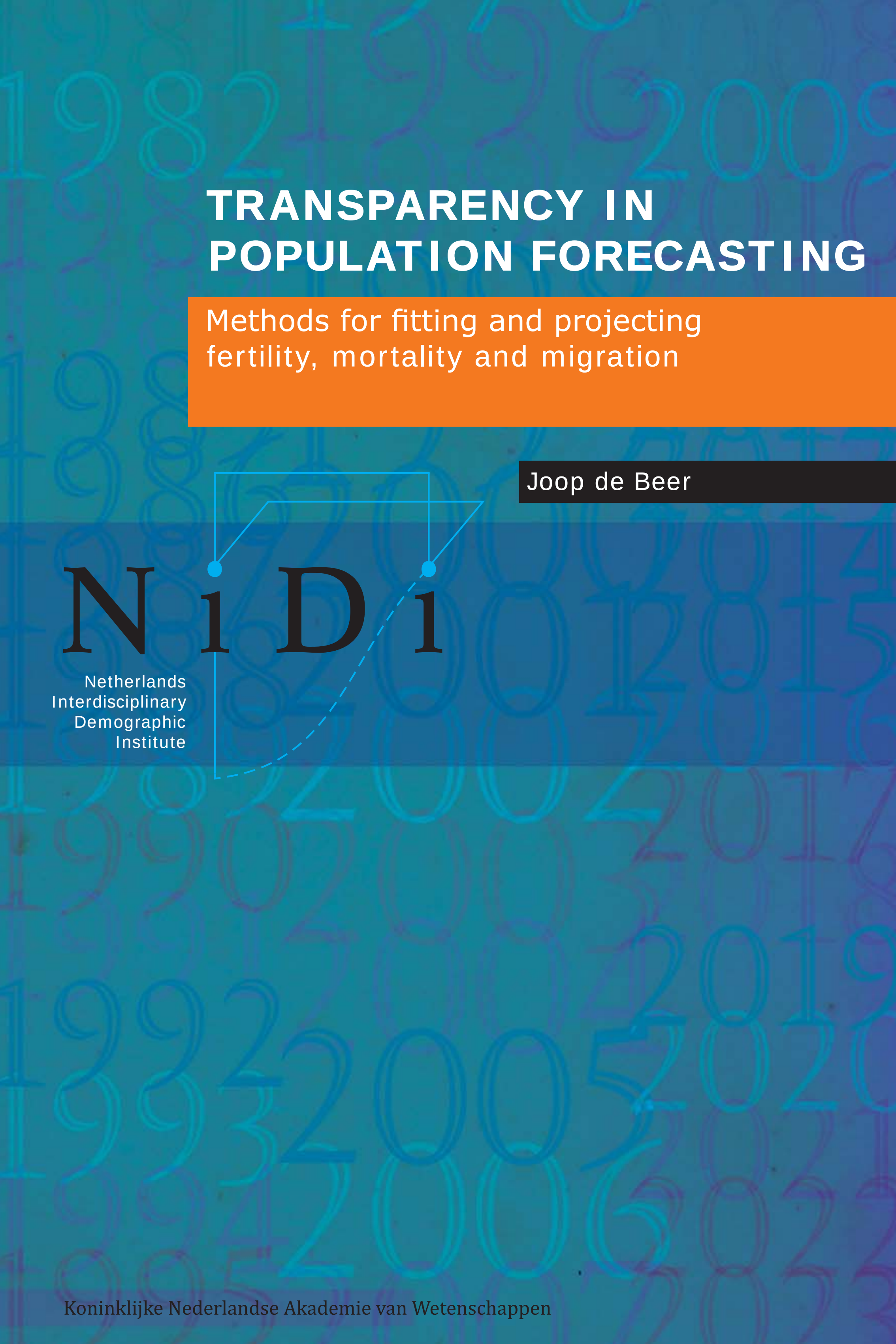


TRANSPARENCY IN POPULATION FORECASTING

Methods for fitting and projecting fertility, mortality and migration

Joop de Beer

NiDi

Netherlands
Interdisciplinary
Demographic
Institute

Koninklijke Nederlandse Akademie van Wetenschappen

Transparency in population forecasting

TRANSPARENCY IN POPULATION FORECASTING

Methods for fitting and projecting fertility, mortality and migration

Joop de Beer

netherlands
interdisciplinary
demographic
institute

NIDI book nr. 83

Koninklijke Nederlandse Akademie van Wetenschappen

Amsterdam, 2011

The series of NIDI books is published by the Netherlands Interdisciplinairy Demographic Institute

Director:
Leo van Wissen

Editors:
Joop de Beer, Tineke Fokkema, Frans van Poppel

Editorial secretariat:
Netherlands Interdisciplinairy Demographic Institute
PO Box 11650, 2502 AR The Hague
Lange Houtstraat 19, 2511 CV The Hague
Telephone: +31-70-3565200
Fax: +31-70-3647187
E-mail: Info@Nidi.nl
Internet: http://www.nidi.nl

Technical coordination:
Jeannette van der Aar, Tonny Nieuwstraten, Jacqueline van der Helm

Publisher:
Amsterdam University Press
Herengracht 221
1016 BG Amsterdam
The Netherlands
www.aup.nl

ISSN 0922-7210
ISBN 978 90 6984 637 8

Table of contents

List of tables

List of figures

Acknowledgements

Four chapters are reprinted from the following publications:

Chapter 2:
J. de Beer, J. Raymer, R. van der Erf and L. van Wissen, 2010, ‘Overcoming the problems of inconsistent international migration data: A new method applied to flows in Europe’, *European Journal of Population* 26: 459-481.

Chapter 3:
J. de Beer, 2008, ‘Forecasting international migration: Time series projections vs. argument-based forecasts’. In: J. Raymer and F. Willekens (eds.), *International migration in Europe: Data, Models, and Estimates*. Chichester: Wiley, 283-306.

Chapter 4:
J. de Beer and I. Deerenberg, 2007, ‘An explanatory model for projecting regional fertility differences in the Netherlands’. *Population Research and Policy Review* 26, 511-528.

Chapter 5:
J. de Beer, 2011, ‘A new relational method for smoothing and projecting age-specific fertility rates: TOPALS’. *Demographic Research* 24: 409-454.

1. Introduction

1.1. Transparency of population projections, scenarios and forecasts

Population projections are widely used. Projections of future population growth and population ageing have an important impact in many policy areas. For example, population ageing may lead to an increase in the future costs of pensions, health care and long term care. The decline in the growth rate of the working age population may have an adverse effect on future economic growth. The ageing of the work force may reduce the future growth rate of labour productivity. The growth of population size may lead to an increase in the future demand of energy. These are only few examples of the effects of changes in the size and age structure of the population. Thus the usefulness of population projections is obvious. Projections of the future size and age structure of the population are based on assumptions about the future levels of fertility, mortality, and migration. Starting from the current population by age and sex, changes in the levels of fertility, mortality and migration determine future changes in the population size and age structure.

If the forecaster makes assumptions about the most likely future development in fertility, mortality and migration, the results of his or her calculations of future population changes can be regarded as forecasts. Although it is not certain that these changes will occur, the forecaster considers these developments to be more likely than other developments, given the knowledge available at the moment that the forecasts are made. Alternatively, the forecaster may calculate future changes in population under the assumption that current trends in fertility, mortality, and migration will continue. If the forecaster does not indicate whether this should be considered as the most likely development, the results of these calculations can be regarded as projections. More generally, projections can be regarded as outcomes of any set of assumptions about future trends without a statement that this is expected to be the most likely future development. Thus publishing projections seems less risky than publishing forecasts. If the assumptions underlying the projections will not come true, *e.g.* if current trends in fertility, mortality and migration will not continue, the projections will not be accurate, but the forecaster cannot be blamed, because he or she has not claimed that this would actually happen. This may be one reason why many statistical agencies label the outcomes of their calculations as projections. However, Keilman (2008) argues that unless the agency presents its assumptions as unrealistic, the projections published by statistical agencies can be viewed as forecasts, indicating a likely development.

In order to emphasize the uncertainty of forecasts, it has become common practice to publish alternative scenarios. Scenarios are aimed to describe possible futures. Often one scenario is labelled as baseline, reference or business as usual scenario. Usually this scenario is a projection, assuming a continuation of trends. If this is assumed to be a likely development, this scenario can be viewed as a forecast. The other scenarios show possible alternative developments. The scenarios can be based on an identification of the main driving forces of changes in fertility, mortality and migration and on assumptions about possible future developments in these driving forces, *e.g.* economic, social, cultural, technological and political changes. Alternatively, the scenarios can be specified on the basis of an assessment of the size of forecast errors. If it is assumed that forecast errors will be large, the interval between scenarios should be large. For example, if the forecaster assumes that the future development of migration is very uncertain, the range between alternative scenarios of future migration should be wide, whereas if the future development of fertility is considered to be not that uncertain, the range between scenarios of future fertility may be relatively small.

The methods used for making projections and scenarios may differ. If projections are based on the assumption that trends will continue, time-series models can be used to estimate the trend and to extrapolate the trend into the future. If scenarios are based on assumptions about alternative developments in the driving forces of fertility, mortality and migration, explanatory models can be used to assess the size of the effects of these driving forces on the levels of fertility, mortality and migration. However, there is no dichotomy. Scenarios may be based on time-series models rather than on explanatory models, *e.g.* by assuming a deceleration of change in one scenario and an acceleration in another. Chapter 6 shows how alternative scenarios of future changes in life expectancy can be based on alternative assumptions about changes in age-specific death probabilities. When making projections, an explanatory model can be used to assess the effect of short term fluctuations. This allows to identify long-term trends which can be the basis for long-term projections. Chapter 3 shows how an explanatory model can be used to estimate the effect of the business cycle on short-term fluctuations in immigration. Similarly Fokkema *et al.* (2008) show how the business cycle affects short-term fluctuations in the total fertility rate.

Whether an extrapolation method is used to make projections or an explanatory model is used to make scenarios, the forecaster needs to make choices and assumptions. When making projections it may make a lot of difference which base period is chosen for estimating the trend. Chapter 3 shows that in projecting emigration from the Netherlands a long base period suggests that migration shows random fluctuations around a constant level, whereas a short base period suggests that there is an increasing trend. Another choice to be made by the forecaster is the type of time-series model. Deterministic time series models, such as a linear trend model, assume that there is a fixed trend that is not affected by random fluctuations. Stochastic time series models, such as the random walk with drift model, are based on the assumption that the trend is subject to random changes. This implies that recent fluctuations affect the level of the trend. The projections made by deterministic models do not react quickly to recent changes in the time series as these are viewed as short term deviations from the long term trend. Chapter 3 shows that the choice of the time series model may lead to quite different projections of migration. Chapter 6 shows that different time series models for age-specific death probabilities lead to different projections of life expectancy. One additional choice to be made is the indicator to be projected. Chapter 3 shows that separate projections of immigration and emigration may lead to a different projection of net migration than directly projecting net migration. Projecting different types of immigration separately (such as labour, family and asylum migration) may result in a different projection of total immigration than projecting total immigration directly. Chapter 6 shows that projecting age-specific death probabilities will lead to a different projection of life expectancy than projecting life expectancy directly.

