to be constant, and if there are no interactions between the main effects, only two parameters have to be forecast. Under these assumptions Willekens and Baydar (1984) use the APC model for forecasting fertility. If the estimates of the cohort effects for recent cohorts are constant, the cohort effects have to be projected for future cohorts only. This would reduce forecast uncertainty considerably. However, if there are interactions between age, period and cohort it is necessary to project a large number of parameters. This makes forecasts rather uncertain. The occurrence of interaction between age and period effects and/or interactions between age and cohorts effects is likely if fertility rates at young ages decline for young cohorts due to the postponement of births (age-cohort interaction). The estimates of the cohort effects for young cohorts will tend to be too low, if the interaction effect is not included in the model, which will result in projecting too low ultimate fertility for those

cohorts. But as young cohorts are not ‘complete’ there is no objective way of assessing whether or not an interaction effect should be included in the model. Thus the estimates of cohort effects depend on the estimation period (De Beer, 1989; Vermunt, 1991b). As the estimation of cohort effects for young cohorts is uncertain, obviously the forecasts based on APC models are uncertain.
The CARIMA-model is aimed to simultaneously projecting a complete set of age-specific rates (De Beer, 1985). The CARIMA-model projects the difference between a fertility rate at a given age in a given year and the value at previous ages and previous years. An objection to this model is that the parameters are not easy to interpret. Furthermore as the projections of this model depend heavily on changes in the most recent observations, the model seems more appropriate for short-term projections than for long-term projections.

Ad 5. Translation model
Translation models attempt to describe the relationship between period and cohort fertility analytically (Ryder, 1964). On the basis of assumptions on the mathematical function describing the trend of the cohort level and timing of fertility, a model can be derived in which period total fertility is related to cohort total fertility and moments of the distribution of births over age. This model can be used for forecasting cohort total fertility on the basis of observed period fertility. The model disentangles the impact of changes in the cohort level and timing of fertility on the period level of fertility (Keilman, 1991). This can be useful for interpreting recent changes in period fertility. However, the results depend heavily on the assumption on the trend curve that describes changes in cohort fertility. For example, Frinking and De Roo (1979) assume that changes in cohort total fertility and mean marriage duration can be described by second or third degree polynomials. De Beer (1983) concludes that the model based on polynomials does not yield plausible forecasts and that the forecasting performance is poor. Moreover, if the trend curves for cohort fertility are known, these curves can be used for projecting cohort fertility in a straightforward way and there is no reason to use a translation model. Hence it is questionable whether the use of translation analysis can reduce uncertainty of fertility forecasts.

Ad 6. Graphical extrapolation.
Graphical extrapolation shows which future changes seem plausible in view of past developments. For example, the ultimate level of fertility of young cohorts can be forecasted on the basis of a graphical extrapolation of cumulative fertility by age. Graphical analyses can give an indication of the degree of uncertainty of forecasts. If the development of a variable shows a smooth, gradual pattern, forecasts may be more certain than if the variable follows a random pattern. However, this is not necessarily true. Long-term forecasts of a variable following a random movement around a constant level may be more certain than long-term forecasts of a variable following a smooth trend, as a moderate change in that trend may have a considerable effect on the long-term level. Graphical extrapolation is subjective. The same curve may be projected into the future in very different ways depending on the person who performs the extrapolation. Thus to an important extent graphical extrapolation can be considered to be judgmental.

In summary, it can be concluded that in using an extrapolation method for forecasting three sources of uncertainty can be distinguished.

1. Uncertainty about the time-series properties. One important choice is that between deterministic and stochastic models. The deterministic model is based on the assumption that the trend is fixed and that there are random fluctuations around the trend. Assuming a deterministic trend, uncertainty of forecasts is determined by the size of the random fluctuations - which tends to be small relative to the change of the trend in the long run - and uncertainty about the stability of the trend. When using a stochastic model a choice has to be made between a stationary and a non-stationary model. In a non-stationary stochastic model it is assumed that the trend itself is affected by random fluctuations. This choice has important implications for the uncertainty of forecasts: a non-stationary model tends to produce a wide forecast interval, particularly in the long run.
2. Uncertainty about the functional form. For long-run forecasts it makes a lot of difference whether the time series is modelled by e.g. a linear, exponential, logistic, or cyclical model. As the future values of the variable to be forecasted may be outside the range observed in the sample period, it is not possible to decide on the basis of objective, statistical criteria which model is to be preferred for forecasting. In many cases the choice implies an assumption about the range of the future development. The choice of the extrapolation method tends to depend greatly on the length of the base period. It makes a great deal of difference whether we choose a model on the basis of a time series over 50 years or one over 10 years. It is by no means obvious which period should be preferred as a basis for forecasting. If the average change of the time series in the last 10 years differs from the average over the last 50 years, the question is whether the direction of the trend has changed or whether the change is due to a temporary fluctuation around the long-term trend. Unfortunately the fit of a model to historical data is no good predictor of its forecasting performance, as the extensive empirical research of Makridakis et al. (1982) has shown. As the choice of the extrapolation method cannot be justified solely on the basis of a statistical analysis of the characteristics of the time series to be forecast, the choice has to be based on assumptions about the way trends are expected to continue. In practice, these assumptions usually remain implicit, suggesting that the choice of the extrapolation method is merely a technical question. This is misleading. The choice of the extrapolation method should be justified explicitly when presenting a forecast. Ryan (1970) argues: “We can rely on trends for predictive purposes only where we can explain how the trend is caused, and thus estimate its reliability. This explains why it is that we can safely rely on unexplained trends only over rather short periods - it is only over short periods that it is at all safe to assume that the causes of which we are ignorant have not altered in any significant way.”

3. Uncertainty about the values of the parameters. Values of parameters may change. One way of dealing with this problem is to specify a more general model, which includes assumptions about the way parameters may change through time. For example, the Kalman filter can be used for estimating a time-series model with time-varying coefficients. This requires the estimation of additional parameters. However, it is uncertain whether these additional parameters are stable. Moreover, the estimation of additional parameters is a problem if the time series is relatively short.

6. Expectation

The two preceding sections show that neither explanation nor extrapolation yields objective, unambiguous population forecasts. For one thing, the choice of the model explaining or describing past developments cannot be made on purely objective, e.g. statistical criteria. For another, application of a model requires expectations about the way parameters and variables may change. Thus forecasts of the future cannot be derived unambiguously from observations of the past. Judgment plays a decisive role in both the choice of the method and the way it is applied. “There can never be a population projection without personal judgment. Even models largely based on past time series are subject to a serious judgmental issue of whether to assume structural continuity or any alternative structure” (Lutz, Goldstein and Prinz, 1996).
No matter whether forecasts are based on an extrapolation of observed trends or on an explanation of underlying causal mechanisms, they require expectations about changes in the future, the validity of which cannot be proved on the basis of data. Obviously expectations of the forecasters themselves play an important role in making forecasts. But forecasters can also take expectations of other experts into account. Forecasters can consult other experts about their expectations about likely - or possible - future developments. Lutz (1995) describes a procedure for the use of expert opinions for making population forecasts.
Two types of expectations can be distinguished:
1. Expectations of people concerning their own future behaviour.
2. Expert opinions.

Ad 1. Intentions.
Human behaviour is not only determined by the past (experience) and present (actual context), but also by the view of individuals about their own future and the future of the socio-economic context (Willekens, 1990). Thus expectations should be taken into account in explaining and consequently in forecasting behaviour. One way of including expectations in forecasts is to ask a sample of people about their own future demographic behaviour; e.g. young women can be asked how many children they expect to have during their lifetime. As expectations are not always realised (Westoff and Ryder, 1977; Van de Giessen, 1992), expectations data cannot be used at face value. To the extent that the deviations between expectations and actual behaviour are systematic, adjustment procedures can be applied so that expectations data can be useful for forecasting purposes. Empirical results indicate that there is a much stronger disposition to change positive intentions downward than to change negative intentions upward (e.g. O'Connell and Rogers, 1983). If expectations data for the same birth cohorts in different years are available, adjustments of expectations for young cohorts can be based on experiences for older cohorts (De Beer, 1991c). Multivariate models can be used to explain changes in expectations and realisations simultaneously (Calhoun and De Beer, 1991). Such models may be useful for adjusting birth expectations. This implies that the uncertainty of forecasts based on expectations depends on the uncertainty about the validity of the model used for adjusting the expectations data. This type of uncertainty is similar to that of the methods discussed in the preceding two sections. Moreover, the realisation of expectations does not only depend on individual factors such as infecundity, but also on changes in the social and economic environment. Assuming that the respondents are not better capable of projecting societal changes than population forecasters, expectations data may improve forecast accuracy to a limited extent only (Keyfitz, 1982). Westoff (1978) points out that the surveys held in the United States in the 1960s did not give any indication of the sharp decline in the 1970s. On the basis of a survey in 1975 among 2000 women who had also been interviewed in 1970, Westoff and Ryder (1977) conclude that the overestimation of the fertility based on the expectations was as large as the change in the period total fertility rate: “the respondents made the understandable but frequently invalid assumption that the future would resemble the present - the same kind of forecasting error that demographers have often made.”

Ad 2. Expert opinion.
As we noted in section 4, even if we would know the ‘true’ theory of fertility, mortality, and migration (assuming there is one) than this would still not lead to certain forecasts. Demographic behaviour cannot be regarded as a closed system. Demographic behaviour depends on the societal context. Changes in fertility, mortality, and migration are influenced by economic, cultural, technological, institutional, environmental, and political developments. For example, fertility is affected by the economy (labour demand, wage rate), culture (emancipation, individualisation, secularisation), institutions (availability of child care, parental leave), technology (IVF), and policy (family allowances). Thus the uncertainty of future demographic developments depends on the question to what extent changes in these areas are predictable. This depends on the question whether the causes of these causal factors can be identified: what are the causes of economic, cultural and other developments? Ultimately this comes down to the question whether an ‘ultimate’ cause can be identified. Often social changes are attributed to ‘modernisation’, characterised by economic growth, scientific progress and individualism. One may question, however, whether this is a cause or only a label. Through history various

philosophers and historians have claimed to have identified the ‘driving forces’ behind economic, cultural and other developments. For example, Marx claimed that economic developments (the ‘productive forces’) have a predominant impact on cultural changes, whereas Weber claimed that cultural factors have a decisive impact on the economic development. Others emphasise the importance of technological innovations. For example, Van de Kaa (1988) claims that the introduction of the contraceptive pill is the major single cause of changes in fertility and marriage behaviour in Western countries since the 1960s. Obviously, even if one or several major causes can be identified, the question is how the future development of these causal factors can be forecast. One should take into account the occurrence of events that have not taken place in the past, e.g. a technological innovation or a political crisis. Thus at the end of the day a forecast is based on expectations about future developments that cannot be derived directly from a model based on observations of past developments.
Basically expectations can be based on four types of assumptions about future changes.