When making scenarios on the basis of assumptions about future changes in the main driving forces of fertility, migration and mortality choices have to be made about the method that will be used to assess the effects of the driving forces. Assumptions can be based on explanatory models, disaggregation or expert opinions. Chapter 4 describes how an explanatory model can be used to assess the effects of demographic, socioeconomic and cultural explanatory variables on regional differences in the level of fertility. The results can be used to specify scenarios of future differences in fertility on the basis of assumptions about future changes in the explanatory variables. Chapter 3 shows how disaggregation of immigration numbers by migration motive can be used to identify explanations of changes in immigration and to make assumptions about future changes. Chapter 7 discusses a method proposed by Lutz (2009) to assess the effects of a set of driving forces on future changes in fertility, mortality and migration from a survey among experts. One benefit of this approach compared with the specification of a quantitative explanatory model is that the choice of explanatory variables is not restricted by the availability of data which are needed for estimating the coefficients of an explanatory model. One limitation of this approach is that experts give qualitative judgements about the direction of changes, but that these qualitative arguments need to be translated into quantitative assumptions about future changes in fertility, mortality and migration. Chapter 7 discusses how Lutz (2009) deals with this problem. Since the number of explanatory variables included in Lutz's argument-based approach is considerably larger than the number of variables that can be included in a quantitative explanatory model, the resulting assumptions about future changes in fertility, mortality and migration can be expected to differ.

Thus whether the forecaster makes a projection or alternative scenarios, he or she has to make choices about the type of method to be used, the base period, the selection of indicators and explanatory variables and to make assumptions about the continuation of past trends in the future and about future changes in driving forces. The arguments given for making these choices and assumptions determine whether a projection or a scenario can be regarded as a forecast. If the forecaster argues that a continuation of a trend is likely because the trend has been manifest for a long period, this projection can be viewed as a forecast. For example, Oeppen and Vaupel (2002) show that best practice life expectancy (*i.e.* the highest level of life expectancy in the world in each year) has followed a linear trend over a period of a century and a half and they argue that there is no reason to assume that this trend will not continue in the coming decades. This is a forecast rather than only an arbitrary projection. Chapter 3 argues that labour migration is affected by the situation on the labour market, that the future decline in the working age population will lead to shortages in the labour market and thus that it is plausible to assume that in the future labour migration will be higher than in the past. This is a forecast rather than only an arbitrary scenario. Chapter 4 argues that the future effects of demographic, socioeconomic and cultural developments on fertility will counterbalance each other and thus that differences in fertility between small and large cities will not disappear. This can be viewed as a forecast rather than as an arbitrary scenario. Alternatively, Chapter 6 shows how the future rise in life expectancy may differ if trends are projected in a different way. If each of these trends is based on valid arguments and the forecaster does not give arguments why one projection is more likely than the others, these are scenarios rather than forecasts.

Thus whether a projection or a scenario can be viewed as a forecast does not depend on the method used but on the arguments underlying the choices and assumptions made by the forecaster. Since the terms projections, forecasts and scenarios are often used interchangeably, the label used by the forecaster to describe his or her calculations does not provide sufficient information. Rather it is important that the forecaster makes decisions and assumptions underlying the choice and application of methods explicit as this will allow the user to determine how projections and scenarios can be used. Armstrong (2001) argues that users often cannot judge the quality of a forecast, but they can decide whether the forecasting process was reasonable. This requires that it is necessary for users to know which decisions are made by the forecaster. Therefore projections and scenarios should be transparent. Transparency requires that in addition to explaining which method is used, the forecaster should specify which underlying choices and assumptions are made, what the arguments for these choices and assumptions are, and what the consequences of these choices and arguments are, *e.g.* by means of sensitivity analyses or by presenting alternative scenarios. Transparency is not an aim in itself. The main aim of a forecast is accuracy: a forecast should give an accurate description of future developments. The aim of scenarios is to show possible future developments, so that the policy maker can take these into account when making plans. However, since the accuracy of forecasts and the plausibility of scenarios are not yet known at the moment that forecasts or scenarios are made, the user can only judge the way forecasts and scenarios are made and this requires that the forecasting process is made transparent.

The aim of this book is to present methods that can be used for making projections and scenarios in a transparent way. Chapters 2 to 6 will discuss choices that are to be made by forecasters and arguments that can be used to determine whether a projection or scenario can be regarded as a forecast. The aim is not to present one model that will outperform all other models and that will produce ‘objective’ forecasts, *i.e.* forecasts that do not depend on choices to be made by the forecaster. Models are very useful instruments for the forecaster, but they are not more than a tool. Forecasts do not automatically follow from a model. It is inevitable that the forecaster has to make choices and it is important that these choices are made explicitly on the basis of arguments and do not remain implicit. In order to improve transparency the methods described in the following chapters are kept as simple as possible. If models are complicated, it is difficult to assess the implications of choices for the outcomes. Even though most chapters in this book include formulas, the basic underlying ideas are simple.

1.2. Outline of this book

This book includes five empirical chapters. The first two chapters deal with international migration, the subsequent two chapters discuss fertility, and the fifth chapter is about mortality. The nature of changes in fertility, mortality and migration differs. Mortality tends to show gradual long-term trends, migration shows large short-term fluctuations, and the level of fertility is affected by changes in the age pattern. Therefore different methods are needed to project future changes in migration, mortality, and fertility. Two chapters in this book are based on Dutch data. One reason is that good and detailed data are available for the Netherlands. The other three empirical chapters use data for several European countries.

The first step in making forecasts is to assess the quality of data. If data have poor quality, forecast accuracy of methods using these data will be poor (Keilman, 2008). Most European countries have good data on fertility and mortality, but data on international migration tend to be less reliable or even lacking. One way of improving statistics on international migration is to compare data from different countries. For example, Germany reported that in the period 2002-2007 on average 136,000 immigrants per year arrived from Poland, whereas Poland reported that on average 14,000 emigrants moved to Germany. Obviously we cannot simply use such reported migration numbers to make projections, particularly if projections for several countries need to be made. Chapter 2 shows how migration data can be improved by using a simple model that compares data from different countries in order to estimate to what extent migration statistics may under- or overestimate the real numbers.

Chapter 3 discusses methods for projecting international migration. It compares time series projections and argument-based forecasts. The analyses are based on Dutch data. Time series of international migration tend to show larger fluctuations over the years than fertility and mortality. One explanation is that the nature of migration has changed over time. Different types of migrants, such as labour migrants, family migrants and asylum seekers, react differently to economic, political and cultural developments. This implies that the projected direction of change may differ by type of migration. Chapter 3 argues that argument-based forecasts should be based on a distinction between the main categories of migrants.

Chapters 4 and 5 deal with fertility. Chapter 4 discusses regional differences in fertility and chapter 5 focuses on international differences. Chapter 4 illustrates how an explanatory model can be used for making argument-based forecasts. In contrast with the other chapters this chapter uses regional data. The chapter examines how differences in the level of the total fertility rate (TFR) between small and large cities in the Netherlands can be explained. Large cities tend to have a lower TFR than small cities. Different types of explanatory variables are included. Whereas projections on the national level focus on projecting the future level of fertility, regional projections focus on projecting regional differences. The model described in chapter 4 is used to develop arguments to answer the question whether the differences in the TFR between large and small cities will be persistent or whether a converging development may be expected. Statistics Netherlands and the Netherlands Environmental Assessment Agency use this model for making assumptions about fertility for the Netherlands regional population forecasts.

Assumptions about future changes in the level of fertility are usually based on assumptions about the future level of the TFR. However changes in the level of the TFR are affected by changes in the age pattern of fertility. Since these effects have temporary effects on the level of the TFR, it is important to assess the size of these effects before making a projection for the long run. If these effects are ignored a temporary decline or increase in the TFR may be projected into the future as if it were a permanent decline or increase. For that reason projections of fertility should take into account changes in age-specific fertility rates. Separate projections of age-specific fertility rates for each age tend to result in irregular patterns. Therefore it is common practice to smooth age specific rates before making a projection. Chapter 5 introduces the relational method TOPALS that produces a smooth age schedule by means of calculating the ratios of the age specific fertility rates to be projected and the rates described by a smooth standard age schedule. The age pattern of the ratios can be described by a linear spline. By making assumptions about the future values of the rate ratios at selected ages, the so-called knots, TOPALS can be used for making smooth projections of age-specific fertility rates. If the standard age schedule is a so-called target age pattern, a partial adjustment model can be used to project the speed with which the age-specific fertility rates will move toward the target level. In chapter 5 projections of fertility rates for six European countries are calculated under the assumption that the current Swedish fertility pattern can be regarded as the target pattern.

Life expectancy has been increasing in most European countries over a long period of time. There is general agreement among most experts that life expectancy will continue to grow, but there is less agreement about the size of the increase. Some optimistic experts expect that life expectancy will continue to grow by 2.5 years per decade. Other experts assume that the rate of increase in life expectancy will slow down, since a linear increase in life expectancy could be achieved by an acceleration of the decrease in age-specific death probabilities only. Chapter 6 shows how TOPALS can be used for projecting age-specific death probabilities. Oeppen and Vaupel (2002) expect that the ‘best practice’ life expectancy of Japanese women will continue to increase in the coming decades. Projected age-specific death probabilities that are consistent with that projection can be used as the target for other countries. A partial adjustment model is used to make projections of age-specific death probabilities for 26 European countries under the assumption that they will move to that age pattern in the (very) long run.

Chapter 7 summarizes the main findings about the use of methods for making projections and scenarios of future migration, fertility, and mortality and discusses the use of these methods for improving transparency of population projections and scenarios.

2. Overcoming the problems of inconsistent international migration data: A new method applied to flows in Europe

Abstract

Due to differences in definitions and measurement methods, cross-country comparisons of international migration patterns are difficult and confusing. Emigration numbers reported by sending countries tend to differ from the corresponding immigration numbers reported by receiving countries. In this chapter, a methodology is presented to achieve harmonised estimates of migration flows benchmarked to a specific definition of duration. This methodology accounts for both differences in definitions and the effects of measurement error due to, for example, under reporting and sampling fluctuations. More specifically, the differences between the two sets of reported data are overcome by estimating a set of adjustment factors for each country's immigration and emigration data. The adjusted data take into account any special cases where the origin-destination patterns do not match the overall patterns. The new method for harmonising migration flows that we present is based on earlier efforts by Poulain (1993, 1999) and is illustrated for movements between 19 European countries from 2002 to 2007. The results represent a reliable and consistent set of international migration flows that can be used for understanding recent changes in migration patterns, as inputs into population projections and for developing evidence-based migration policies.

2.1. Introduction

Our understanding of the mechanisms and patterns of international migration over time are impeded both by the lack of data and by inconsistencies in the measurement and collection of the data that are available. In fact, it is well known that the patterns of migration vary significantly depending on which country is reporting the data (Kupiszewska and Nowok, 2008; Nowok *et al.*, 2006 and Zlotnik, 1987). Considering that international migration is the main factor contributing to population growth in Europe, this is very unfortunate. In response to the problem of inconsistent migration data, we have developed a methodology for harmonising the data available to us from countries in Europe. More specifically, we make use of doubly-counted information obtained from migrant sending and migrant receiving countries to estimate adjustment factors necessary for producing a consistent set of migration flows. These estimated flows are benchmarked to a particular definition.