1. Assuming that history develops in a linear way (following a trend which may be labelled ‘modernisation’), one may assume that observed trends will continue.
2. One may assume that history follows a cyclical pattern: ‘history repeats itself’. For example, in the short term economic developments tend to show business cycles. But also in the long term the economic development may be characterised by a cyclical pattern. Kondratieff claimed that there is a long-term economic cycle (e.g. Van Duijn, 1983). Alternatively, in the very long run one may assume that a cyclical pattern is caused by the rise and fall of civilisations as described by e.g. Toynbee and Spengler (e.g. Huntington, 1996).
3. If it is assumed that history changes in new, unprecedented directions, forecasts are not to be based on an extrapolation of historic trends.
4. If it is assumed that there are analogies in the developments in different countries, forecasts for some countries may be based on developments that have already been observed in other countries. For example, assuming that North European countries are ‘forerunners’, since new types of demographic behaviour (e.g. cohabitation, births outside marriage) are spread earlier than in other European countries, developments in North Europe can be helpful for making forecasts for the other countries. In addition, differences in government policy in the field of child care, parental leave, child allowances, juridical differences between married and unmarried couples and immigration policy between countries may give an indication of effects of possible future changes in policy on demographic behaviour. The question is, of course, to what extent developments in different countries are similar. Demographic convergence can be expected on the basis of the consideration that demographic characteristics are primarily determined by economic developments and that economic characteristics are becoming similar across the industrial world (Coleman, 1997). Coleman argues that economic convergence can be assumed to arise from internationalisation of markets, international investments, EU policies, global media and telecommunication. However, as Coleman argues, even though there is economic convergence in Europe, it is uneven and shows periods of stagnation. Instead of explaining changes in fertility by economic factors, cultural explanations have received much attention. The concept of the second demographic transition is based on the assumption that shifts in values are similar across countries: ‘post-material’ values emphasising individualism are gaining ground at the expense of more conservative values emphasising duty. Coleman (1997) notes, however, that “it is difficult to see how such liberating forces would lead to convergence, unless all agreed to be liberated in the same direction.” Hofstede (1981) claims that cultural differences between countries are very stable through time. He distinguishes four dimensions of values (power distance, uncertainty avoidance, individualism, and masculanity) that determine inter-country cultural differences. Even though there are global trends that affect all countries (e.g. globalisation, increasing communication, technological developments, changes in the physical environment), the reactions of people to these global changes differ due to durable cultural differences. According to Hofstede there is only a convergence of superficial aspects of culture (e.g. consumption patterns, amusement), but not of the fundamental values. The main source of uncertainty of the use of analogies, is the question whether it is justified to assume that differences in the social context between countries will not cause significant differences between observed developments in one country and future developments in another.

As experts tend to have different opinions, the Delphi method may be used. After a first round in which experts give their forecasts, a second round follows in which the experts are told what the other forecasts were and they are asked whether they want to revise their own forecasts. The aim of this procedure is to reach consensus. As judgments may be biased in a predictable manner, judgmental forecasts should not be accepted at face value, but should be adjusted (Kahneman and Tversky, 1979). For example, there may be a tendency to overestimate the continuity of trends, resulting in extreme predictions. Keilman (1990) claims that in making population forecasts there is an ‘assumption drag’: “forecasters use assumptions already been refuted by the data.” As a consequence, there are overreactions of forecasters in revising forecasts. As a result successive forecasts tend to follow an alternating pattern. As yet there is only little empirical research of the impact of experts opinions on the accuracy of population forecasts. One problem is that expectations are often not stated explicitly in the publication of forecast results. As Ahlburg and Land (1992) state, “there is definitely a need to know more about how official agencies use expert judgment in the production of forecasts.” Alho (1992) finds that the use of experts worsened forecasts of US mortality. Collopy and Armstrong (1992), however, conclude that expert judgment can reduce forecast errors if the judgment is structured. Goodwin (1997) shows that forecast errors can be reduced by the use of a statistical model to adjust judgmental forecasts.

If other institutions make population forecasts for the same country, the assumptions underlying these forecasts can serve as starting point in the assumption making process. In addition, assumptions underlying forecasts for other countries with comparable demographic developments can be examined. If other institutions have produced reasonably accurate forecasts in the past or if they have more expertise in particular fields, taking into account the assumptions underlying their forecasts may help in improving forecast accuracy. Moreover, it may be easier to convince users of the validity of the forecast if the assumptions agree with the ideas of other forecasters. However, agreement of different forecasters does not imply that the forecast will come true.

Forecasts are not only based on expectations of the most likely future development, they may also be affected by expectations about possible reactions of users of forecasts to errors. Forecasters should take into account the possibility that forecasts can be self-denying. This may e.g. occur if forecasters assume that immigration will be high, whereas policy makers aim at reducing immigration. Policy makers may react to the forecast by sharpening immigration policies. If these policies are effective, the forecast will not become true. If the forecaster takes this effect into account and reduces his forecast of immigration, the policy maker may conclude that immigration policies do not need to be made more strict, as a consequence of which immigration may be higher than forecast. The strength of this dilemma depends on the effectiveness of policies. If policies are far from one hundred percent effective, the risk of forecasts being self-denying is small. Note that self-denying forecasts illustrate that uncertainty and forecast errors are not synonymous: uncertainty does not imply that a forecast is wrong, but that it is not certain whether the forecast is right or wrong; a self-denying forecast is certain to be wrong.

In conclusion we can say that even if quantitative models are used, opinions play an important part (e.g. Kahneman and Tversky, 1979; Lawrence and O'Connor, 1992). As McNees (1990) puts it, the man (or judgment) versus model dichotomy is a false one. In applying univariate time-series models judgment plays an important role in the operationalisation of the variables to be forecast, the selection of the model and the choice of the base period. If multivariate models are used, the choice of the explanatory variables as well as the assessment of the future values of the exogenous variables are based on judgment. Particularly in making long-term forecasts judgment plays an important role. It is certainly not obvious that regularities observed in the past are stable enough to be valid in the far future. Thus even if a forecaster employs an ‘objective’ forecasting method, the forecast is implicitly based on expectations. If the forecaster does not really expect that trends will continue in the very long term, an extrapolation should not be labelled a forecast but rather a projection. As noted above, however, the users tend to interpret projections as forecasts. Thus the forecaster should indicate explicitly whether or not he really expects that trends will continue. If not, the user would like to know what the forecaster does expect.
Since as yet there is not a single method outperforming all other methods, in preparing forecasts an eclectic approach can be followed. The forecasts can be based on a combination of demographic analysis, explanatory analysis, time-series models and consultation of experts. Judgment is required to combine the outcomes of the separate analyses into one forecast.
In order to allow the user of the forecast to judge the uncertainty of a forecast it is important that the reasoning behind the expectations is made explicit. As expectations are subjective, in the sense that they are not based on an explicit quantitative model, the validity can only be judged on the basis of an evaluation of the underlying arguments. Even though the validity of expectations cannot be proven, the forecaster should try to argue his expectations. It is not sufficient to state simply that a certain development is expected. This is particularly important as the expectations are decisive for the accuracy of the forecast. Ascher (1978) concludes that “the core assumptions underlying a forecast, which represent the forecaster’s basic outlook on the context within which the specified forecasted trend develops, are the major determinants of forecast accuracy (...) When the core assumptions are valid, the choice of the methodology is either secondary or obvious.” Thus a presentation of the method used for making the forecasts does not provide sufficient justification of the forecasts. Rather the fundamental assumptions underlying the way the methods are used should be stated explicitly.
As forecasts are largely based on judgment, the assessment of uncertainty of forecasts depends heavily on judgment. There is no unambiguous objective way of assessing the uncertainty of forecasts. The assessment of uncertainty is based on assumptions of which the validity is not certain. However, this does not imply that uncertainty of forecasts is purely subjective. By making explicit assumptions underlying the assessment of uncertainty, experts and users can make their own judgment on the uncertainty of forecasts. This makes uncertainty intersubjective.

7. Deterministic forecast variants of fertility, mortality, and migration

In most industrialised countries the official population forecasts include one medium, main, central or baseline variant. This variant usually aims to project either a continuation of observed trends in fertility, mortality, and migration, or the future developments that are considered most probable by the forecasters. Of course, these two criteria may coincide: forecasters may assume that a continuation of trends is the most probable future.
Uncertainty on future fertility, mortality, and migration can be taken into account by means of calculating the effect of alternative values of these three components on the future size and age structure of the population. Two approaches can be distinguished.
First, deterministic forecast variants can be calculated, i.e. variants that are calculated in a similar way as the medium variant, the only difference being that different values of fertility, mortality, and/or migration are used. Each combination of values for the three components results in one set of outcomes of population size and age structure. It is possible to give each variant the same importance, i.e. not to label one variant the most probable.

If, however, variants are calculated, users will regard the middle one as the central, most probable variant. That could be a reason for a forecaster to publish two (or four) variants. Many users do not appreciate that. If they prefer to use only one variant they need to either make their own choice or calculate some average. To let users make their own choice does not seem an optimal use of expertise. The forecasters should be better equipped to make that choice than the user. “This approach is analogous to a physician giving alternatives for treating a disease without guiding the patient toward what the physician considers advisable” (Daponte et al., 1997). One important disadvantage of calculating the arithmetic (or some other) average is that it is not clear how the result should be interpreted. If the variants can be interpreted as boundaries of a symmetric confidence interval, there is no reason why the forecaster should not publish the arithmetic average as the most probable variant. If not, the arithmetic average does not correspond with the most probable variant.
Secondly, probability distributions of future fertility, mortality, and migration can be determined. On the basis of these probability distributions, the probability distribution of future population size and structure can be assessed. Probability distributions of fertility, mortality, and migration will be discussed in the next section. This section focuses on deterministic variants.

Two types of deterministic variants can be distinguished:
1. scenarios of future fertility, mortality, and migration representing alternative futures;
2. low and high values of fertility, mortality, and migration, indicating the bandwidth of possible future values.

Keilman (1990) distinguishes both types of variants by claiming that scenarios are based on qualitatively different assumptions, while uncertainty variants are based on the same set of qualitative assumptions and differ only by the way the qualitative assumptions are translated into quantitative values. This interpretation seems too strict. It raises the question which criterion is to be used for labelling assumptions qualitatively different. For example, if forecasts are based on an explanatory model, variants can differ by the values of the exogenous variables (using the same model) or variants can be based on alternative models. The interpretation of Keilman suggests that only the latter type of variants should be labelled scenarios (Keilman is not specific about what is qualitatively different). It seems, however, perfectly justified to label the former type scenarios too, since there is no reason why variants based on different assumptions about non-demographic variables would not describe ‘alternative’ futures, even though the same explanatory model is used.

Ad 1. Scenarios.
Even though the term scenario is increasingly used for various types of forecast variants, here we restrict the use of this term to variants of fertility, mortality, and migration that are based on different sets of qualitative assumptions about the future development of the determinants of these components which provide internally consistent, comprehensive pictures of alternative futures. Scenarios can be based on quantitative explanatory models specifying the relationship between fertility, mortality, and migration and non-demographic explanatory variables. Alternatively scenarios can be based on a qualitative assessment of the impact of economic, social, cultural, and technological changes on future fertility, mortality, and migration. If scenarios are based on qualitative assumptions about social developments, additional assumptions are needed in order to translate the qualitative considerations into quantitative assumptions. Usually these additional assumptions remain implicit.
Scenarios can take into account two types of uncertainty:
1. uncertainty about the relationship between demographic and non-demographic variables;
2. uncertainty about the future development of the non-demographic variables.