Harmonisation of migration data is required for the development of policies on immigration (Kraier *et al.*, 2006). Differences in both the concepts and techniques used to measure migration make any international comparison of migration difficult. There has been a lot of work on data issues and migration definitions, for example see Champion (1994), Kelly (1987), Kraly and Gnanasekaran (1987), Poulain (1993), Raymer and Willekens (2008), United Nations (2002) and Willekens (1994, 1999). Several international institutes such as the *International Labour Organization, the Organisation for Economic Co-operation and Development, the United Nations* and *the European Commission* have all invested heavily in the harmonisation of international migration data, but without much success or progression (Bilsborrow *et al.*, 1997; Herm, 2006a and Fassmann, 2009). In fact, the situation today in terms of migration definitions and measurement is not much better than it was, say, 20 years ago.

Recently, some renewed efforts have been made to improve the migration data situation in Europe. In 2007, the European Parliament adopted a new regulation on migration statistics. This regulation provides clear definitions of immigration and emigration (Official Journal of the European Union, 2007), and lists the migration data that must be supplied to Eurostat, the statistical office of the European Union (EU), by Member States. However, this regulation leaves the Member States free to decide how they will provide these data, including the use of estimation methods (Fassmann, 2009). The methodology presented in this paper should help national statistical offices to improve and harmonize the data they currently provide to international organisations, such as Eurostat.

The migration definition set out in the 2007 Regulation corresponds to the definition recommended by the United Nations (1998), where an international migrant is defined as ‘a person who moves to a country other than that of his or her usual residence for a period of at least a year.’ One problem affecting the implementation of this definition is that some countries are unable to identify their nationals who have left (Fassmann, 2009). Furthermore, many European countries exclude the immigration of nationals from the published statistics, as they are not considered to be ‘migrants’. Another important obstacle has to do with the recommended duration of residence in the country of destination. It may take up to two years to identify all persons who have stayed at least one year, as they may arrive anytime during the annual time period of interest. This means that the publication of migration statistics based on the actual duration of stay may be delayed for some time. To provide statistics to the user community in a quicker fashion, many countries simply count those migrants who have stayed for at least three months, which leads to higher numbers than if the one-year criterion was applied. Other countries use the intended duration of stay as the criterion (Fassmann, 2009).

Many European countries do not have reliable statistics on emigration. This is mainly caused by the fact that migrants have little incentive to report their move to the administration of the country they have emigrated from. Moreover, it is difficult to count persons leaving the country because they are no longer present in the country collecting the data. In this situation, comparisons of sending country data with receiving country data provide important information on the degree of underestimation found in reported emigration flows (UNECE, 2009). In fact, the analysis of the so-called ‘double-entry matrix’ of migration flows produced by UNECE since the early 1970s, and more recently by Eurostat, has been found to be very useful and informative. Kelly (1987) and Poulain (1999), for example, have used the information contained in this matrix to assess the degree of harmonisation amongst reported data. In doing so, the possibility that very narrow or loose definitions of migration may be used for reported immigration statistics must be taken into account, which results in lower or higher levels of migration flows, respectively, in relation to, say, the United Nations’ recommended one-year definition (UNECE, 2009).

The aim of this chapter is to illustrate how reliable estimates of harmonized migration statistics may be obtained from a set of origin-destination flows, where two reported flows are available for each particular flow, *i.e.*, from receiving and sending countries. The new method that we present is based on earlier efforts by Poulain (1993, 1999), and is applied to reported flows between 19 European countries from 2002 and 2007. Note, however, that this chapter does not consider flows outside the 19 country system, or those that are missing. Raymer (2008) describes a method for estimating missing migration flow data.

2.2. Comparability of international migration data

The reliability of migration statistics can be measured by how well they correspond to a particular country’s definition or concept of migration. However, as definitions differ across countries, reliability does not guarantee comparability. Moreover, under-registration, under-coverage and accuracy of the collection system also affect the measurement of migration (Bilsborrow *et al.*, 1997 and Nowok *et al.*, 2006). First, there may be under-registration of migrants. This may be the case if the data depend on declarations by the migrants themselves. The willingness to report changes in places of residence varies both between countries and between groups of migrants. In general, migrants have more incentive to report their arrival than their departure, as there are usually direct benefits in doing so (*e.g.*, access to social services). Therefore, immigration statistics are generally considered more reliable than emigration statistics (Thierry *et al.*, 2005 and UNECE, 2009). Second, there may be under-coverage. This measurement category refers to the non-inclusion of particular migrant groups. Here, the differences are most often caused by the absence or inclusion of nationals, students, asylum seekers or irregular (illegal) migrants in the data. In general, asylum seekers are included only when they have been granted refugee status and received a temporary or permanent residence permit. However, in some instances, they are registered at an earlier stage of the asylum process. In other instances, even recognised refugees are not included. Irregular migrants are generally not included in migration statistics, as they are especially difficult to measure (for obvious reasons). In fact, Spain is the only EU country that includes irregular migrants in the official statistics. Finally, data based on sample surveys may be unreliable due to sampling errors. Furthermore, unless the sample size is very large, the data are likely to show irregularities in the patterns across ages or in the distribution of origins or destinations over time, as flows of migrants represent a relatively small proportion of the overall population being surveyed.

The main sources of the differences in the definitions used by EU countries to measure migration are the concepts of place of residence and duration of stay (Zlotnik, 1987; Bilsborrow *et al.*, 1997 and Kupiszewska and Nowok, 2008). The *de jure* (legal) approach to residence implies that in order to become a resident, a migrant must comply with certain regulations, which tend to differ between nationals and foreigners, and among foreigners, between EU- and non-EU-nationals. For example, it is not uncommon for emigrants to be registered in their country of citizenship (origin) even after several years of living abroad (Thierry *et al.*, 2005). Thus, having a place of residence does not necessary imply a presence in that country. The *de facto* (actual) approach is connected with physical presence in a country, usually for a specified minimum period of time. To prevent the delay caused by measuring actual duration of stay, most European countries use the intended duration of stay instead (Nowok *et al.*, 2006). Alternatively, the intended duration of stay may be used to provide provisional statistics, which are updated at a later point with the actual duration of stay statistics. Another group of countries measure 'permanent' change of residence only (*e.g.* Poland and Slovakia), which is very restrictive and tends to produce flow levels that are much lower relative to other definitions. The duration of stay criterion used by the majority of EU countries is between three months and one year. Only three countries (Cyprus, Sweden and UK) apply strictly the one-year criterion for immigration, as well as for emigration and for both nationals and non-nationals (Thierry *et al*., 2005). In fact, some countries do not take duration of stay into account at all. Germany is such an example, where everybody taking up a residence is counted as a migrant.

Because of differences in definition, coverage, registration and accuracy of the collection mechanism, the origin-destination matrix of migration flows between European countries based on immigration data reported by the countries of origin tends to differ from the matrix reported by the countries of destination. With respect to definitions, the differences are expected to be systematic over time. For example, the German definition is wider than the Dutch definition which, in turn, is wider than that of Sweden. In fact, Germany reports higher figures than the Netherlands, and the figures of the Netherlands are higher than those reported by Sweden (Kupiszewska and Nowok, 2008). A comparison of the size of these reported flows provides information on the effects of differences in definition on the size of migration flows (Bilsborrow *et al*., 1997 and UNECE, 2009). However, as mentioned above, not all differences can be explained by differences in definition. In some cases, countries report relatively large percentages of unknown countries of origin or destination. Furthermore, sudden jumps in observations may be caused by changes in definitions or by changes in the registration method.

Data on immigration and emigration flows by country of origin and destination are usually presented in an origin-destination matrix with off diagonal entries containing the number of people moving from any origin i to any destination j in a given calendar year. For this study, we have collected migration data for the 19 countries set out in *table 2.1*. As each flow can be reported by both sending and receiving countries, two migration tables may be produced. Such data are set out in *table 2.2*. Here, the average 2002-2007 values of migration between the 19 European countries set out in table 2.1 are presented. Table 2.2.a contains flows reported by the countries of destination and table 2.2.b contains the flows reported by the countries of origin. Clearly,

Table 2.1. List of European countries reporting both immigration flows by country of origin and emigration flows by country of destination, 2002-2007

Country	Abbreviation
Austria	AT
Cyprus	CY
Czech Republic	CZ
Germany	DE
Denmark	DK
Spain	ES
Finland	FI
Iceland	IS
Italy	IT
Lithuania	LT
Luxembourg	LU
Latvia	LV
Netherlands	NL
Norway	NO
Poland	PL
Sweden	SE
Slovenia	SI
Slovakia	SK
United Kingdom	UK

there are large differences between the two sets of reported numbers (see, *e.g.*, Spain to the United Kingdom or Poland to Germany).

2.3. Method

The differences between reported immigration and emigration numbers are useful for improving and harmonizing the migration data. If reported emigration numbers for a given country turn out to be systematically lower than the corresponding immigration numbers reported by the countries of destination, this suggests that the reported emigration numbers are too low. Adjusting these numbers in an upward direction moves them closer to the actual numbers. The same applies to reported immigration numbers. For each country we can estimate one adjustment factor for immigration and one for emigration in such a way that the adjusted immigration and emigration numbers are closer to each other than the reported numbers. To prevent arbitrary judgments biasing the results, we believe the estimation of adjustment factors for immigration and emigration flows should be estimated simultaneously. Moreover, it should be noted that immigration is not necessarily recorded more accurately than emigration. In some situations, sending country data may be considered better (Nowok *et al.*, 2006).

Poulain (1993, 1999) was the first to develop a method to adjust reported immigration and emigration numbers for the purpose of obtaining a consistent set of migration flows. 'Correction factors' were estimated by minimizing the sum of squares $\sum_{i,j}(\hat{\alpha}_j I_{ij} - \hat{\beta}_i E_{ij})^2$, where I_{ij} denotes migration from country i to country j reported by the receiving country j, E_{ij} denotes the same flow reported by the sending country i, α_j is the adjustment factor for all immigration to country j and β_i is the adjustment factor for all emigration from country i. Poulain and Dal (2008) refined this method by dividing the squared differences by the sum of the reported numbers, *i.e.*,

$$\sum_{i,j}(\hat{\alpha}_j I_{ij} - \hat{\beta}_i E_{ij})^2 / (I_{ij} + E_{ij}) \qquad (1)$$

This refinement prevents flows from (or to) large countries from biasing the estimates.

Various constraints have been tried by Poulain and colleagues (Abel, 2009). For instance, following the iterative approach to harmonizing migration flows suggested by Van der Erf and Van der Gaag (2007), Poulain and Dal (2008) proposed that the estimates should be normalized to Swedish immigration data, as they are generally considered to be highly reliable and in agreement with the UN recommended measure, as well as with the new EU regulation (Herm, 2006b). The parameters α_j and β_i may be estimated by solving a system of linear equations, which result from applying the method of Lagrange multipliers. Multiplying I_{ij} by $\hat{\alpha}_j$ and E_{ij} by $\hat{\beta}_i$ produces two sets of migration flow estimates from country i to country j. The final set of estimates are obtained by simply taking the average of the two, *i.e.*, $\hat{n}_{ij} = (\hat{\alpha}_j I_{ij} + \hat{\beta}_i E_{ij})/2$, where $\hat{n}_{ij}$ denotes the harmonised migration flows. Note, Poulain and Dal (2008) applied their correction method first to countries with relatively reliable data to prevent countries with less reliable data influencing the overall patterns. Here, the main concern is that the less reliable data have origin-destination patterns that are not consistent with the actual patterns. Thus, less reliable flows were adjusted in a hierarchical fashion, *i.e.*, by using the harmonised reliable data as a basis.