Usually scenarios differ only by the values of the exogenous variables. This implies that only the second source of uncertainty is taken into account and that other sources of uncertainty that were discussed in section 4, e.g. uncertainty about the validity of the underlying theory, uncertainty about the functional form of the relationships, and uncertainty about the stability of the values of the parameters of the model are ignored. Thus only calculating the effect of alternative values of the explanatory variables may result in underestimating uncertainty.
Uncertainty about the validity of alternative explanatory theories can be taken into account by specifying alternative variants which are based on different theoretical approaches. For example, one fertility scenario may be based on the Easterlin hypothesis, assuming a negative relationship between cohort size and the level of fertility, resulting in long-term cycles in fertility, whereas another scenario may be based on the assumption that fertility is mainly determined by long-term trends such as modernisation which develop into one direction, resulting in a permanent low level of fertility. Similarly mortality scenarios can be based on two alternative theories (e.g. Nusselder, 1998). One scenario may be based on the limited-life-span theory which assumes that life expectancy will not increase beyond a certain age, e.g. 85 years (e.g. Olshansky and Carnes, 1994). Further reductions in mortality rates are constrained by biological limits. Life expectancy can increase by a rectangularisation of the survival curve. But as human life span is fixed, life expectancy will reach a maximum. At this point people will die from ‘natural death’, i.e. death without disease (Fries, 1980). Rectangularisation implies that increasingly larger proportional declines of mortality rates are required to produce the same increase of life expectancy. A life expectancy at birth beyond 85 years would require enormous reductions in mortality rates. Another scenario can be based on the mortality-reduction theory which states that the decline of mortality rates will continue and that life expectancy at birth may increase beyond 100 years (e.g. Manton et al., 1991). Vaupel and Lundstrom (1994) argue that mortality rates at advanced age are still decreasing by 1 to 2 per cent per year, even in countries where mortality rates are already very low. In addition they argue that sub-populations with healthy life styles reach very low mortality rates.

Ad 2. Low and high variants.
The choice of low and high values of future fertility, mortality and migration can be based on three approaches:
a) identification of sources of forecast errors;
b) extreme variants;
c) arbitrary values.

Ad a. Identification of sources of forecast errors.
On the basis of an analysis of sources of uncertainty one may attempt to assess the bandwidth of possible future values of fertility, mortality, and migration. Sections 3-6 discussed sources of uncertainty of forecasts based on disaggregation, explanation, extrapolation, and expectation. On the basis of an identification of sources of forecast errors, the width of the interval between low and high variants can be assessed.
The impact of the uncertainty due to changes in the heterogeneity of the population can be taken into account by specifying variants that differ by the assumption about changes in heterogeneity. One variant may assume an increase of heterogeneity in demographic behaviour, whereas an alternative variant may assume heterogeneity to decrease. For example, one variant may assume that the impact of differences in the level of fertility between ethnic groups and the native population on aggregate fertility increases, because the relative size of ethnic groups increases or because the differences in fertility between the ethnic groups and the native population increase. Another variant may assume that the differences in fertility will disappear, because the fertility of the ethnic groups moves into the direction of the level of fertility of the native population. Alternatively one may assume that the degree of heterogeneity does not differ between the forecast variants, but rather that one variant is based on assuming high levels of fertility for all ethnic groups, including the native population, while another variant is based on low levels. When calculating the width of the interval between the forecast variants at the aggregate level by adding up the intervals for the separate population categories, it should be noted that the high and low variants at the aggregate level are less likely than the high and low variants for the separate population categories, unless one assumes perfect correlation between all population categories. We will come back to this.
Uncertainty about changes in the societal context, more specifically uncertainty about the future values of exogenous variables and the impact of omitted variables, can be taken into account by specifying variants that differ by the rate or direction of change of the economic, social, cultural, political, and technological context. For example, in assessing variants of fertility one variant may be based on the assumption that the process of individualisation will continue, as a consequence of which fertility will remain low, whereas another variant may assume that in the foreseeable future a saturation level will be reached after which the trend may change into the opposite direction, which may lead to an increase of fertility. Similarly for mortality one variant may assume a fundamental change in medical technology which leads to a considerable decline of mortality, whereas another variant may assume that technological progress will only have a limited impact on future longevity. For migration one variant may assume that the increasing migration pressure from Third World countries will lead to an increasing trend of migration to Europe, whereas another variant may assume that restrictive immigration policies will be effective and lead to a decrease in immigration.
Uncertainty about the stability of observed trends can be taken into account by specifying variants in which the trend changes in different directions. For example, one variant may assume that the future rate of increase of life expectancy at birth will be higher than in the past, whereas an alternative variant may assume the rate of increase to decline or even the increase to turn into a decline. Trend extrapolations are often based on deterministic time-series models. Such models assume a linear, exponential or other trend, around which random fluctuations occur. These models are deterministic, since the trend itself is not affected by random fluctuations. Analytically deriving confidence intervals from such models would result in underestimating uncertainty as only random fluctuations and the estimation error of the slope of the trend are taken into account, but not the possibility that the future trend may be different from that in the past. However, it is possible to use these models to generate forecasts of observed time series in order to determine empirical forecast intervals (Williams and Goodman, 1971).

Ad b. Extreme variants.
In assessing the width between low and high variants the question is how extreme the variants should be. When assessing the uncertainty of extrapolating an observed trend, the question is whether one should take into account differences in the rate of change only or whether one should allow the trend to change in a different direction than in the sample period. For example, in projecting life expectancy at birth the question is whether the variants should only differ by the rate of increase or whether one variant should assume that future life expectancy will continue to increase whereas another variant assumes a decrease. By the same token, when assessing the impact of uncertainty about societal changes the question is whether one should take into account some variation around a general trend like modernisation only or whether one should take into account the possibility of cultural changes into a different direction, e.g. a trend towards more family-oriented values. In addition, the question is whether in each variant one should assume that all societal developments will affect demographic behaviour in the same direction or whether one should take opposing effects into account. For example, in assessing forecast variants of life expectancy, in one variant one may assume that healthy life styles, medical progress, and a favourable economic development each have a positive effect on life expectancy, whereas another variant is based on assuming negative effects. It is, however, also possible to assume that in each variant economic growth and life styles have opposite effects on mortality. The former assumption will result in a wider bandwidth than the latter. Obviously the probability that the interval between extreme variants will cover the true value is larger than the probability for a narrow interval. However, if the range is large, the forecast is uninformative. Thus there is a trade-off between the precision of the forecast and its probability. The question

how wide the interval between forecast variants should be depends on the question how probable the interval between the variants should be, i.e. how likely it should be that the interval will cover the true future development. The probability of the interval between variants that differ only by the rate of increase of life expectancy is smaller than that of the interval between one variant that assumes a rapid increase and another variant that assumes a decline of life expectancy. Thus an assessment of the width of the interval between variants requires that some indication of the probability be specified. The assessment of the probability can be based on an analysis of errors of past forecasts, on a time-series model or on expert judgment. These methods will be the subject of the section 8.

Ad c. Arbitrary values.
Forecast variants can be based on arbitrary values of fertility, mortality, and migration. For example, often one variant is calculated in which the total fertility rate is assumed to reach replacement level. In some cases this is the medium variant, in other cases the high variant. Usually there is no explicit argument for this choice. Another common variant is zero net migration. This may be the medium variant or the low variant. A third possibility is to assume constant fertility and mortality rates and/or constant net migration. Usually these values are set equal to the current value. A comparison of the population forecasts based on constant rates with the medium variant gives an indication of the effect of the changes assumed in the medium variant. Hence these variants can be useful for a sensitivity analysis, showing the possible effects of future changes in fertility, mortality, and migration in comparison with the effect of the current age structure.

Summarising we can conclude that the uncertainty of forecasts of fertility, mortality, and migration is usually indicated by specifying alternative variants in addition to the medium or most probable variant. The width of the interval between the variants provides an indication of the degree of uncertainty of the forecast. The variants can be assessed on the basis of an identification of sources of forecast errors. If scenarios are based on assumptions about future economic and cultural developments, the width of the interval between the demographic scenarios depends on the degree of uncertainty of the assumptions about the non-demographic explanatory variables and the degree of uncertainty about the relationships between the demographic and non-demographic variables. If low and high variants are specified in order to indicate the bandwidth of possible developments, these variants can be based on an identification of possible different directions of trends. The probability that the interval between the variants will cover the true value is an important criterion for deciding how wide the interval should be. The assessment of the probability is the subject of the next section. Sometimes forecast variants are based on arbitrary values of fertility, mortality, and migration, e.g. constant rates or numbers. If, however, the choice of these values is not based on explicit arguments it is not clear how these variants can give an indication of the degree of uncertainty of the forecasts.

8. Probabilistic forecasts of fertility, mortality, and migration

In the preceding section it was argued that an assessment of the bandwidth between low and high variants of fertility, mortality, and migration depends on the question what the corresponding probability should be. It makes a lot of difference whether the bandwidth between the low and high variants should correspond with a chance of 50 percent or 99 percent. Similar considerations are also relevant for scenarios. In making scenarios the forecaster has to decide whether the scenarios should describe extreme situations, e.g. based on widely different sets of assumptions about the economic, social, and cultural context or whether the scenarios should reflect only gradual rather than fundamental different views about the future. Obviously if extreme scenarios are specified the probability that the area between the scenarios will cover the true development is much bigger than if the scenarios are close to each other. Thus in specifying scenarios the forecaster makes a - usually implicit - decision on the (order of magnitude of the) probability of the scenarios. In order to inform the user how to interpret the scenarios, an explicit statement about the intended level of probability is preferable to ignoring the issue as is usually done in presenting scenarios. “It is important for users to know the probability of a scenario. For example, a high-fertility scenario with a 20 per cent probability should be taken more serious in policy considerations than one with only 2 per cent probability” (Lutz, Goldstein and Prinz, 1996).

There are other reasons why an indication of the probability of forecast variants is useful. One is that it provides a criterion for achieving consistency between the forecast intervals for fertility, mortality, and migration, in the sense that the differences between the separate intervals should reflect differences in uncertainty between these three components. Forecasts of fertility, mortality, and migration are expressed in different measures: the ultimate number of children per woman, life expectancy at birth, and numbers of immigrants and emigrants, respectively. Thus the intervals between the variants for the three components cannot be compared directly. Probability provides a criterion for comparing the forecast intervals between these indicators, e.g. by requiring that the interval for each component should correspond with a given probability.
Another reason for taking into account the probability of forecast intervals is that it allows to take into account the effect of aggregation on uncertainty. Adding up forecast intervals with a given probability does not result in an aggregate forecast interval with the same probability, unless there is perfect positive correlation. For example, if forecast variants of fertility for separate population categories are added up, the resulting variants of fertility at the aggregate level are less likely than the variants for the separate categories, the reason being that it is less likely that fertility of all categories will be high simultaneously than that fertility of some categories will be high while that of others will be below its highest level. In other words, if the forecast intervals for the separate population categories are assumed to correspond with a certain probability, adding up the intervals results in overestimating

uncertainty at the aggregate level, if one ignores the fact that the probability at the aggregate level is higher than that of the separate intervals.
An additional reason for specifying the probability of a forecast is that it creates greater flexibility in reflecting uncertainty. Deterministic models are based on a choice by the forecaster about whether certain developments will take place or not. They either include or exclude certain parameter values. “Using a probability distribution permits one to express states of knowledge in between these two alternatives” (Daponte et al., 1997).
The probabilistic approach enhances communication among demographers. “Making one’s beliefs explicit using probability distributions allows other demographers to observe exactly how one views the sources of uncertainty. Others can then know on what they agree or disagree” (Daponte et al., 1997).
A final reason for assessing the probability of a forecast interval is that the user of a forecast needs an explicit statement of the probability of a forecast in order to be able to assess the risks of a decision based on the forecast. For example, if the user has to make a decision about an investment project and its financial results depend heavily on a population forecast, he will weigh the possible loss of a decision based on a wrong forecast with the probability of the forecast being wrong. He should know whether he has to take the possibility into account that future population size will be outside the forecast interval or whether this is highly improbable so that he can safely ignore it.

In assessing the bandwidth between variants or scenarios for each component one has to take into account that the way uncertainty increases with the forecast lead time differs between components. In the short run migration tends to be the most uncertain component, as migration usually shows very large fluctuations in successive years. However, the degree of uncertainty only increases to a limited extent with the forecast lead-time as increases tend to be followed by decreases, and vice versa. For fertility the uncertainty in the short run is smaller than that for migration, as the number of births is much less strongly changing between one year and the next. However, uncertainty strongly increases with the forecast horizon due to the uncertainty on the size of long-run changes in the trend. Finally, the uncertainty about mortality is much smaller than that about fertility or migration, as mortality rates tend to change only gradually. If for each forecast year the bandwidth between the variants or scenarios is to correspond with a given probability, the increase of the bandwidth with the forecast lead-time reflects the increase of uncertainty.
While uncertainty of forecasts is positively related with the length of the forecast period, the opposite may be true for the length of the unit forecast interval. Due to annual fluctuations, errors of annual numbers may be larger than average errors over longer periods. For example, net migration tends to fluctuate strongly. Consequently a forecast of the size of net migration in one specific year is rather uncertain, even in the short term. But since periods of increase tend to be followed by periods of decrease, the average size of net migration over a period of five or ten years tends to fluctuate less strongly than annual numbers. As a result a forecast of average net migration over the next ten years may be less uncertain than a forecast of net migration for the next separate years.

Since the uncertainty of forecasts increases with the length of the forecast period, one question is whether beyond a certain point in time we are too unsure to specify values of fertility, mortality, and migration. One reason is that the state of the world may change strongly in the long run, which may cause drastic changes in demographic behaviour. Another reason is that the impact of the current age structure on future population growth decreases in the long term. For example, the current age structure has an important effect on the development of the number of births in the next 20 to 30 years, because most future parents are already born. But beyond that time children will be born to parents who are not yet born at the moment the forecast is made. Consequently the uncertainty of forecasts of births is large beyond some 30 years. Forecasts of the number of deaths beyond 30 years in the future are less uncertain than forecasts of births, as most people dying then are already part of the current resident population.
When assessing deterministic forecast variants, the forecaster has to decide beyond which time horizon, he thinks it is not warranted to specify assumptions about probable values of fertility, mortality, and migration. Calculations beyond that time, e.g. assuming constant rates of fertility, mortality, and migration, may be useful, as they show the consequences of the current age structure for population growth in the long run. But these calculations are illustrations rather than forecasts of the most probable future. If the values of fertility, mortality, and migration in the extended period are kept constant in all variants, the bandwidth between the variants does not indicate the degree of uncertainty. This may be misleading for the users. Another cause of confusion may be that the time beyond which forecasts are very unsure differs between fertility, mortality, and migration. Thus the period for which values are kept constant may be different for the three components. This may be confusing for the users, because it implies that for the population size and age structure there is no clear-cut divide between the period for which the results can be interpreted as forecasts of the likely future and the period for which the figures are only the result of calculations based on an arbitrary assumption, viz. constant values of fertility, mortality, and migration. If forecast intervals are based on an assessment of the probability, the increase of the uncertainty in the long term is reflected in the increase of the forecast interval. If the forecasts for the very long term are very uncertain, the interval should be very wide. Obviously in the very long run the interval may become so wide, that it is not informative for the user. However, it is up to the user to decide at which point the forecast interval is not informative for his purpose.

The ‘true’ probability of a forecast is not known. The assessment of the probability of a forecast is based on assumptions of which the validity is uncertain (Lutz, Sanderson and Scherbov, 1996). Thus the probability of a forecast is a forecast itself. Therefore it is important that the forecaster not only gives an indication of the probability of his forecast, but that he gives an explicit account of the assumptions underlying the assessment of the probability. The user needs this information in order to be able to assess whether the forecaster is optimistic or pessimistic. Armstrong (1985) states that experts tend to be too optimistic about the accuracy of their own forecasts. Keilman (1990) suggests that the optimism or pessimism of the forecaster may depend on the size of errors of his previous forecasts.

The probability of a forecast interval can be assessed in three ways:
1. An analysis of errors of past forecasts.
2. Model-based estimate of *ex ante* forecast errors.
3. Expert judgment.

Ad 1. An analysis of errors of past forecasts.
The probability of a forecast interval can be assessed on the basis of a comparison with the errors of forecasts published in the past. On the assumption that the errors are approximately normally distributed – or can be modelled by some other distribution - and that the future distribution of the errors is the same as the past distribution, these errors can be used to calculate the probability of forecast intervals of new forecasts. Keilman (1990) examines the errors of forecasts of fertility, mortality, and migration of Dutch population forecasts published between 1950 and 1980. He finds considerable differences between the errors of the three components. For example, errors in birth rates are considerably larger than errors in life expectancy. Furthermore, Keilman examines how strongly forecast errors increase with the length of the forecast period, to what extent errors vary between periods and whether errors of recent forecasts are smaller than those of older forecasts, taking into account the effect of differences in the length of the forecast period.
One problem in comparing old and new forecasts is that they refer to different periods. Thus a simple comparison of forecast errors of old and new forecasts does not tell whether or not forecast accuracy has increased. For that reason Keilman (1990) uses an APC (Age Period Cohort)-type model for estimating the separate effects of the length of the forecast period (the 'age effect'), the - social and demographic conditions during the - period to be forecasted (the 'period effect'), and the jump-off year (the 'cohort-effect'). One problem in assessing these separate effects is that interactions may play an important role and that it is not obvious how these interactions should be implemented in the model. Keilman's model recognises that some periods are more 'difficult' to forecast than others. For that reason his model includes a period effect. This effect estimates the differences in the size of forecast errors between different periods. The question what is 'difficult' to forecast, however, is not independent of the type of forecast. For example, developments in a period in which fertility rates hardly change would be forecast accurately by a forecast based on constant levels of fertility rates, but not by a forecast based on assuming constant changes in fertility rates (unless fertility rates did not change in the period before the forecast was made because that would result in forecasting a constant level). Thus a period in which rates do not change much is not necessarily more 'easy' to forecast than a period with changing rates. If in the period before the forecast was made, rates changed in one direction, a subsequent period with rates changing in the same direction would have been more 'easy' to forecast than a period with constant rates.
In order to compare the quality of forecasts made at different times it is not sufficient to examine the size of forecast errors. A forecast may be right for the wrong reason. The reasoning behind the assumptions underlying the forecast should be taken into account. If a forecast based on constant rates turns out to be rather accurate because the rates did not change much in the forecast period, whereas in the period before the forecast was made the rates changed strongly, it is important to examine whether the accuracy of the forecasts was due to luck (because simply constant rates were assumed irrespective of the past development) or whether explicit arguments were given for assuming the change in the rates to come to an end and whether these arguments were correct (if the arguments were not correct, the forecast accuracy may have been a matter of coincidence).
One problem in using errors of published forecasts, is that it is difficult to assess the causes of forecast errors, because the publications of forecasts usually do not explicitly report whether the forecasts are primarily based on extrapolation, explanation or judgment. Thus it may not be clear whether errors were caused by a wrong estimate of the trend or by an unexpected change of the trend, e.g. caused by unexpected changes in exogenous variables, or by unexpected events, e.g. a change of policy which was not taken into account in making the forecast. As a result it is difficult to assess to what extent the errors of past forecasts are a good indicator of future uncertainty of new forecasts.
One additional problem is that the sample of ex post forecasts tends to be biased towards the older ones, as for recent forecasts the accuracy cannot yet be checked except for the short run (Lutz, Goldstein and Prinz, 1996). Forecast errors for the very long run result from forecasts made a long time ago. Both the variability of fertility, mortality, and migration and the methodology of the forecast may have changed. Hence it is questionable whether the assumption that the future distribution of errors is the same as the past distribution is valid. For example, the sharp decline of fertility after 1965 in many European countries caused large forecast errors. In assessing the uncertainty of new forecasts the question is how likely it is that large errors such as those of the 1960s may occur again in the future, for after 1975 fluctuations in fertility have been much smaller. Moreover, errors of successive forecasts may not be independent. For example in the 1960s and early 1970s in the Netherlands forecasts reacted only gradually to the decline in fertility. Consequently fertility forecasts were too high in successive forecasts. The correlation of errors of successive forecasts has an adverse effect on the size of *ex ante* errors of new forecasts on the basis of errors in historic forecasts.
Keilman (1990) notes that the sophistication of forecasting methods is likely to have changed over time. For example, if a new forecast is based on the cohort perspective while old forecasts were based on a projections of period rates or if a new forecast is based on explanation, whereas previous forecasts were based on

extrapolation, one may question whether the past errors are a good predictor of future errors. As noted above, it is difficult to assess whether recent forecasts are more accurate than older forecasts, because the forecasts refer to different periods. Keilman (1990) gives an overview of methodological changes in the national Dutch population forecasts that were published by Statistics Netherlands between 1950 and 1980. In the 1950s and 1960s forecasts of fertility were based on age-specific period (marital) fertility rates. In the 1970s the cohort approach was introduced. From the population forecasts of 1975 onward the quantitative assumptions were based on qualitative assumptions about economic, social and cultural developments. Similarly since the population forecasts of 1980 the forecasts of mortality have been based on an identification of the main determinants of changes in mortality (such as medicine, hygiene, nutrition, socio-economic situation, cultural trends) and assumptions about the future effects of these determinants on mortality. International migration was not included in the medium variant until the population forecasts of 1980. This implies that implicitly zero net migration was assumed. In the 1950s and 1960s some additional calculations of the effect of migration were done which were presented separately from the forecast. The 1975-based forecasts did not include a medium variant. Two variants were published, only one of which included an assumption about migration. From an analysis of the forecast errors of Dutch population forecasts published between 1950 and 1980 Keilman (1990) concludes that there was some improvement of the accuracy of fertility forecasts around 1975. It is, however, difficult to conclude whether this is a ‘real’ improvement of the forecasts, because, as noted above, the variability of fertility rates has decreased strongly since 1975 and it is questionable whether the ‘period effect’ in Keilman’s model is capable of ‘correcting’ the estimate of the size of errors of different forecasts sufficiently for this effect. For mortality Keilman did not find a monotonical decrease of forecast errors. In general, one may conclude that it is difficult to draw a clear-cut conclusion from a comparison of the errors of different historic forecasts. Consequently it is rather uncertain which conclusions can be drawn for the uncertainty of new forecasts.