Table 2.2a. Reported migration by country of destination, averages 2002-2007

	To									
From	AT	CY	CZ	DE	DK	ES	FI	IS	IT	LT
AT		41	310	14 257	303	774	109	33	774	17
CY	22		13	276	23	25	23	1	30	3
CZ	1 316	118		9 218	262	833	56	42	672	24
DE	15 447	332	1 362		4 001	15 982	921	255	12 809	490
DK	203	25	46	2 687		964	365	1 413	265	85
ES	700	45	71	14 703	1 758		644	68	2 044	252
FI	270	21	38	2 173	414	844		45	235	43
IS	31	0	4	236	1 665	131	50		35	10
IT	1 608	49	254	22 196	986	9 320	250	74		82
LT	179	35	47	4 496	1 034	2 274	73	272	378	
LU	67	3	2	2 282	162	123	50	27	213	5
LV	83	104	13	2 155	457	300	87	93	183	175
NL	791	70	255	13 681	864	4 762	261	55	905	41
NO	98	14	24	1 378	3 148	1 696	845	364	167	87
PL	5 231	752	1 608	136 927	2 436	8 277	187	2 229	9 045	120
SE	489	88	67	3 348	3 313	1 826	3 502	492	379	91
SI	556	9	17	1 798	46	136	6	9	321	2
SK	3 192	432	14 064	11 148	149	788	22	45	690	4
UK	1 222	3 170	506	13 263	3 482	38 674	946	228	4 553	875
Total	31 504	5 306	18 702	256 221	24 502	87 725	8 397	5 741	33 695	2 407

From	To LU	LV	NL	NO	PL	SE	SI	SK	UK	Total
AT	8	9	559	111	180	307	100	208	1 395	19 496
CY	0	2	51	15	7	61	2	2	2 533	3 087
CZ	4	15	511	116	45	164	6	979	4 109	18 489
DE	454	166	9 182	2 268	2 876	3 374	299	446	19 039	89 701
DK	11	46	475	2 943	34	5 264	3	21	1 874	16 721
ES	24	18	3 101	768	119	1 300	8	36	14 581	40 239
FI	3	43	379	799	6	3 204	1	6	684	9 208
IS	0	6	75	373	11	462	1	2	417	3 509
IT	67	33	1 811	246	309	599	79	109	5 829	43 900
LT	1	236	302	926	43	574	0	5	2 507	13 380
LU		2	161	18	5	90	5	1	682	3 897
LV	2		125	233	6	264	0	5	1 227	5 511
NL	27	20		711	163	979	12	41	6 799	30 436
NO	2	24	453		48	5 098	1	24	1 667	15 135
PL	19	45	5 744	4 602		3 718	3	276	36 759	217 977
SE	14	54	696	4 917	113		15	20	3 213	22 635
SI	1	1	90	14	2	42		16	0	3 064
SK	4	4	465	238	18	110	4		4 584	35 961
UK	39	190	5 820	1 624	1 126	3 114	22	116		78 969
Total	682	913	30 000	20 921	5 111	28 723	559	2 311	107 897	671 315

Source: Eurostat.

Table 2.2b. Reported migration by country of origin, averages 2002-2007

	To									
From	AT	CY	CZ	DE	DK	ES	FI	IS	IT	LT
AT		18	937	6 665	166	429	231	27	1 022	111
CY	6		21	57	6	19	12	0	39	9
CZ	186	13		560	24	35	28	2	112	10
DE	17 787	271	8 104		3 095	16 807	2 371	287	31 235	2 455
DK	228	24	179	2 612		1 669	368	1 347	716	655
ES	155	9	57	2 686	157		110	9	1 163	120
FI	97	23	42	758	400	671		53	203	21
IS	13	2	17	205	1 800	59	48		105	64
IT	588	6	67	10 206	149	1 508	136	17		11
LT	48	8	54	1 269	158	628	87	23	204	
LU	31	3	13	911	99	79	35	19	175	4
LV	18	8	6	302	45	18	46	5	51	138
NL	616	50	298	10 493	533	3 774	322	54	1 278	54
NO	69	17	43	709	3 093	789	855	412	146	108
PL	538	15	63	14 417	111	341	20	46	505	6
SE	298	73	104	1 634	3 159	1 348	3 403	413	463	48
SI	311	3	14	589	5	27	4	1	186	1
SK	177	1	629	255	4	16	1	0	42	0
UK	1 593	4 060	2 692	12 579	1 932	33 431	682	103	5 270	1 074
Total	22 758	4 600	13 339	66 905	14 933	61 649	8 758	2 818	42 914	4 887

	To									
From	LU	LV	NL	NO	PL	SE	SI	SK	UK	Total
AT	45	42	426	87	2 401	388	402	1 778	901	16 076
CY	2	18	10	2	111	13	0	32	371	724
CZ	3	7	81	16	583	24	8	9 539	219	11 449
DE	1 686	1 494	9 293	2 122	100 827	3 974	2 004	9 456	17 233	230 499
DK	138	316	602	2 947	833	5 253	31	95	3 889	21 898
ES	87	19	869	159	398	203	10	45	3 430	9 684
FI	71	27	233	777	63	3 216	4	10	1 175	7 842
IS	37	29	49	482	872	478	25	56	232	4 570
IT	218	8	531	121	417	199	151	40	3 508	17 879
LT	18	163	116	199	122	233	3	5	2 638	5 975
LU		4	97	12	23	73	5	11	166	1 760
LV	6		20	34	26	67	1	2	196	987
NL	191	33		731	1 020	900	45	138	7 953	28 482
NO	23	69	287		281	5 083	5	61	1 395	13 444
PL	23	3	557	127		303	2	10	5 219	22 306
SE	127	62	522	4 746	354		27	29	3 905	20 713
SI	24	0	30	5	5	38		6	70	1 319
SK	2	0	13	3	15	8	3		69	1 235
UK	362	324	5 943	1 993	6 507	2 666	0	1 053		82 264
Total	3 062	2 619	19 676	14 561	114 854	23 117	2 724	22 364	52 567	499 105

Source: Eurostat.

There are several limitations in the model described above. First, the reported numbers included in the denominator of Equation (1) are known to be incorrect (Abel, 2009). Second, the row and column totals of the two estimated matrices are not equal. As a result, the row and column totals of the average harmonised migration matrix do not correspond to the row and column totals estimated using the adjustment factors. Finally, the method can only be applied to a limited set of countries with reasonably reliable data. This implies that the estimates of the adjustment factors depend on the selection of countries, which may not reflect the broader patterns of interest. For these reasons, we have revised Poulain's method in two important ways. First, the row-sums and column-sums of the two estimated matrices are set to be equal. Second, we introduce additional constraints on individual cells in the migration matrices, so that more countries (with less reliable data) may be included.

The adjustment factors for our method can be estimated by solving a system of linear equations and imposing a constraint. If we have a $N \times N$ receiving country and an equivalent $N \times N$ sending country matrix, the adjustment factors for receiving country, α_j and the adjustment factors for sending country data, β_i can be estimated by:

$$\sum_j \hat{\alpha}_j I_{ij} = \hat{\beta}_i \sum_j E_{ij} \quad \text{for } i = 1,\dots, N; \quad i \neq j \tag{2}$$

$$\hat{\alpha}_j \sum_i I_{ij} = \sum_i \hat{\beta}_i E_{ij} \quad \text{for } j = 1, \dots, N; \quad i \neq j \tag{3}$$

Equation (2) states that for each country the emigration total estimated on the basis of the adjusted matrix of flows reported by receiving countries equals the emigration total estimated on the basis of the adjusted matrix of flows reported by sending countries. Equation (3) does the same for immigration totals.

Equations (2) and (3) can be written as a homogeneous system of $2N$ linear equations with $2N$ unknowns, *i.e.*,

$$\hat{\alpha}_2 I_{12} + \hat{\alpha}_3 I_{13} + \ldots + \hat{\alpha}_N I_{1N} - \hat{\beta}_1 \sum_j E_{1j} = 0 \tag{4}$$

....

$$\hat{\alpha}_1 I_{N1} + \hat{\alpha}_2 I_{N2} + \ldots + \hat{\alpha}_{N-1} I_{NN-1} - \hat{\beta}_N \sum_j E_{Nj} = 0$$

$$\hat{\alpha}_1 \sum_i I_{i1} - \hat{\beta}_2 E_{21} - \hat{\beta}_3 E_{31} - \ldots - \hat{\beta}_N E_{N1} = 0$$

....

$$\hat{\alpha}_N \sum_i I_{iN} - \hat{\beta}_1 E_{N1} - \hat{\beta}_2 E_{N2} - \ldots - \hat{\beta}_{N-1} E_{NN-1} = 0$$

This system has an infinite number of solutions for α_j and β_i. For each set of values of $\hat{\alpha}_j$ and $\hat{\beta}_i$ that solve this system, $k\hat{\alpha}_j$ and $k\hat{\beta}_i$ are solutions as well. In order to find a unique solution one restriction needs to be imposed. In accordance with Poulain and Dal (2008), we assume that the adjustment factor for Swedish immigration is equal to one, since Sweden uses a definition of migration that is consistent with the new EU regulation and the quality of Swedish immigration data is considered to be adequate. This also means that the resulting estimates are harmonised in line with the new European regulation.

The basic assumption underlying our estimation procedure (as described above) is that the distributions of reported immigration by country of origin and reported emigration by country of destination correspond to the distribution of actual migration flows under the harmonised definition. This implies that the reported emigration of country *A* is *x%* higher or lower than the actual number (based on the standard definition) for all countries of destination. The same assumption applies to receiving country numbers. However, as we find in the next section, the estimated receiving country flows by country of origin and the estimated sending country flows by country of destination are not always consistent with each other. In a number of cases, specific origin-destination flows have to be considered separately. For that reason, we introduce additional constraints, corresponding to particular origin-destination flows that differ from the remaining flows.

Let us assume that the estimated receiving country migration flow from country p to q, $\hat{\alpha}_q I_{pq}$, differs substantially from the estimated sending country flow, $\hat{\beta}_p E_{pq}$. To make them consistent, we can multiply $\hat{\alpha}_q I_{pq}$

by $\hat{\gamma}_{pq}$ or $\hat{\beta}_p E_{pq}$ by $\hat{\delta}_{pq}$ so that both estimates of migration are equal. The question whether we should adjust the estimate based on the reported receiving country or the estimate based on the reported sending country depends on our knowledge of the data.

Given the estimated values $\hat{\alpha}_q$ and $\hat{\beta}_p$ and we can calculate the value of $\hat{\gamma}_{pq}$ easily from $\hat{\gamma}_{pq} = \hat{\beta}_p E_{pq} / \hat{\alpha}_q I_{pq}$ or the value of $\hat{\delta}_{pq}$ from $\hat{\delta}_{pq} = \hat{\alpha}_q I_{pq} / \hat{\beta}_p E_{pq}$

However, introducing $\hat{\gamma}_{pq}$ or $\hat{\delta}_{pq}$ changes the estimates of $\hat{\alpha}_q$ or $\hat{\beta}_p$. This also means that the row and column totals of both estimated migration matrices no longer tally. Therefore, we adjust the system of linear equations (2) and (3) by adding constraints on individual cells of the matrices. If we assume that the emigration number reported by country *p* needs to be adjusted, Equations (2) and (3) can be rewritten as:

$$\sum_j \hat{\alpha}_j I_{ij} = \hat{\beta}_i \sum_j E_{ij}(1+\hat{\delta}^*_{pq} D_{ij}) \text{ for } i = 1,\ldots,N; \quad i \neq j \tag{5}$$

$$\hat{\alpha}_j \sum_i I_{ij} = \sum_i \hat{\beta}_i E_{ij}(1+\hat{\delta}^*_{pq} D_{ij}) \quad \text{for } j = 1, \ldots, N; \quad i \neq j \tag{6}$$

where $D_{ij} = 1$ if $i = p$ and $j = q$, $D_{ij} = 0$ otherwise, and $\hat{\delta}^*_{pq} = \hat{\delta}_{pq} -1$.

The equations including I_{pq} and E_{pq} in the system of equations (4) can be rewritten as follows:

$$\hat{\alpha}_1 I_{p1} + \ldots + \hat{\alpha}_q I_{pq} + \ldots + \hat{\alpha}_N I_{pN} - \hat{\beta}_p E_{p1} - \ldots - \hat{\delta}_{pq}\hat{\beta}_p E_{pq} - \ldots - \hat{\beta}_p E_{pN} = 0 \tag{7}$$

$$\hat{\alpha}_q \sum_i I_{iq} - \hat{\beta}_1 E_{1q} - \ldots - \hat{\delta}_{pq}\hat{\beta}_p E_{pq} - \ldots - \hat{\beta}_N E_{Nq} = 0$$

In contrast with Equation (4), these are non-linear equations, because they include the term $\hat{\delta}_{pq}\hat{\beta}_p E_{pq}$. The values of the coefficients can be estimated by an iterative procedure. The model can be extended in a straightforward way to include additional constraints. However, for any particular country, the number of constraints should not be too high, as this reduces the available information to estimate α and β.

2.4. Data

The sending and receiving country migration data have been provided by the national statistical institutes of the EU Member States in response to annual rounds of data collection conducted jointly by five international organizations and coordinated by Eurostat (Kupiszewska and Nowok, 2008). As concerns Europe, Eurostat processes and disseminates data received from 37 countries on their website (epp.eurostat.ec.europa.eu). Data sources used by EU member states to produce migration statistics are very diverse (Kupiszewska and Nowok, 2008 and Nowok *et al.* 2006). The major types of sources are population registration systems, statistical forms, other administrative registers related to foreigners (such as alien registers, residence permit registers and registers of asylum seekers), sample surveys and censuses. Thirteen EU countries use a population register as the source of migration statistics. Alien registers and residence permit registers are used in seven countries, sometimes in addition to population registers. These registers only provide information on the migration of non-nationals. Cyprus and the UK rely on passenger surveys conducted at the borders, while Portugal and Ireland rely on household surveys. Greece, France and Portugal do not have any data on migration by nationals. Some countries derive their emigration statistics from data on residence permits by assuming a migrant has left the country when a residence permit has expired. Moreover, they often assume that the country of next residence is the country of their citizenship. The result, we believe, is an overestimation of actual emigration to those particular countries. Finally, several countries include in their so-called ‘administrative corrections’ emigration that has not been declared, which cannot be disaggregated by country of next residence.

Data on immigration by country of previous residence or emigration by country of next residence are not always available or complete (Nowok *et al.*, 2006). Thus the sending country and receiving country matrices, when combined into a double-entry matrix may be incomplete. For some countries, a large share of emigrants have an unknown country of destination: Around 75 percent in Slovenia, 40 percent in Luxembourg, 35 percent in Austria, 31 percent in the Netherlands and 39 percent in Spain, for example. Fortunately, the estimation of adjustment factors takes this into account.

In the next section we present our harmonised estimates of migration between 19 European countries that provide data on both immigration by country of origin and emigration by country of destination for the calendar years 2002-2007. The reported data contains both nationals and non-nationals.

Table 2.1 provides a list of the countries. Although there are some data for Ireland, Portugal and Romania, these have not been used because they cover only a part of the migration flows (*e.g.* only foreigners or nationals). For Iceland, Italy, and Luxembourg, data for one or more years in the period 2002-2007 are missing. For these countries, the adjustment factors are estimated for averages over the available years.

2.5. Results

The results presented in this section are obtained by applying the estimation method described in section 2.3. Table 2.2a shows the average values of migration between 19 European countries reported by receiving countries for the years 2002-2007 and table 2.2b shows the corresponding numbers reported by the sending countries. The countries listed in the row headings refer to origins and those listed in the column headings refer to destinations. A comparison of tables 2.2a and 2.2b reveals large differences between numbers reported by sending and receiving countries. According to the numbers reported by receiving countries, 671,315 migrants per year moved between these 19 countries, whereas the numbers reported by sending countries total 499,105. For 11 countries, the reported receiving country immigration totals are higher than the corresponding sending country totals. For example, Germany reported that 256,221 immigrants arrived from the 18 countries in this study, whereas these countries reported that only 66,905 emigrants moved to Germany. Poland reported that 22,306 persons emigrated to the other 18 countries which, for their part, reported receiving 217,977 immigrants from Poland, suggesting that Polish emigration data are around 10 times too low. For 15 of the 19 countries, the emigration total reported by the sending country is lower than the corresponding totals reported by receiving countries. Keep in mind that receiving country data should not always be considered better than sending country data. Consider, for example, the flows from Poland to Germany in tables 2.2a and 2.2b. Here, Germany received an average of 136,927 migrants from Poland, whereas Poland reported that they only sent an average of 14,417. This difference could be explained by the duration criteria used by these countries, with Germany having a very loose definition (instant) and Poland having a very restrictive definition (permanent). So, in comparison with the harmonised definition of a one year period, Germany's reported number is considered too high and Poland's too low.

The estimated adjustment factors are set out in *table 2.3*. We indicated above that in order to estimate the adjustment factors a restriction was introduced, *i.e.* the adjustment factor for Swedish immigration is set equal to one. For 16 of the 19 countries, the E_{ij} adjustment factor exceeds one, indicating that sending country numbers tend to be underestimated. However, table 2.3 also shows that I_{ij} numbers seem to be underestimated in the majority of countries as well. This may seem contradictory since for 11 of the 19 countries the reported immigration totals exceed the corresponding emigration numbers reported by the sending countries. This is because the reported receiving country numbers should be compared with the adjusted sending country numbers rather than the reported numbers. For example, the immigration total reported by the UK (107,897) exceeds the reported emigration from sending countries to the UK (52,567). The reported emigration to the UK includes 5,219 emigrants from Poland to the UK. However, since the reported emigration from Poland is too low (the adjustment factor equals 10.46, see table 2.3) the reported emigration from Poland to the UK is

Table 2.3. Estimates of adjustment factors for immigration and emigration, 2002-2007

	Immigration	Emigration
Austria	1.06	1.74
Cyprus	1.06	5.29
Czech Republic	2.14	3.33
Germany	1.03	0.69
Denmark	0.74	0.80
Spain	0.82	4.90
Finland	1.26	1.22
Iceland	0.57	0.74
Italy	1.42	2.92
Lithuania	2.33	2.45
Luxembourg	5.65	2.43
Latvia	2.92	6.22
Netherlands	0.97	1.25
Norway	0.84	1.19
Poland	17.85	10.64
Sweden	1.00	1.21
Slovenia	5.18	2.71
Slovakia	18.90	43.69
United Kingdom	1.21	1.18

adjusted from 5,219 to 55,506. Moreover, the adjustment factor for Spanish emigration data equals 4.90, so the reported emigration from Spain to UK is adjusted from 3,430 to 16,792. For several other countries, emigration to the UK is adjusted upwards as well. As a consequence, the adjusted emigration numbers to the UK exceed the total of immigration reported by the UK and thus the reported immigration is adjusted upwards as well. Note that the adjustment factors for immigration for most countries are closer to one than the adjustment factors for emigration, which indicates that the reported immigration numbers are more accurate than the emigration numbers.

Multiplying the reported numbers in table 2.2a by the adjustment factors for receiving country data and the reported numbers in table 2.2b by the adjustment factors for sending country data results in two tables for which the row and column totals are equal (not presented here for space reasons). The differences between the cells in these two matrices are considerably smaller than those in table 2.2. In fact, the root mean squared error (RMSE) is reduced from 8,966 to 2,131. In other words, the differences between the two reported migration flow tables are reduced by 77 percent. However, we still found some substantial differences in the two estimated migration flow tables. For example, the migration from Poland to Germany estimated on the basis of German immigration data equals 141,035, whereas the estimate based on Polish emigration data is equal to 153,399. These differences reflect the fact that the distribution of reported Polish emigration by country of destination is not consistent with the share of immigration from Poland in the total reported immigration numbers of other countries. As a result, the estimate of the migration flow from Poland to Germany based on Polish data exceeds that based on German data, whereas for most other countries, the adjusted Polish emigration numbers are lower than the corresponding adjusted immigration numbers. This means that one substantial inconsistency in the estimates is likely to influence the estimates of other migration flows. To prevent such inconsistencies from affecting the overall estimates, we have added constraints to individual cells (flows) in the model.

The introduction of constraints to individual cells in the matrix allows us to consider special cases, such as the Poland to Germany flow described above. In total, we found six migration flows where the estimates differed by more than 10,000. Specifically, these flows were Poland to Germany, Poland to UK, Germany to Poland, Germany to UK, Czech Republic to Slovakia and UK to Poland. After identifying the flows with large differences, we then had to decide whether the constraint should be applied to the numbers of the receiving country or of the sending country. Since we believe that reported

Table 2.4. Estimates of adjustment factors for immigration and emigration, 2002-2007, including six additional constraints on individual flows

	Immigration	Emigration
Austria	1.17	1.35
Cyprus	0.88	4.71
Czech Republic	1.97	8.92
Germany	0.81	0.71
Denmark	0.72	0.74
Spain	0.73	4.32
Finland	1.18	1.12
Iceland	0.59	0.69
Italy	1.48	2.44
Lithuania	2.16	2.15
Luxembourg	5.45	2.08
Latvia	2.78	5.44
Netherlands	1.04	1.06
Norway	0.81	1.10
Poland	14.25	18.31
Sweden	1.00	1.10
Slovenia	4.90	2.33
Slovakia	8.34	39.40
United Kingdom	1.09	0.91
Coefficients for additional constraints (Lagrange multipliers)		
Immigration to Poland from Germany	1.74	
Immigration to Poland from the UK	0.37	
Emigration from Poland to Germany		0.42
Emigration from Poland to the UK		0.42
Emigration from Germany to the UK		1.70
Emigration from the Czech Republic to Slovakia		0.10

emigration numbers are generally considered to be less reliable than reported immigration numbers, we apply the constraints to the sending country data, except for the Germany to Poland and UK to Poland flows (*i.e.*, Poland's immigration data is considered to be of lower quality than the corresponding emigration data reported by both Germany and the UK).