Ad 2. Model-based estimate of *ex ante* forecast errors.
Instead of assuming that future forecast errors will be similar to errors of past forecasts, one may attempt to estimate the size of future forecast errors on the basis of the assumptions underlying the methods used in making new forecasts. If the forecasts are based on an extrapolation of observed trends, *ex ante* forecast uncertainty can be assessed on the basis of the time-series model used for producing the extrapolations. If the forecasts are based on a stochastic time-series model, the model produces not only the forecast, but also the entire probability distribution. For example, ARIMA-models are stochastic univariate time-series models that can be used for calculating the probability distribution of a forecast (Box and Jenkins, 1970). Alternatively, structural time series models can be used for this purpose (Harvey, 1989). The latter model is based on a Bayesian approach: the probability distribution may change as new observations become available. The Kalman filter is used for updating the estimates of the parameters. Similarly multivariate time-series models, such as transfer function models and VAR (Vector Autoregression) models can be used (Granger and Newbold, 1977; Sims, 1980). One problem in using stochastic models for assessing the probability of a forecast is that the probability depends on the assumption that the model is correct. Obviously the validity of this assumption is uncertain, particularly in the long run. Alternatively the probability of *ex ante* forecasts can be estimated by calculating the errors that would have resulted from applying the model in the past. For example, one can calculate the forecast errors that would have resulted from using the model *n* years ago to extrapolate the subsequent *n* years (obviously for assessing the size of *ex ante* errors, the data for the last *n* years should not be used for the identification of the model and the estimation of its parameters). This procedure can also be used to assess the probability of forecasts of a deterministic trend model. Moreover, this procedure might be applied to an explanatory model, although in that case use of the observed values of the exogenous variables would lead to underestimating the size of *ex ante* errors. A similar approach may be followed to estimate the uncertainty due to heterogeneity. For example, one can calculate the errors that would have resulted if *n* years ago one would have assumed that differences in demographic rates between population categories would not change or that the changes would be equal to that in preceding years. This provides an estimate of the uncertainty due to changes in the size of population categories. In addition one can calculate the errors that would have resulted if *n* years ago one would have assumed that the relative size of population categories would not change or that the rate of change would be constant. This would provide an estimate of the uncertainty due to changes in the demographic rates of the separate population categories. On the basis of these errors the probability of new forecasts can be estimated.

Ad 3. Expert judgment.
In assessing the probability of forecast intervals on the basis of either an analysis of the errors of past forecasts or an estimate of the size of *ex ante* errors, it is assumed that the future will be like the past. Instead, the probability of forecasts can be assessed on the basis of experts’ opinions about the possibility of events that have not yet occurred. For example, the uncertainty of long-term forecasts of mortality depends on the probability of technological breakthroughs that may have a substantial impact on survival rates. The uncertainty of long-term forecasts of migration depends on the probability of major shifts in international economic relations, e.g. the possibility of an economic boom in Africa. Even though these developments may not be assumed to occur in the medium variant, an assessment of the probability of such events is needed to determine the uncertainty of the forecast. More generally, an assessment of *ex ante* uncertainty requires assumptions about the probability that the future will be different from the past. If a forecast is based on an extrapolation of past trends, the assessment of the probability of structural changes which may cause a reversal of trends cannot be derived directly from an analysis of historical data and therefore requires judgment of the forecaster. By the same token, if a forecast is based on an explanatory model the assessment of the probability that the exogenous variables will develop in

another direction than in the past and the probability that the values of the structural parameters will change is based on judgment. Obviously, if a forecast is based on qualitative considerations, e.g. expectations about societal changes, the probability of alternative developments is to be assessed on the basis of judgment. Lutz, Goldstein and Prinz (1996) argue that subjective distributions are to be preferred to a time-series approach, because “structural changes and unexpected events are likely to happen.” Lutz, Sanderson, Scherbov and Goujon (1996) assess the probability of forecasts on the basis of opinions of a group of experts. The experts are asked to indicate the upper and lower boundaries of 90 percent confidence intervals for the total fertility rate, life expectancy, and net migration up to the year 2030. Subjective probability distributions of a number of experts are combined in order to diminish the danger of individual bias.
When the probability of forecasts is assessed on the basis of expert judgment it is important to take into account possible sources of bias. On the basis of an analysis of Dutch population forecasts Keilman (1990) concludes that after a steady movement of trends in a consistent direction, the confidence of forecasters increases. For example, in the 1980-based Dutch population forecasts the interval between the low and high variants of fertility was rather narrow. Keilman concludes that “this reflects the optimism among the forecasters about the predictability after a period of stable trends.” By the same token the relatively wide interval between the low and high variants in the 1975-based forecasts can be considered as a reaction to the sharp decline of fertility rates in the preceding years.

In conclusion it can be said that the assessment of the probability of forecasts is a forecast itself. There is no single method for assessing the probability of forecasts that outperforms all other methods. The probability can be forecasted on the basis of an analysis of (ex post) historic errors or the size of *ex ante* errors can be estimated on the basis of a time-series model. Alternatively, the probability can be assessed on the basis of expert judgment. The use of one method does not need to exclude the use of another method, as these methods may complement each other. For example, an analysis of errors of past forecasts may be more useful for short- and medium-term forecasts than for long-term forecasts, since there are only few data on errors for the long run and, moreover, these errors result from forecasts that were made some decades ago on the basis of considerably less methodological and substantial knowledge than is available at present. One way of trying to overcome this problem is to extrapolate forecast errors by means of a time-series model. If the increase of forecast errors with the length of the forecast period follows a regular pattern, the size of forecast errors for the long run can be projected on the basis of forecast errors of recent forecasts for the short run. Thus estimates of *ex ante* forecast errors can be based on an extrapolation of ex post errors. Alternatively the use of past forecast errors for assessing the probability of short-term forecast intervals can be complemented by expert judgment for assessing the uncertainty of long-term forecasts.

9. Uncertainty of population size and age structure

The uncertainty of future population size and age structure depends on the uncertainty of future fertility, mortality, and migration. The uncertainty of the latter three components can be made explicit by specifying forecast intervals. The width of the separate intervals is an indication of the degree of uncertainty of each component. The upper and lower bounds of the forecast intervals of fertility, mortality, and migration can be assessed on the basis of scenarios which specify how fertility, mortality, and migration are related with the economic and social context or on the basis of low and high values, indicating the bandwidth of possible developments. The forecast intervals of fertility, mortality, and migration can be used to assess the degree of uncertainty of the future size and age structure of the population on the basis of a deterministic or a stochastic approach. The basic difference between these two approaches is that a deterministic approach is based on perfect (positive or negative correlation) between the demographic components and between the values of the components in successive forecast years. A stochastic approach takes into account that correlations usually are not perfect.

If a scenario approach is followed, the correlation between the components depends on the assumptions about the relationships between fertility, mortality, and migration on the one hand and economic, social, cultural, and technological developments on the other. In a scenario the assumptions about fertility, mortality, and migration are based on the same set of assumptions about the economic, social, cultural, and technological explanatory variables. For example, if it assumed that the level of fertility, mortality, and net migration is positively related with economic growth, in a scenario based on high economic growth, fertility and net migration will be high and mortality will be low. The probability of a scenario depends on the probability of the assumed values of fertility, mortality, and migration, and their mutual correlation. The probability of the values of fertility, mortality, and migration depends on the probability of the values of the explanatory variables. The correlation between the three demographic components depends on the assumptions about the relationships between these components and the explanatory variables. If the correlations are strong, the probability of population size and age structure in each scenario can be expected to be rather close to the probability of the underlying assumptions of fertility, mortality, and migration. In practice the degree of probability of scenarios is usually not specified.

If for fertility, mortality, and migration low and high variants are specified as boundaries of the bandwidth of possible developments, a deterministic or a stochastic approach can be followed for calculating variants of population size and age structure. For calculating a complete probability distribution of the future age structure a stochastic approach is needed. If only a forecast interval of total population size is needed, a deterministic approach can be followed to approximate a confidence interval for total population size. Thus three approaches can be followed in assessing forecast intervals of population size:

1. Deterministic variants.
2. Quasi confidence intervals of population size.
3. Probability distribution of population size and age structure.

Ad 1. Deterministic variants.
This is the method most widely used in practice: the uncertainty variants are based on the same deterministic cohort-component model that is used for calculating the medium variant. Different assumptions for the components, fertility, mortality and migration, can be combined in various ways. One possibility is to publish extreme variants by combining high (low) fertility and net migration with low (high) mortality resulting in high (low) population growth. Clearly this procedure will not result in statistical confidence intervals for population size and age structure as it implies perfect correlation between forecast errors of the three components and between errors in successive forecast years. This is not very likely. There is no reason to assume that an overestimate of fertility always will go together with an underestimate of mortality. Neither is it plausible that in each successive forecast year the deviation between observation and forecast will have exactly the same direction as in the preceding forecast year. It can be expected that under- and overestimates partly will cancel out. This implies that if the intervals between the variants for fertility, mortality and migration are assumed to correspond with a given probability, the interval between the low and high variants of population size correspond with a higher probability. Nevertheless, deterministic low and high variants provide useful information on the degree of uncertainty of different forecast results. The uncertainty on fertility, mortality, and migration affects different age groups. Uncertainty on fertility in the short and medium term affects the number of children only, uncertainty on mortality mainly affects the aged and uncertainty on migration mainly affects the number of young adults. Hence the effect of combining high fertility, high migration and low mortality within the same variant on the size of separate age groups is limited, except for the very long run. It should, however, be noted that assuming perfect correlation between successive forecast years affects the probability of age groups even if the size of the age groups is mainly determined by one component. For example, assuming perfect correlation of fertility in successive years implies that the uncertainty of the age group 0-4 years in the 5th forecast year is overestimated, i.e. the probability of the forecast interval of that age group is larger than the probability of the forecast intervals of the total fertility rate in the separate years.
The degree of uncertainty of different age groups varies strongly. Users of the population forecasts who are mainly interested in the numbers of persons in the intermediate ages can use rather certain forecasts for a long period. For the forecasts of the number of children uncertainty is rather large. For the aged persons the uncertainty is not very large in absolute numbers, but it is large compared to the size of the group.
In addition to the low and high variants other forecast variants can be specified. Instead of combining the low and high variants of fertility, mortality and migration in such a way that they result in minimum and maximum population growth, it is possible to combine fertility and mortality in such a way that the differences in the age structure are larger. In the low variant both the numbers of young and old persons are small and in the high variant both numbers are large. Variants resulting in a young population can be obtained by combining high fertility and migration and low mortality. Similarly an old variant can be obtained by combining low fertility and migration, and high mortality. The reason for including high migration in the young variant is that immigrants tend to be younger than emigrants. If the young and high variants are based on the same fertility rates, the percentages aged 0-19 years are similar. This does not apply to the persons aged 65 or over. Even though the old and the high variant may be based on the same mortality rates, the percentages of aged persons tend to differ. The percentage of aged persons in the old variant may even differ less from the low variant than from the high variant. The explanation is that the uncertainty intervals for fertility and migration are larger than for mortality. In the high variant the number of aged persons is larger than in the low variant, but the numbers of persons younger than 65 years differ much more strongly between both variants. Clearly the uncertainty on the future percentage of aged persons according to the young and old variants is considerably larger than according to the low and high variants.
Instead of combining low and high variants of fertility, mortality, and migration there are still other possibilities. For example, high and low variants for one component can be combined with the medium variant of the other components in order to illustrate the effect of the uncertainty on each component separately. The intervals between these variants do not add up to the interval between the low and high variants, because there is a combination effect. The high variants of the three components reinforce each other. For example, in the high variant there are more migrants than in the medium variant, but they also have more babies and live longer, like the rest of the population. Hence the interval between the high and medium variant is larger than the sum of the intervals corresponding with the three separate high variants. However, the difference is not large, only a few percent of the total interval for total population size.

Ad 2. Quasi confidence intervals.
Even though deterministic variants do not provide a statistical confidence interval of future population size, there are two methods to approximate it:
a) Probabilities can be attached to deterministic variants.
b) The high and low variants can be assessed in such a way that they correspond with the boundaries of a two-thirds confidence interval of population size.

Ad a. Stoto (1988) proposes a simple method for assessing probabilities for the variants of a deterministic cohort-component model. He uses all 27 variants resulting from combining low, medium and high assumptions of fertility, mortality and migration. The method is based on the assumption that for each component the probability is 50% that the medium variant will be realised and that there is a 25% chance that each of the low and high variants will be realised. The underlying assumption is that the low and high variants correspond with the 10th and 90th percentiles of a normal distribution. On the basis of these assumptions a probability can be calculated for each of the 27 variants. One advantage of this method is that no perfect correlation between components in the low and high variants is assumed. One problem is, however, that the method is based on perfect serial correlation of forecast errors.