The adjustment factors taking into account the six constraints on individual flows are set out in *table 2.4*. The coefficients (Lagrange multipliers) for the Poland to Germany and Poland to UK flows are both equal to 0.42. This raises the adjustment factor for emigration from Poland from 10.64 (table 2.3.) to 18.31 (table 2.4), while at the same time, the adjustment factor for Polish emigration to Germany and the UK falls to 7.69 (i.e., 18.31 x 0.42). For Polish immigration, the adjustment factor becomes smaller. The high adjustment factor for Polish receiving data was mainly a consequence of the big difference between the two figures for migration from Germany to Poland. Including a constraint for this flow raises the adjustment factor for Poland's reported flow from Germany by a factor of 1.74 (i.e., the adjustment factor of 14.25 is multiplied by 1.74 to get 24.80). In contrast, the adjustment factor for Poland's reported flow from the UK falls to 5.27 (i.e., 14.25 x 0.37). For the Czech Republic, the reported emigration numbers are considerably lower than the corresponding reported immigration numbers with one big exception: the number of emigrants reported to Slovakia is relatively large. Clearly, the emigration flows from the Czech Republic to all other countries need to be adjusted by a different factor than the emigration flow to Slovakia.

The adjustment factors in table 2.4 illustrate how substantial improvements in the estimated adjustment factors can be made by introducing constraints on specific 'problem' flows in the matrix. For example, the inclusion of a constraint for the migration flow from the Czech Republic to Slovakia lowered the adjustment factor for Slovakia's receiving migration data from 18.90 to 8.34. Another example is German's receiving data. Here, the adjustment factor is reduced from 1.03 to 0.81. This is mainly explained by the reduction of the estimate of Polish emigration to Germany. Since Germany has a wide definition of migration, one would expect the adjustment factor to be below one. Thus the adjustment factors in table 2.4 appear more plausible than those set out in table 2.3.

The harmonised migration tables that used the additional constraints are set out in *tables 2.5a* and 2.*5b*. The introduction of these constraints led to a further strong reduction in the differences between both tables, as indicated by the RMSE, which fell from 2,131 to 952 or by a further 54 percent. To obtain a final single set of harmonised flows, we believe it is better to rely on table 2.5a than on table 2.5b. This table gives more weight to the receiving country data, which we consider more reliable. Poulain, on the other hand, advocated taking the average of the two estimated matrices. This approach implies that the origin-destination patterns in the reported sending country data are as reliable as those in the reported receiving country data.

The average adjustment factors estimated for the period 2002-2007 (table 2.4) can be applied to the annual reported migration data to create a time series of harmonised flows. In *figure 2.1*, the estimated total immigration and emigration flows for Germany from and to the other 18 countries in this study are compared. As expected, the estimated numbers are lower than the reported numbers because the definition for Germany is much wider than the harmonised definition. The figure also shows that estimated emigration increases more gradually over time than the reported numbers. In *figure 2.2*, the immigration and emigration flows for the UK are presented. Here, the average levels of the reported and estimated numbers do not differ much, but the estimated flows show a more gradual pattern over time than the reported flows. One reason for the sharp fluctuations in the reported numbers is that they are based on sample surveys.

2.6. Discussion

The aim of this chapter has been to obtain a reasonable and consistent set of international migration statistics. For this purpose we have developed a model using statistical information from different countries. The method is based on an idea originally proposed by Poulain (1993, 1995). Our method differs from his in three important ways. First, we have estimated a set of adjustment factors for receiving and sending country data in a way that

Figure 2.1. Reported and estimated immigration from and emigration to 18 European countries, Germany, 2002-2007 (x 1,000)

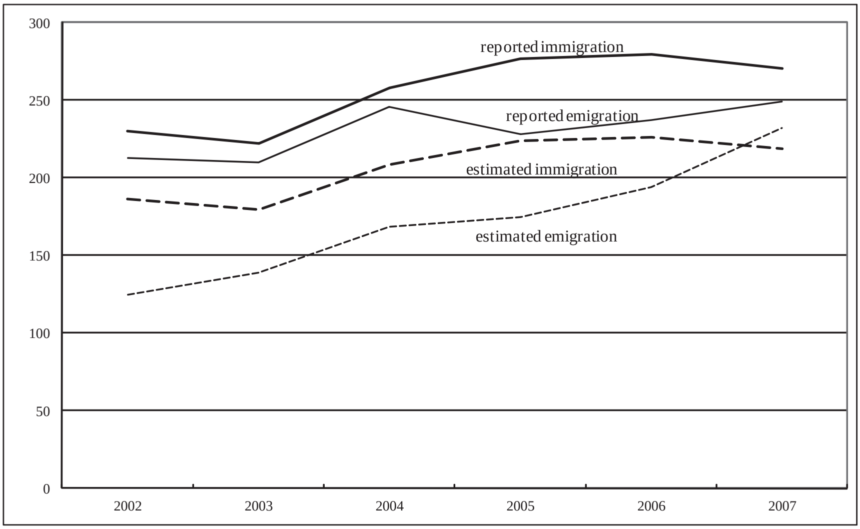


Table 2.5a. Estimated migration by country of origin and destination, including constraints on six individual flows, 2002/2007, based on numbers reported by receiving countries

	To									
From	AT	CY	CZ	DE	DK	ES	FI	IS	IT	LT
AT		36	610	11 526	217	563	129	20	1 144	37
CY	26		25	223	16	18	28	1	44	7
CZ	1 547	103		7 453	187	606	66	25	992	51
DE	18 148	291	2 679		2 864	11 631	1 092	152	18 920	1 057
DK	238	21	90	2 172		701	432	840	391	183
ES	822	39	139	11 887	1 259		763	40	3 020	544
FI	317	19	75	1 757	296	614		27	347	92
IS	37	0	8	191	1 192	95	59		51	22
IT	1 889	43	500	17 945	706	6 782	297	44		177
LT	211	31	92	3 635	741	1 655	86	162	558	
LU	79	2	5	1 845	116	89	59	16	315	11
LV	97	91	26	1 742	327	218	103	55	270	378
NL	930	61	502	11 060	619	3 465	309	33	1 336	89
NO	115	12	47	1 114	2 254	1 234	1 001	216	246	187
PL	6 146	659	3 163	110 701	1 744	6 024	221	1 324	13 361	259
SE	574	77	132	2 706	2 372	1 328	4 150	292	560	197
SI	653	8	32	1 454	33	99	7	5	473	4
SK	3 750	379	27 658	9 012	107	574	26	26	1 020	9
UK	1 436	2 780	996	10 722	2 493	28 145	1 121	135	6 725	1 886
Total	37 015	4 654	36 780	207 145	17 542	63 841	9 950	3 411	49 772	5 190

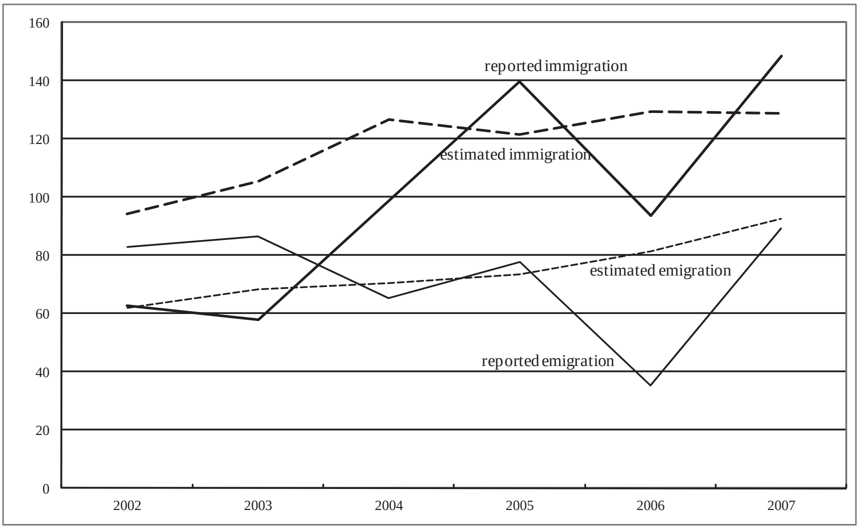
	To									
From	LU	LV	NL	NO	PL	SE	SI	SK	UK	Total
AT	44	26	584	90	2 570	307	488	1 730	1 522	21 647
CY	1	4	53	12	100	61	8	18	2 764	3 408
CZ	23	43	534	94	644	164	28	8 158	4 482	25 200
DE	2 475	462	9 586	1 843	71 514	3 374	1 462	3 715	20 770	172 034
DK	60	128	496	2 391	480	5 264	15	176	2 044	16 124
ES	132	49	3 237	624	1 698	1 300	41	297	15 907	41 798
FI	18	119	395	649	88	3 204	4	51	746	8 819
IS	2	17	78	303	162	462	6	18	455	3 157
IT	365	93	1 891	200	4 402	599	386	906	6 359	43 582
LT	7	656	316	752	610	574	1	39	2 734	12 858
LU		4	168	15	69	90	24	10	744	3 661
LV	9		131	189	90	264	0	38	1 338	5 367
NL	146	55		577	2 323	979	59	343	7 417	30 305
NO	12	68	473		679	5 098	2	201	1 818	14 778
PL	101	125	5 997	3 739		3 718	12	2 298	40 102	199 695
SE	74	149	727	3 996	1 608		74	168	3 505	22 690
SI	6	4	94	12	24	42		129	0	3 079
SK	24	12	486	193	254	110	20		5 001	48 662
UK	213	528	6 077	1 320	5 907	3 114	105	968		74 670
Total	3 715	2 540	31 321	17 000	93 222	28 723	2 735	19 264	117 708	751 530

Table 2.5b. Estimated migration by country of origin and destination, including constraints on six individual flows, 2002/2007, based on numbers reported by sending countries

	To									
From	AT	CY	CZ	DE	DK	ES	FI	IS	IT	LT
AT		24	1 261	8 973	223	578	311	36	1 376	150
CY	26		99	266	26	87	56	0	181	42
CZ	1 658	116		4 999	213	309	253	21	998	89
DE	12 616	192	5 748		2 195	11 921	1 681	203	22 154	1 741
DK	168	17	132	1 924		1 229	271	991	527	482
ES	669	37	247	11 592	678		475	37	5 020	516
FI	109	25	47	852	449	754		60	228	24
IS	9	1	11	142	1 243	41	33		73	44
IT	1 433	15	164	24 877	364	3 675	331	41		26
LT	102	16	115	2 730	339	1 352	187	50	438	
LU	65	5	26	1 895	205	165	73	40	364	8
LV	98	44	30	1 640	245	100	250	27	277	748
NL	656	53	317	11 165	567	4 016	342	57	1 360	57
NO	76	18	47	780	3 399	867	940	453	160	118
PL	9 857	272	1 160	110 701	2 023	6 244	369	845	9 247	116
SE	327	80	114	1 789	3 460	1 477	3 728	453	507	52
SI	725	7	33	1 374	12	63	10	2	435	2
SK	6 974	46	24 784	10 028	144	617	20	0	1 642	0
UK	1 446	3 685	2 444	11 418	1 753	30 345	619	93	4 784	975
Total	37 015	4 654	36 780	207 145	17 542	63 841	9 950	3 411	49 772	5 190