Ad b. Using the same deterministic cohort-component model that is used for calculating the medium variant, the intervals between the variants for fertility, mortality, and migration can be chosen in such a way that the interval between the low and high variants of population size corresponds rather closely to a statistical confidence interval. Long (1992) speaks of 'quasi' confidence intervals. He points out that when publishing new forecasts the U.S. Census Bureau compares the forecast intervals with errors of previous forecasts in order to provide an indication of the probability that the interval covers the future observation. It has become common practice to aim for two-thirds intervals rather than 95 percent intervals which are usual in tests of statistical significance. Keyfitz (1977) notes that no one is interested in 95 percent limits for forecasts since they are too wide for planning purposes.

The confidence interval of population size can be based on:

i. a time-series model of population growth;

ii. an analysis of historic forecast errors;

iii. a time-series model of forecast errors.

Ad i. Saboia (1974) specifies a statistical time-series model for population growth. On the basis of this model the variance of forecast errors can be assessed. The main problem of this approach is that the mean of the forecast distribution generally will not correspond with the medium variant of the forecasts of population size based on forecasts of fertility, mortality, and migration. Since the assessment of the confidence interval of population size based on a time-series model is only valid on the assumption that the time-series model is correct, if the projections of the time-series model are not consistent with the medium variant, the confidence interval based on the time-series model is not consistent with the medium variant.

Ad ii. On the basis of an analysis of forecast errors of population size of historic forecasts, empirical confidence intervals can be assessed. For example, on the basis of an analysis of forecast errors Stoto (1983) concludes that the high and low variants of the forecasts of the Bureau of the Census correspond approximately to a two-thirds confidence interval for the total population. One problem in using ex-post forecast errors to measure accuracy is that it requires a long series of forecasts. Other problems were mentioned in the previous section in discussing the use of ex-post forecasts for assessing the uncertainty of forecasts of fertility, mortality, and migration.

Ad iii. A statistical model of forecast errors can be specified. Stoto (1983) proposes a method that is based on an estimation of the standard deviation of forecast errors of the average growth rate. Stoto assumes that the forecast interval of the logarithm of population size increases linearly with the forecast lead time. However, this assumption is valid only if the growth rate is constant and if the uncertainty concerns the size of the growth only (De Beer, 1991d). Stoto ignores that forecast errors tend to be serially correlated. In the Netherlands, as in most other West European countries, constant growth rate is not a plausible assumption. If the growth rate decreases or increases in a more or less systematic way, forecast errors tend to be serially correlated. De Beer (1991d) presents a model that does not assume perfect correlation between fertility, mortality, and migration, but assumes that the correlation in the future will be the same as in the past. Furthermore the model does not assume perfect serial correlation. As a consequence the width of the forecast interval produced by the model increases less strongly with the length of the forecast period than the interval between deterministic low and high variants.

Ad 3. Probability distribution.

Based on statistical models of fertility, mortality, and migration, statistical forecast intervals of population size and age structure can be derived. either analytically or by means of simulations. In order to obtain a statistical forecast interval for the age structure of a population analytically a stochastic cohort-component model is needed. Application of such models, however, is very complicated. Analytical solutions require a large number of simplifying assumptions. Examples of applications of such models are given by Cohen (1986) and Alho and Spencer (1985). In both papers assumptions are specified of which the empirical basis is questionable. Instead of an analytical solution, confidence intervals can be derived from simulations. Pflaumer (1988) assumes rectangular distributions for the total fertility rate, life expectancy and net migration. Furthermore he assumes that the components are assumed to be independent of each other. One problem in Pflaumer’s approach is that the expectations of the distributions of the components are assumed to be constant through time. Hence the average of the simulations does not correspond with the medium variant.

Section 8 discussed the assessment of the probability of the bandwidth of future values of fertility, mortality, and migration. On the basis of these assumptions, the probability distribution of the future population size and age structure can be calculated by means of Monte Carlo simulations. For each year in the forecast period values of the total fertility rate, life expectancy at birth of men and women, and net migration are drawn from the probability distribution. Subsequently age-gender-specific fertility and mortality rates and migration numbers are specified. Each draw results in a population by age and gender at the end of each year. Thus the simulations provide a distribution of the population by age and gender in each forecast year.

The calculation of the simulations is based on several assumptions.

1. Type of probability distribution. For calculating the simulations it is not sufficient to specify the probability corresponding with the bandwidth between low and high variants for each component. A complete probability distribution needs to be specified for each component for each forecast year. If a normal distribution is

assumed, only two parameters have to be specified: the mean (corresponding with the medium variant) and the variance. Similarly the uniform distribution requires only two parameters to be specified, viz. the minimum and maximum values. If an asymmetrical distribution is assumed, at least one additional parameter indicating the skewness has to be specified. One disadvantage of an asymmetrical distribution is that the mean of the distribution does not correspond with the most probable value. This may be confusing for the users of the forecast.

2. Correlation between age-specific rates. The age-specific fertility and mortality rates in a given forecast year can be expected to be positively correlated. If the economic and social situation is favourable for having children, it can be expected that all age-specific fertility rates will be relatively high. For cohort data a negative correlation may be plausible. If the economic situation drives people to postpone having children, age-specific fertility rates at young and old ages may be negatively correlated. Similarly a selection mechanism may cause a negative relationship between mortality rates at young and old ages for the same cohort.
3. Serial correlation. The probability distributions of fertility, mortality, and migration in successive forecast years are correlated. If fertility is very high in one forecast year, it will not be very probable that fertility will be very low in the next year. Thus if a high value of fertility is drawn in one forecast year, the probability of drawing a high value in the next year should be higher than that of drawing a low value. In the short run a negative correlation may also be possible. For example, if the number of deaths is relatively high in one year due to a severe winter, the number of deaths may be relatively low in the next year due to a selection mechanism, as many frail people died in the previous year.
4. Correlation between components. The values of fertility, mortality, and migration can be correlated. For example, if immigrants have more children than the native population, an increase in the number of young immigrants may lead to an increase in the fertility rates in later years. Note that even if independence between fertility and mortality *rates* and migration numbers is assumed, there is no independence of *numbers* of births and deaths, and numbers of migrants. For example, if immigration is high in a certain year, this will result in larger numbers of births and deaths in later years for given values of fertility and mortality rates.

Summarising we can conclude that the uncertainty of population size and structure can be assessed by calculating deterministic forecast variants, based on alternative assumptions about fertility, mortality, and migration. These assumptions can be based on consistent assumptions about the determinants of the demographic components. This results in scenarios describing alternative possible futures. Or the assumptions can specify low and high values determining the bandwidth of possible values. One major disadvantage of deterministic variants is that usually the probability of the variants or of the area between the variants is not indicated. Thus for the user it is not clear how serious he should take all scenarios or variants. This information is needed since optimal decision making is based on weighing the benefits or costs of decisions in different circumstances and the probability that these circumstances will actually occur. Alternatively, probabilistic forecasts can be made. On the basis of assumptions about the probability of future values of fertility, mortality, and migration, and their mutual correlation, a probability distribution of the future population size and structure can be calculated. One problem is that the ‘true’ probability is not known. A probability distribution of the future is a forecast. Nevertheless a probability distribution is the best way in which forecasters can communicate the uncertainty of forecasts to the users.

10. Conclusions

Forecasts of population size and age structure are based on assumptions about fertility, mortality, and migration. Hence the main source of uncertainty of population forecasts is uncertainty about the future development of fertility, mortality, and migration (assuming there are accurate observations of the present size and structure of the population). The future is uncertain because the future development of fertility, mortality, and migration may be different from developments in the past or different in another way than expected. In what way the future may be different from the past depends on the way we look at the future. It is a metaphysical question whether the future of the real world is uncertain, i.e. whether the future is really different from the past or whether the future only appears to be different due to our lack of knowledge. But anyhow our knowledge of the future is uncertain. Future changes in demographic rates are uncertain because demographic rates for homogenous population categories may change or because the relative size of population categories may change in a different way than expected. If forecasts are based on an extrapolation of trends into the future, the future is uncertain because the future trend may differ from the trend in the observation period. If forecasts are based on an explanatory model the future is uncertain because the structural parameters or the functional form may be change or because the exogenous variables may develop in a different way than expected or because variables not included in the model may turn out to have an unexpected impact.
Thus uncertainty depends on the way we make assumptions about the future. Forecasts are based on the assumption that parameters describing changes in demographic variables through time, or relationships between the demographic variables to be forecasted and explanatory variables, are constant or on an assumption about the way the parameters may change and an assumption about the impact of future changes in exogenous variables. If a qualitative approach is followed, parameters are not quantified explicitly, but some implicit assumptions are made in order to arrive at quantitative forecasts. The uncertainty of forecasts depends on the validity of these assumptions.
Forecast variants or scenarios can be specified on the basis of an identification of sources of uncertainty. Variants may differ according to the uncertainty about the magnitude of changes in trends or the direction of

changes in exogenous variables. The width of the interval between variants gives an indication of the degree of uncertainty of the forecasts. The width of the interval, however, does not provide sufficient information. A wide interval does not necessarily imply that the forecast is very uncertain. As extreme variants may be very unlikely, the probability that the interval between the variants will cover the true value may be very high. By the same token, a small interval does not necessarily imply that the forecast is very certain: the probability may be very small. Thus forecast uncertainty is to be measured by the width of the forecast interval together with an indication of the corresponding probability. Hence the probability of a forecast interval is needed as a criterion for assessing the width of the forecast intervals. Moreover, probability provides a criterion for achieving consistency between the intervals for fertility, mortality, and migration. Differences in forecast intervals of fertility, mortality, and migration should reflect differences in uncertainty between these components. Another reason for specifying the probability of forecasts is that it allows to take into account the impact of aggregation on uncertainty. For example, the degree of uncertainty of a forecast of total population cannot simply be assessed by adding up the forecast intervals of the underlying components.

The true probability of future fertility, mortality, and migration is not known. The probability of a forecast is a forecast itself. This forecast can be based on an analysis of ex post forecast errors, on a model-based estimate of *ex ante* forecast errors or on experts' opinions about the possibility that the future will be different from the past. Because the assessment of the probability of a forecast is based on assumptions of which the validity is unknown, it is important that the forecaster gives an explicit account of the assumptions underlying the assessment of forecast uncertainty.

On the basis of assumptions about the probability distribution of fertility, mortality, and migration, forecast intervals of population size and age structure can be calculated using the cohort-component model. The calculations can be based on Monte Carlo simulations. For each age category a forecast interval corresponding with a given probability or, alternatively, a probability corresponding with a given interval can be calculated.

A probability distribution allows users to make their own decisions. A forecaster does not claim that certain events will or will not take place, but that some events are more likely than others. It is up to the user to decide how much risk he is prepared to take, i.e. whether or not he will take events that are not very likely, but that nevertheless are not impossible, into account in making his decisions. This is particularly important since population forecasts can be used for a variety of purposes. The group of users is heterogeneous. "All users do not want the same forecast" (Keyfitz, 1977). One user needs a conservative forecast, wanting to guard against the possibility that the realisation will be lower than the forecasts, another needs a high forecast. "Only the user can know how much he stands to lose through a projection being wrong in one direction or the other. This loss function may be strongly asymmetric." The loss function indicates how much loss a user would suffer with a particular departure of the forecast from the true population. Only if the user is given the population forecast in the form of a probability distribution, the user can decide how he can minimise the total expected loss. If the loss function is asymmetric, it may be advisable for the user not to use the most probable, medium variant, but another value. The choice depends on the weighing of the possible losses due to forecast errors and their probability. If the probability of the medium variant is considerably higher than the probability of other variants, the user should only choose another variant if the loss function is extremely asymmetric. If, on the other hand, the probability of alternative variants differs only slightly, the user may decide to use to use another variant if the asymmetry is not extreme.