	To									
From	LU	LV	NL	NO	PL	SE	SI	SK	UK	Total
AT	61	57	573	118	3 232	523	541	2 393	1 213	21 642
CY	10	86	45	11	520	59	0	148	1 745	3 408
CZ	25	58	721	140	5 203	213	70	8 158	1 956	25 200
DE	1 195	1 060	6 591	1 505	71 514	2 818	1 421	6 707	20 770	172 034
DK	101	233	443	2 170	613	3 868	22	70	2 863	16 124
ES	377	83	3 749	686	1 718	875	42	195	14 802	41 798
FI	80	30	262	873	70	3 617	5	11	1 321	8 819
IS	25	20	34	333	602	330	17	39	160	3 157
IT	531	20	1 294	295	1 015	484	368	96	8 551	43 582
LT	39	351	250	428	263	502	7	11	5 676	12 858
LU		9	201	26	47	152	10	22	345	3 661
LV	33		109	182	141	362	4	11	1 068	5 367
NL	203	35		777	1 085	958	48	146	8 462	30 305
NO	26	75	316		309	5 587	5	67	1 534	14 778
PL	412	61	10 205	2 319		5 539	34	189	40 102	199 695
SE	139	68	571	5 199	388		29	31	4 277	22 690
SI	57	1	70	11	11	88		13	164	3 079
SK	72	0	493	118	584	328	112		2 699	48 662
UK	328	294	5 395	1 809	5 907	2 420	0	956		74 670
Total	3 715	2 540	31 321	17 000	93 222	28 723	2 735	19 264	117 708	751 530

Figure 2.2. Reported and estimated immigration from and emigration to 18 European countries, United Kingdom, 2002-2007 (x 1,000)



ensures consistency in the two sets of marginal totals. Second, we have introduced additional constraints on special origin-destination cases where the average adjustment factors do not apply. This allows us to include countries with less reliable data in our analysis. Third, instead of calculating the arithmetic average of the two estimated matrices, we believe it is better to use the matrix giving more weight to the reported immigration numbers (*i.e.* table 2.5a). In this way we take advantage of the fact that the information on countries of origin in receiving country data tend to be more reliable than the country of destination information in sending country data. Finally, our estimates are consistent with the harmonised migration definition based on an (intended) minimum duration of stay of 12 months.

Due to differences in definition, coverage and registration, the origin-destination matrix of migration flows between European countries based on receiving country data tends to differ from the matrix based on sending country data. Germany has a wide definition of migration, as it does not include a time constraint and thus the reported number may well include short term migrants. In contrast, Poland has a very narrow definition of migration and, as a consequence, the reported numbers are very low. By comparing corresponding reported immigration and emigration flows for 19 European countries, we have assessed to what extent German migration statistics are higher than they would be under a harmonised definition and to what extent Polish migration statistics are lower.

However, the large differences between European countries cannot be explained by differences in definitions alone. First, these differences cannot explain why emigration flows are more likely to be underestimated than immigration flows. Second, whereas eleven countries employ a duration limit that is shorter than that of the harmonised definition (Kupiszewska and Wisniowski, 2009), only five of these countries have an adjustment factor of immigration below one. The other six countries with durations of six months or shorter have adjustment factors for immigration greater than one. These include Austria, Czech Republic, Italy, Luxembourg, the Netherlands and Slovenia. Thus, to an important extent, the differences must also be caused by problems of coverage. This is confirmed by a study comparing migration statistics between Sweden, Denmark and Belgium which suggests that less than 25 percent of differences are due to differences in the duration criterion (Nowok *et al*., 2006). The effects of differences in definition and coverage may offset each other to some extent. One would expect the under-registration of short term migrants to exceed that of long-term migrants. A wide definition of migration (*i.e.* a short duration of stay) would lead to a higher reported number of migrants than would be expected on the basis of the harmonised definition. Under-registration, however, would lead to a smaller number. This may explain why the adjustment factors for Germany are not as low as one might expect from applying the very wide definition.

The main reason for the relatively low numbers reported by sending countries is that emigrants do not have strong incentives to report leaving a country. In particular, this applies to EU citizens who can live in another EU country without asking for a residence permit. One solution might be to introduce a removal card system (Nowok *et al*., 2006). Here, any person leaving country *A* would be required to fill in a form to be given to the authorities in country *B* at arrival. After country *B* has determined whether or not the person is an international migrant under a harmonised definition, it would then inform country *A* of the arrival. The Nordic countries have such a system and their immigration and emigration statistics are mutually consistent (Herm, 2006a). However, policy makers tend to be more interested in migrants from outside Europe and asylum seekers than intra-European migrants, and therefore such a system is not likely to have a high priority in the future. As long as such a system is lacking, cross-country comparability of migration statistics can only be achieved by comparing statistics from different countries. To the extent that the differences between countries are caused by differences in definitions and coverage, the differences may be expected to remain systematic over time. The method developed in this paper aims to assess the size of these systematic differences. Table 2.4 shows that for 10 out of the 19 countries in this study, the adjustment factor for sending country data exceeds two, meaning that reported emigration numbers are underestimated by more than 50 percent in relation to the one-year duration definition. As a consequence, reported net migration totals may be overstated.

In addition to 'correcting' the reported receiving and sending country migration data for differences in definition and coverage, our method contributes to producing estimates that tend to fluctuate less strongly over time. One clear example concerns the UK. Since the UK uses a general purpose passenger survey, the reported flows fluctuate considerably over time. Moreover, flows to some (smaller) countries may not be observed in some years. We believe our method produces more stable estimates of migration flows for the UK (and other countries relying on sample data). Interestingly, the estimated adjustment factors for the UK are close to one. This implies that the sample survey used for estimating migration to and from the UK provides a reasonably reliable estimate of total migration flows on average, but that the annual estimates are affected by sizeable random fluctuations.

The adjustment factors shown in table 2.4 can be used to adjust migration numbers to and from countries not included in the matrix, so that total immigration and emigration numbers and total net migration can be estimated for the 19 countries in this study. Before doing so, one first has to make sure that the share of unknowns in the migration statistics can be distributed evenly across all origins or destinations. If so, the adjustment factors will take this into account. Thus, for estimating total immigration and emigration numbers, the adjustment factors should be applied to total migration numbers excluding unknowns.

The matrix may be extended to include flows with missing data. Raymer (2008) developed a two-step estimation method for countries with missing data (see also De Beer *et al.*, 2009 and Raymer and Abel, 2008). The first step estimates missing immigration and emigration totals based on harmonised migration flows and covariate information. The second step uses the origin-destination interaction patterns of the harmonised migration flows and covariate information to estimate the missing interaction patterns. This estimation step takes into account the fact that migration is relatively high, for example, between neighbouring countries and countries belonging to a similar language group.

Finally, work is currently being carried out to integrate harmonisation and estimation of missing data into a single (Bayesian) model that also includes measures of uncertainty and expert judgements. The Integrated Modelling of European Migration (IMEM) project recently funded by New Opportunities for Research Funding Agency Co-operation in Europe (NORFACE) is expected to develop such a model (see http://www.norface.org/migration12.html) over the next couple of years. We hope this study will provide an important foundation for work such as this, and other projects aiming to improve our knowledge and understanding of the complexity of international migration.

3. Forecasting international migration: Time series projections versus argument-based forecasts

Abstract
Forecasts of immigration and emigration can be based on extrapolations of changes observed in the past. Extrapolations can be based on different time series models, ranging from simple linear trends to stochastic time series models. Extrapolations of immigration, emigration and net migration for the Netherlands show that different methods can lead to very different outcomes. Thus it is useful to examine the explanations behind the changes in past migration, which can then be used to determine future changes. Different types of migration, such as labour migration, family migration, and asylum migration, are affected by various factors. Thus for assumptions about future changes in migration it is useful to distinguish the main types of migration.

3.1. Introduction

Since the 1980s, immigration to most European countries has increased substantially. As a consequence, migration has become the main source of population growth in Europe and, therefore, assumptions on the future size of migration are an important input to population projections. Howe and Jackson (2005) argue that "most official immigration projections (…) are based on little theory and virtually no definable methodology." The present chapter aims to demonstrate how migration projections may be improved by distinguishing different migration flows. The illustrations come from Dutch data, for which detailed migration data are available over time.

Most national statistical institutes and international organizations, such as the European Union (EU) and the United Nations (UN), use the cohort-component model for making population projections. Typically, these organisations only incorporate assumptions on the future size of *net migration* by age and sex into their models. Those projections that incorporate the separate flows of immigration and emigration are preferred because they can be readily associated with predictive variables and can be analysed according to different types of flows. Also, with some time lag, foreign emigration can be linked with foreign immigration. The same is possible for emigration and immigration flows of nationals. Moreover, immigration tends to be positively related with the business cycle and emigration negatively. Thus, there are many advantages in making separate assumptions on future changes in immigration and emigration, which is more difficult with net migration.

The availability of accurate data on immigration and emigration is a problem in many countries. For some countries, information on net migration may only be available. As immigration tends to fluctuate more strongly than emigration, changes in net migration are mainly due to changes in immigration. Hence, assumptions on changes in net migration tend to be mainly based on assumptions on the future direction of immigration. However, in making assumptions on future changes in net migration, one should take into account that after a period of increase in immigration, there may be an increase in emigration and, thus, a reduction in net migration. The age structure of emigration may also differ from that of immigration. This implies that if the size of emigration changes in a different way than that of immigration, the age structure of net migration will change.

Projections of immigration and emigration can be based on extrapolations of changes observed in the past. Extrapolations may be based on different time series methods, ranging from simple linear trends to sophisticated stochastic time series models. One special case is to assume that net migration is zero. If net migration fluctuated around zero in the past, this may well be a valid assumption. However, even if total net migration is zero, this usually does not apply to separate age groups. Since emigrants are older than immigrants on average, net migration for young people will be positive and for older people negative. Thus assuming net migration to be zero for all ages may lead to some bias in the population projections.

The use of time series models for projecting immigration, emigration and net migration on the basis of Dutch data is illustrated in section 3.2. As migration tends to fluctuate rather strongly, different extrapolation methods may produce a wide range of projected outcomes. For making forecasts of future changes in migration, one has to decide to what extent past changes found in immigration and emigration patterns will continue in the future. This requires an identification of the main factors explaining past changes in migration flows. A discussion of how forecasts of migration may be based on explanations or migration theories is presented in section 3.3.

Different types of immigration and emigration are affected by various factors. For example, while labour migration is primarily affected by the situation in the labour market, marriage migration is affected by the size and composition of the resident migrant population and asylum migration is affected by asylum policies. For assumptions made on future changes in migration, it is useful to distinguish migrants according to their primary motives, such as labour, family or asylum. This allows one to make argument-based forecasts, rather than simply extrapolating changes observed in the past. As Howe and Jackson (2005) point out, one important benefit of explanatory arguments of forecasts is that these can be objectively evaluated and tested against historical evidence. In many countries, data on different categories of migrants are lacking. In these situations, the population characteristics of age, sex and country of origin may be used as proxies. The main factors affecting the future size of different types of immigration and emigration are discussed in sections 3.4 and 3.5, respectively. A discussion of how migration categories can be used for making argument-based forecasts is included in section 3.6.

As time series of net migration tend to exhibit large fluctuations, projections of migration are rather uncertain, even in the short run. The degree of uncertainty of migration forecasts can be assessed on the basis of historic forecast errors or on stochastic time series models. However, different types of immigration and emigration may be assumed to change in different ways in the future. For example, labour migration may be assumed to increase due to the ageing of the labour force, whereas asylum migration may be assumed to decrease due to more strict policies. This raises the question to what extent past developments in migration provide sufficient basis for assessing the uncertainty of future migration. Therefore, the argument-based approach set out in section 3.7 may be used for both making assumptions about future changes in migration flows, as well as, the degrees of uncertainty.