References

Ahlburg, D.A., 1983, Good times, bad times: A study of the future path of U.S. Fertility. Social Biology 30, 17-23.

Ahlburg, D.A., 1986, A relative cohort size forecasting model of Canadian total live births. Social Biology 33, 51-56.

Ahlburg, D.A. and K.C. Land, 1992, Population forecasting: guest editor's introduction. International Journal of Forecasting 8, 289-299.

Al-Osh, M., 1986, Birth forecasting based on birth order probabilities, with applications to U.S. data, Journal of the American Statistical Association, 81, 645-656.

Alho, J.M., 1990, Stochastic methods in population forecasting. International Journal of Forecasting, 6, 521-530.

Alho, J.M., 1992, Estimating the strength of expert judgment. Journal of Forecasting 11, 157-167.

Alho, J.M., 1997, Scenarios, uncertainty and conditional forecasts of the world population. Journal of the Royal Statistical Society, A, 160, 71-85.

Alho, J.M. and B.D. Spencer, 1985, Uncertain population forecasting. Journal of the American Statistical Association, 80, 306-314.

Ascher, W., 1978, Forecasting: an appraisal for policymakers and planners (Johns Hopkins University Press, Balitomre).

Becker, G.S., 1960, An economic analysis of fertility. In: Demographic and economic change in developed countries (Princeton University Press, Princeton).

Bell, W.R., 1992, ARIMA and Principal Components Models in forecasting age-specific fertility. In: N. Keilman and H. Cruijsen (eds.), National population forecasting in industrialized countries, Swets & Zeitlinger, Amsterdam, 177-200.

Bell, W.R., J.F.Long, R.B. Miller and P.A. Thompson 1988, Multivariate time series projections of parameterized age-specific fertility rates. Research report no. 88/16, Statistical Research Division, Bureau of the Census, Washington, DC.

Blaug, M., 1980, The methodology of economics (Cambridge University Press, Cambridge)

Blommestein, H.J. and F.A. van Vught, 1988, Fallacies and errors in economic forecasting. Paper presented at the 8th International Symposium of Forecasting, June 12-15, 1988, Amsterdam.
Boudon, R., 1986, Theories of social change: a critical appraisal (Polity Press, Cambridge).
Box, G.E.P. en J.M. Jenkins, 1970, Time series analysis. Forecasting and control. Holden-Day, San Francisco.
Bozik, J.E. and W.R. Bell, 1987, Forecasting age specific fertility using principal components. SRD Research Report, Statistical Research Division, Bureau of the Census, Washington DC.
Brass, W., 1974, Perspectives in population prediction: illustrated by the statistics of England and Wales, Journal of the Royal Statistical Society A, 137, 532-583.
Brooks, C., D. Sams and L.S. Williams, 1982, An econometric model of fertility, marriage, divorce and labour force participation for Australian women, 1921/22 to 1975/76. IMPACT Research Centre, Melbourne.
Butz, W.P. and M.P. Ward, 1979, The emergence of countercyclical US fertility. American Economic Review 69, 318-328.
Calhoun, C.A. and J. De Beer, 1991, Birth Expectations and Population Forecasts: The Case of the Netherlands. In: Wolfgang Lutz (ed.), Future Demographic Trends in Europe and North America: What Can We Assume Today? Academic Press, London, pp. 197-223.
Caselli, G., 1993. Future longevity among the elderly. Department of Demography, University of Rome.
Collopy. F. and J.S. Armstrong, 1992, Expert opinions about extrapolations and the mystery of the overlooked discontinuities. International Journal of Forecasting 8, 575-582.
Coale, A.J. and J. Trussell, 1974, Model fertility schedules: variations in the age structure of childbearing in human populations, Population Index, 40, 185-206.
Cohen, J.E., 1986, Population forecasts and confidence intervals for Sweden: a comparison of model-based and empirical approaches. Demography, 23, 105-126.
Coleman, D.A., 1997, Converging and diverging patterns in Europe's populations. Paper presented at the European Population Conference, Cracow, 10-13 June 1997.
Daponte, D.O., J.B. Kadane and L.J. Wolfson, 1997, Bayesian demography: projecting the Iraqi Kurdish population, 1977-1990, Journal of the American Statistical Association 92, 1256-1267.
De Beer, J., 1983, Translation models for projecting cohort fertility rates Working paper, Netherlands Central Bureau of Statistics, Voorburg.
De Beer, J., 1985, A, Time Series Model for Cohort Data. In: Journal of the American Statistical Association 80, 525-530.
De Beer, J., 1988, Predictability of demographic variables in the short run. In: European Journal of Population 2, 283-296.
De Beer, J., 1989, Projecting age-specific fertility rates by using time-series models. In: European Journal of Population 5, 315-346.
De Beer, J., 1990, Uncertainty of international-migration projections for the 12 EC-countries. Paper presented at the Fourth Annual Meeting of the European Society for Population Economics, June 1990, Istanbul.
De Beer, J., 1991a, Geboorteontwikkeling wordt beïnvloed door vertrouwen in de economie. Maandstatistiek van de bevolking 39 (9), 25-30.
De Beer, J., 1991b, Births and cohort size. In: Social Biology 38, 146-153.
De Beer, J., 1991c, From birth expectations to birth forecasts: a partial-adjustment approach. In: Mathematical Population Studies 3, 127-144.
De Beer, J., 1991d, Forecasts of confidence intervals for population projections. Paper presented at the European Population Conference, October 21-25, Paris.
De Beer, J., 1992, General time series models for forecasting fertility. In: N. Keilman and H. Cruijsen, eds., National population forecasting in industrialized countries. Swets & Zeitlinger, Amsterdam, 147-175.
De Beer, J., 1994, Recent developments in the practice of household projections. Paper presented at the 1994 Annual Meeting of the Population Association of America, May 5-7, Miami.
De Beer, J., W. van Hoorn and J. Vermunt, 1991, Demographic changes in the Netherlands, with special reference to the relationship between female labour supply and fertility. In: J.J. Siegers and F. Tazelaar (eds.), The Dutch labour market in 2000. Demographic changes and policy implications. (Wolters-Noordhoff, Groningen), 23-52.
De Beer, J., H. Kuijper, R. Noordam, K. Prins and A. Sprangers, 1993, The linking of immigrant stock and flow data in the Netherlands; present and future possibilities. In: Statistical Journal 10, 321-334.
De Feijter, H., 1991, Voorlopers bij demografische veranderingen (NIDI, The Hague).
De Jong, A. and H. Visser, 1997, Long-term international migration scenarios for the European Economic Area. Eurostat working papers, Population and social conditions, E4/1997-6.
Devaney, B., 1983, An analysis of variation in U.S. fertility and female labor force participation trends. Demography 20, 147-161.
Easterlin, R.A., 1976, The conflict between aspirations and resources. Population and Development Review 2, 417-442.
Edmonston, B., 1992, Why refugees flee: an exploratory analysis of refugee emigration data. National Academy of Sciences, Committee on National Statistics, Washington, DC.
Ermisch, J., 1979, The relevance of the 'Easterlin hypothesis' and the 'new home economic' to fertility in Great Britain. Population Studies 33, 39-58.
Ermisch, J., 1992, Explanatory models for fertility projections and forecasts. In: N. Keilman and H. Cruijsen (eds.), National population forecasting in industrialized countries, Swets & Zeitlinger, Amsterdam, 201-222.
Farid, S.M., 1973, On the pattern of cohort fertility, Population Studies, 27, 157-168.
Feyerabend, P.K., 1975, Against method (NLB, London).

Fries, J.K., 1980, Aging, natural death, and the compression of morbidity. New England Journal of Medicine 303, 130-135.
Frinking, G. and Y. de Roo, 1979, The translation process as forecast instrument. (Working paper, NIDI, Voorburg).
Gardner, E.S., 1985, Exponential smoothing: the state of the art. Journal of Forecasting 4, 1-28.
Gleick, J., 1987, Chaos (Viking, New York).
Goodwin, P., 1997, Adjusting judgmental extrapolations using Theil's method and discounted weighted regression. Journal of Forecasting 16, 37-46.
Granger, C.W.J., 1969, Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37, 424-438.
Granger, C.W.J and P. Newbold, 1977, Forecasting economic time series (Academic Press, New York).
Gregory, P.R., J.M. Campbell and B.S. Cheng, 1972, A simultaneous equation model of birth rates in the United States. Review of Economics and Statistics 54, 374-380.
Hall, N. (ed.), 1992, The New Scientist guide to chaos (Penguin Books, London).
Harvey, A.C., 1984, A unified view of statistical forecasting procedures. Journal of Forecasting 3, 245-275.
Harvey, A.C., 1989, Forecasting, structural time series models and the Kalman filter (Cambridge University Press, Cambridge).
Heligman, L. and J.H. Pollard, 1980, The age pattern of mortality. Journal of the Institute of Actuaries 107, 49-80.
Hoem, J.M., 1973, Levels of error in population forecasts (Central Bureau of Statistics of Norway, Oslo).
Hoem, J.M. and E. Berge, 1975, Some problems in Hadwiger fertility graduation, Scandinavian Actuarial Journal, 129-144.
Hofstede, G., 1981, Culture's consequences. International differences in work-related values (Sage Publications, Beverly Hills, CA).
Huntington, S.P., 1996, The clash of civilizations and the remaking of world order (Simon & Schuster, New York).
Jencks, C., 1981, Destiny's tots. The New York Review of Books 28 (15), 30-32.
Johansson, S.R., 1997, Fertility and family history: using the past to explain the present. Population and Development Review 23, 627-637.
Kahneman, D. and A. Tversky, 1979, Intuitive prediction: biases and corrective procedures. In: S. Makridakis and S.C. Wheelwright (eds.) Forecasting (North Holland, Amsterdam), 313-327.
Keilman, N.W., 1990, Uncertainty in national population forecasting: issues, backgrounds, analyses, recommendations. Swets & Zeitlinger, Amsterdam.
Keilman, N., 1991, Emerging issues in demographic methodology. Paper presented at the European Population Conference, October, Paris.
Keyfitz, N., 1972, On future population. Journal of the American Statistical Association, 67, 347-363.
Keyfitz, N., 1977, Applied mathematical demography (John Wiley & Sons, New York)
Keyfitz, N., 1982, Can knowledge improve forecasts? Population and Development Review 4, 729-751
Keyfitz, N., 1990, Future mortality and its uncertainty. Paper presented at the annual meeting of the Population Association of America, Toronto, May.
Keyfitz, N., 1990b, A comment. In: C.A. Hazeu and G.A.B. Frinking (eds.), Emerging issues in demographic research (Elsevier Science Publishers, Amsterdam), 67-69.
Kirk, D., 1996, Demographic transition theory. Population Studies 50, 361-387.
Klaassen, L.H. and Z. Pawlowski, 1981, Long-term forecasting: medidations of two pitfall collectors. De Economist 129, 455-475.
Kuhn, T.S., 1970a, The structure of scientific revolutions (University of Chicago Press, Chicago).
Kuhn, T.S., 1970b, Reflections on my critics. In: A. Musgrave and I. Lakatos (eds.), Criticism and the growth of knowledge (Cambridge University Press, Cambridge).
Lakatos, I., 1970, Falsification and the methodology of scientific research programmes. In: A. Musgrave and I. Lakatos (eds.), Criticism and the growth of knowledge (Cambridge University Press, Cambridge).
Lawrence, M. and M. O'Connor, 1992, Exploring judgmental forecasting, International Journal of Forecasting 8, 15-26.
Leach, D., 1981, Re-evaluation of the logistic curve for human populations. Journal of the Royal Statistical Society A 144, 94-103.
Lee, R.D., 1974, Forecasting births in post-transitional populations: stochastic renewal with serial correlated fertility. Journal of the American Statistical Association 69, 607-617.
Lee, R.D., 1976, Demographic forecasting and the Easterlin hypothesis. Population and Development Review 2, 459-468.
Lee, R.D. and L. Carter, 1990, Modeling and forecasting US mortality. Working paper no. 30, University of California, Berkeley.
Lee, R.D., L. Carter and S. Tuljapurkar, 1995, Disaggregation in population forecasting: do we need it? And how to do it simply? Mathematical Population Studies 5, 217-234.
Leibenstein, H., 1980, Relaxing the maximization assumption in the economic theory of fertility. Paper presented at Seminar on determinants of fertility trends: major theories and new directions for research, Bad Homburg.
Long, J.F., 1992, Accuracy, monitoring, and evaluation of national population projections. In: N. Keilman and H. Cruijsen (eds.), National population forecasting in industrialized countries, Swets & Zeitlinger, Amsterdam, 129-145.
Lutz, W., 1995, Scenario analysis in population projection. Paper presented at the Symposium on analysis of errors in demographic forecasts with implication on policy, March 30- April 2, 1995, Koli, Finland.

Lutz, W., J.R. Goldstein and C. Prinz, 1996, Alternative approaches to population projection. In: W. Lutz (ed.), The future population of the world. What can we assume today? Revised edition (Earthscan, London), 14-44.
Lutz, W., W. Sanderson and S. Scherbov, 1996, Probabilistic population projections based on expert opinion. In: W. Lutz (ed.), The future population of the world. What can we assume today? Revised edition (Earthscan, London), 397-428.
Lutz, W., W. Sanderson, S. Scherbov and A. Goujon, 1996, World population scenarios for the 21st century. In: W. Lutz (ed.), The future population of the world. What can we assume today? Revised edition (Earthscan, London), 361-396.
Mackenbach, J., 1988, Mortality and medical care (thesis, Leiden).
Macunovich, D.J. and R.A. Easterlin, 1988, Application of Granger-Sims causality tests to monthly fertility data, 1958-1984, Journal of Population Economics, 1, 71-89.
Macura, M., 1994, Overview. In: M. Macura and D. Coleman, eds., International migration: regional processes and response. Economic Studies no. 7, United Nations, Geneva, 1-20.
Madduri, V.B.N. and K.L. Gupta, 1974, An econometric model of fertility behaviour in Canada. Proceedings of the Business and Economic statistics section, American Statistical Association, 284-288.
Makridakis, S. et al., 1982, The accuracy of extrapolation (time series) methods: results of a forecasting competition. Journal of Forecasting 1, 111- 153.
Manton, K.G., E. Stallard and H.D. Tolles, Limits to human life expectrancy: evidence, prospects and implications. Population and Development Review 17, 603-637.
Massey, D.S., 1990, Social structure, household strategies, and the cumulative causation of migration. Population Index 56, 3-26.
Massey, D.S. et al. , 1993, Theories of international migration: a review and appraisal. Population and Development Review 19, 31-466.
Massey, D.S., J. Arango, G. Hugo, A. Kouaouci, A. Pellegrino and J.E. Taylor, 1994, An evaluation of international migration theory: the North American Case. Population and Development Review 20, 699-751.
McDonald, J., 1979, A time series approach to forecasting Australian total live births. Demography 16, 575-601.
McNees, S.K., 1990, The role of judgment in macroeconomic forecasting accuracy, International Journal of Forecasting 6, 287-299.
McNeil, D.R., T.J. Trussell and J.C. Tumer, 1977, Spline interpolation of demographic data, Demography, 14, 245-252.
McNown, R. and A. Rogers, 1989, Forecasting mortality: a parameterized time series approach. Demography 26, 645-660.
McNown, R and A. Rogers, 1992. Forecasting cause-specific mortality using time series models. In: International Journal of Forecasting
Miller, R.B., 1986, A bivariate model for total fertility rate and mean age of childbearing. Insurance: Mathematics and Economics 5, 133-140.
Mitra, S. and A.Romaniuk, 1973, Pearsonian Type I curve and its fertility projection potentials, Demography, 10, 351-365.
Murphy, M, 1990. Methods for forecasting mortality for population projections. In: OPCS Population projections: Trends, methods and uses. Occupational paper 38, London.
Nusselder, W., 1998, Compression or expansion of morbidity? A life-table approach (thesis, Erasmus University Rotterdam).
O'Connell, M. and C.C. Rogers, 1983, Assessing cohort birth expectations data from the current population suvey, 1971-1981. Demography 20.
Olshansky, S.J. and B.A. Carnes, 1994, Demographic perspectives on human senescence. Population and Development Review 20, 57-80.
Passel, J.S., 1976, Accuracy and precision of fertility projections made with time-series models. Proceedings of the Social Statistics Section, American Statistical Association, 660-665.
Peto, R., A.D. Lopez, J. Boreham, M. Thun and C. Heath, 1994, Mortality from smoking in developed countries, 1950-2000 (Oxford University Press, Oxford).
Pflaumer, P., 1988, Confidence intervals for population projections based on Monte Carlo methods. International Journal of Forecasting 4, 135-142.
Pitcher, B.L., 1978, Family formation patterns: a social learning model. Unpublished dissertation, University of Arizona.
Pittenger, D.B., 1977, Population forecasting standards: some considerations concerning their necessity and content, Demography, 14,363-368.
Popper, K.R., 1959, The logic of scientific discovery (Hutchinson, London).
Rogers, A., 1986, Parameterized multistate population dynamics and projections, Journal of the American Statistical Association, 81, 48-61.
Rogers, A. and L. Castro, 1981, Model migration schedules. Research report, IIASA, Laxenburg.
Romaniuk, A., 1973, A three parameter model for birth projections,Population Studies, 27, 467-478.
Ross, J.A. and S. Madhavan, 1981, A Gompertz model for birth interval analysis, Population Studies, 35, 439-454.
Russell, S.S., 1997, International migration: implications for the World Bank. Working paper, World Bank.
Ryan, A., 1970, The philosophy of the social sciences (Macmillan, London).
Ryder, N.B., 1964, The process of demographic translation, Demography, 1,74-82.
Ryder, N.B., 1965, The cohort as a concept in the study of social change. American Sociological Review 30, 843-861.

Saboia, J., 1974, Modeling and forecasting populations by time series: the Swedish case. Demography, 11, 483-492.
Saboia, J.L.M., 1977, Autoregressive Integrated Moving Average (ARIMA) models for birth forecasting. Journal of the American Statistical Association 72, 264-270.
Santow, G., 1997, The mortality, epidemiological and health transitions: Their relevance for the study of health and mortality. Stockholm Research Report in Demography no. 122.
Sims, 1980, Macroeconomics and reality. Econometrica 48, 1-48.
Smith, S.K., 1997, Further thoughts on simplicity and complexity in population projection models. International Journal of Forecasting 13, 557-565.
Stoto, M.A., 1983, The accuracy of population projections. Journal of the American Statistical Association, 78, 13-19.
Stoto, M.A., 1988, Dealing with uncertainty: statistics for an aging population. The American Statistician, 42, 103-110.
Tabeau, E., P. Ekamper, C. Huisman and A. Bosch, 1997. Forecasting of mortality by gender, age/generation, and main causes of death (NIDI, The Hague).
Theil, H., 1966, Applied economic forecasting (Rand-McNally, Chicago).
Theil, H., 1971, Principles of econometrics (North Holland, Amsterdam).
Thompson, P.A., W.R. Bell, J.F. Long en R.B. Miller, 1989, Multivariate time series projections of parameterized age-specific fertility rates. Journal of the American Statistical Association, 84, 689-699.
Tuljapurkar, S. and C. Boe, 1997, Mortality change and forecasting: how much and how little do we know? Mountain View Research, Los Altos, CA.
Valkonen, T. and F.W.A van Poppel, 1997. The contribution of smoking to gender differences in life expectancy; four Nordic countries and the Netherlands 1970-1989. In: European journal of public health.
Van Alphen, H.J., 1982, Population prediction using a logistic with variable coefficients: an application to Dutch data (Free University, Amsterdam).
Van Duijn, J.J., 1983, The long wave in economic life.
Van de Giessen, H., 1992, Using birth expectations information in national population forecasts. In: N. Keilman and H. Cruijsen (eds.), National population forecasting in industrialized countries, Swets & Zeitlinger, Amsterdam, 223-241.
Van Hoorn, W. and J. de Beer, 1997, The latest official projections of Dutch mortality: method and assumptions. Paper presented at the Workshop ‘Forecasting of mortality in developed countries: searching for better methods and realistic assumptions’, The Hague, 5 September 1997.
Van de Kaa, D.J., 1988, The second demographic transition revisited: theories and expectations. Paper presented at the Conference on Population and European Society, Florence, 7-9 December 1988.
Vaupel, J.W. and H. Lundstrom, 1994, The future of mortality at older ages in developed countries. In: W. Lutz (ed.), The future population of the world. What can we assume today? (Earthscan, London), 295-315.
Vermunt, J.K., 1991a, Leefstijl en demografisch gedrag: een toepassing van latente-klasse-analyse. Maandstatistiek van de bevolking 39 (11).
Vermunt, J.K., 1991b, Een model ter bepaling van de cohortcomponent in de sterfte. In: L. Th. van Leeuwen and H.G.J.M. Cruijsen (eds.), Sterfte en gezondheid - nu en straks. (Nederlandse Vereniging voor Demografie, The Hague).
Venieris, Y.P., F.D. Sebold and R.D. Harper, 1973, The impact of economic, technological and demographic factors on aggregate births. Review of Economics and Statistics 55, 493-497.
Ward, M.P. and W.P. Butz, 1980, Completed fertility and its timing. Journal of Political Economy 88, 917-940.
Westoff, C.F., 1978, The predictability of fertility in developed countries. Population Bulletin of the United Nations 1978 (11), 1-5.
Westoff, C.F. and N.B. Ryder, 1977, The predictive validity of reproductive intentions. Demography 14, 431-453.
WHO European Centre for Environment and Health, 1995. Concern for Europe's Tomorrow: Health and the Environment in the WHO European Region. Wissenschaftliche Verlagsgesellschaft, Stuttgart.
Wilkinson, M., 1973, An econometric analysis of fertility in Sweden. Econometrica 41, 431-453.
Willekens, F. and N. Baydar, 1984, Age-period-cohort models for forecasting fertility (Working paper, Netherlands Interuniversity Demographic Institute).
Willekens, F.J., 1990, Demographic forecasting; state-of-the-art and research needs. In: C.A. Hazeu and G.A.B. Frinking (eds.), Emerging issues in demographic research (Elsevier Science Publishers, Amsterdam) 9-66.
Williams, W.H. en M.L. Goodman, 1971, A simple method for the construction of empirical confidence limits for economic forecasts. Journal of the American Statistical Association 66, 752-754.