3.2. Extrapolations

In this section, the applications of different extrapolation methods are illustrated with data from the Netherlands. The data represent annual immigration, emigration and net migration totals for the period 1950-2004, which were obtained from the Statistics Netherlands StatLine data base (www.cbs.nl). Extrapolations of these data can be made by applying time-series models, which include both deterministic and stochastic models. A well-known example of a deterministic model is fitting a straight line to a time series of data. Deterministic models are based on the assumption that there is a fixed trend. Random fluctuations do not affect this trend. Stochastic time-series models are based on the assumption that the direction of the trend of the time series is subject to random changes. For this, the ARIMA-models introduced by Box and Jenkins (1970) are widely applied.

Linear trends for immigration and emigration in the Netherlands were estimated from 1950 to 2010 (see *figure 3.1*). The predicted 2004 values differs substantially from the observed values. Immigration in 2004 dropped sharply below the trend, whereas emigration was well above the trend. Accepting the long-run linear trend implies that the decrease in immigration or the increase in emigration in 2003 and 2004 are assumed to be temporarily occurrences.

Stochastic time-series models focus on the short run. ARIMA-models are identified on the basis of autocorrelation coefficients. These indicate the correlation of a time series with the same time series lagged 1 or 2 or more years. The autocorrelations for immigration, emigration and net migration for both the time series of observations and the time series of first differences are shown in *table 3.1*.

For immigration and emigration, the patterns of the autocorrelation coefficients suggest two models: A first-order autoregressive model and a random walk model. The parameters of the first-order autoregressive model estimated for immigration are:

Figure 3.1. Migration from and to the Netherlands, 1950-2010: Observations, linear trends and ARIMA

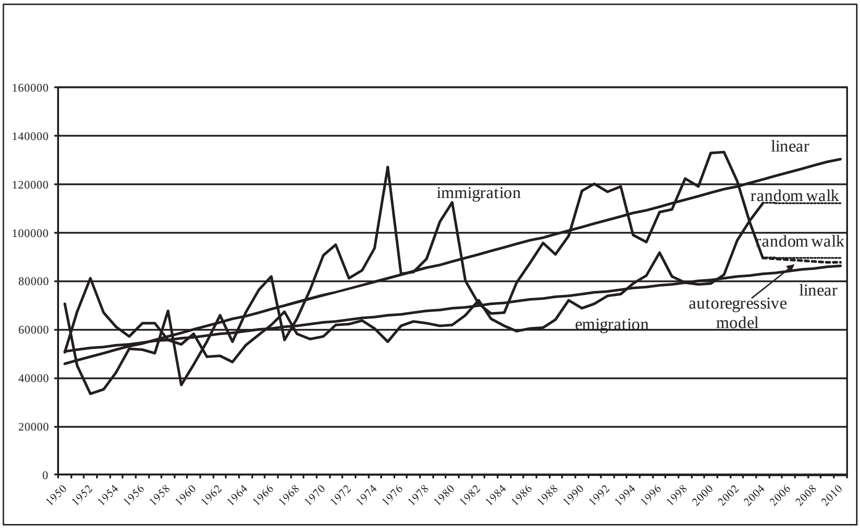


Table 3.1. Autocorrelation coefficients for the Netherlands immigration, emigration and net migration data, 1950-2004

	Immigration		Emigration		Net migration
Lag	Levels	1^{st} Difference	Levels	1^{st} Difference	Levels
1	0.86	-0.07	0.89	0.13	0.72
2	0.74	-0.15	0.75	-0.17	0.43
3	0.68	-0.25	0.68	-0.22	0.30
4	0.69	0.06	0.66	-0.08	0.32
5	0.69	0.16	0.67	0.02	0.36
6	0.61	-0.15	0.66	-0.05	0.27
7	0.58	-0.11	0.66	0.06	0.29
8	0.58	0.00	0.63	0.09	0.32
9	0.58	0.07	0.53	-0.09	0.26
10	0.56	0.09	0.46	-0.04	0.23

$$IM_t = \underset{(.07)}{.86}\, IM_{t-1} + \underset{(6258)}{12363} + e_t, \qquad (1)$$

where IM_t = immigration in year t and e_t = random term, which is serially uncorrelated and has an expected value of zero. The numbers between parentheses are the standard errors. This model implies that, in the long run, the projection tends to a level of 86445 (*i.e.*, 12363 / (1 - 0.86)). The random walk model for immigration is specified as:

$$IM_t - IM_{t-1} = c + e_t, \qquad (2)$$

where the constant term c (usually labelled as ‘drift’) does not differ significantly from zero. If c is excluded from the model, the projections are equal to the last observed value, *i.e.*, $IM_{t+1} = IM_t$ (because the expected value of e_{t+1} is zero). The autoregressive model and the random walk projections for immigration differ only slightly, as illustrated in figure 3.1.

For emigration, the estimated autoregressive coefficient for the autoregressive model turned out to equal 0.99. This suggests that a random walk model is more appropriate for projecting emigration flows. Since the constant term does not differ significantly from zero, the following model is used for projecting emigration:

$$EM_t - EM_{t-1} = e_t \qquad (3)$$

where EM_t is emigration in year t.

Net migration totals can be projected on the basis of outputs from the projections of immigration and emigration or from a time series model applied to the net migration totals themselves. The autocorrelation coefficients in table 3.1 indicate that a first order autoregressive model is appropriate for net migration:

$$NM_t = \underset{(.08)}{.82}\, NM_{t-1} + e_t, \qquad (4)$$

where NM_t is net migration in year t. As the constant term does not differ significantly from zero, it is not included in the model. *Figure 3.2* shows that the projection based on this model tends to zero in the long run, whereas the projection of net migration based on separate projections of immigration and emigration is equal to around negative 20 thousand. Note, since immigration and emigration have been modelled as random walk models without drift, the projection of net migration based on these models remains at a constant level.

If the model is estimated for the 1950-2004 period, the random walk model projects a constant level of immigration and emigration, since the constant term in the model does not differ significantly from zero. However, if the model is estimated for the 1980-2004 period, the drift parameter turns out to be positive for the emigration time series (but not for the immigration series). Similar to the linear trend model, the random walk model with drift projects a straight line. The main difference is that the projections from the random walk model start from the last observed value. The projections of the linear trend model and the random walk model with drift are compared in *figure 3.3* for the observation period 1980 to 2004. The trend directions are similar, but the levels are considerably different.

One conclusion that can be made from the extrapolations above is that different methods can lead to very different outcomes. As shown in *table 3.2*, the projection of immigration and emigration for the year 2010 ranges from 90 thousand to 133 thousand and from 87 thousand to 125 thousand,

Figure 3.2. Net migration in the Netherlands, 1950-2010: Observations and projections

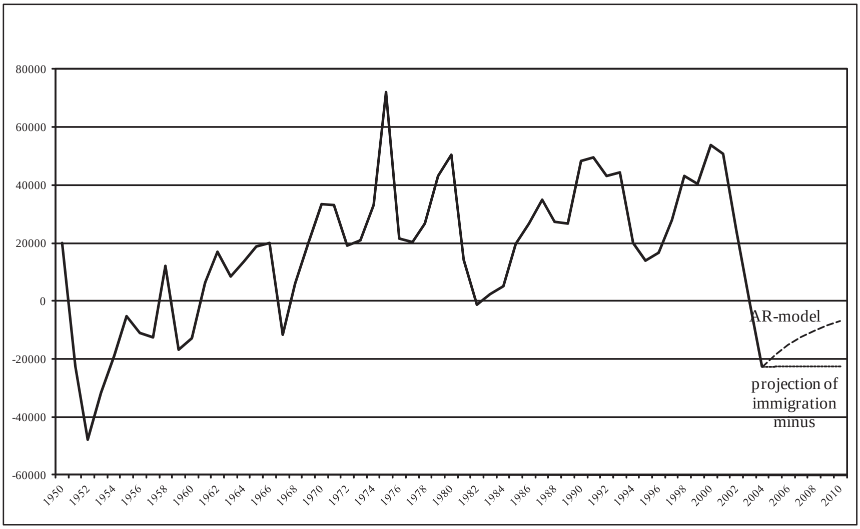


Figure 3.3. Emigration from the Netherlands, 1980-2010: Observations and projections

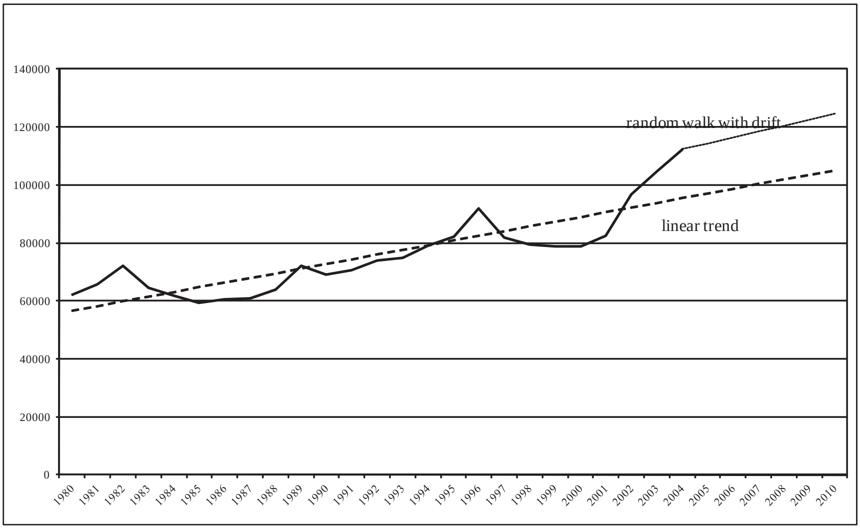


Table 3.2. Projections of immigration, emigration and net migration for 2010 (in thousands) for the Netherlands

	Immigration	Emigration	Net migration
Base period 1950-2004			
Linear deterministic model	131	87	44
Random walk model	90	112	-23
Base period 1980-2004			
Linear deterministic model	133	105	28
Random walk model	90	125	-35

respectively. The projection of net migration ranges from -35 thousand to 44 thousand. Clearly, these are considerable differences.

The results of the extrapolations depend on various choices made by the researcher. First, one has to choose between a deterministic or stochastic trend. The deterministic trend emphasises long-run developments. Projections based on this model tend to react slowly to recent changes in the time series. In contrast, projections based on a stochastic model tend to react very quickly, which may result in widely varying projections made in successive years. For example, with a start point of 2001, the random walk model projects emigration in 2010 to be 82 thousand, whereas with a start point of 2004, emigration is projected to be 112 thousand. The deterministic model resulted in 2010 emigration levels that changed much less: From 80 thousand with a start point of 2001 to 87 thousand with a start point of 2004. Second, the choice of the base period makes a difference. For example, on the basis of the 1950-2004 period, it appears that the random walk model does not require a constant term, whereas on the basis of the 1980-2004 period, it appears that a positive constant term is needed. Third, extrapolations of the time series of net migration totals may differ from the difference between separate extrapolations of immigration and emigration. Finally, there is no single extrapolation method that outperforms all other methods under all circumstances. Each one has its pros and cons. One way to decide on a particular model is to examine how the methods performed in the past. However, even this does not lead to a clear solution, as the results tend to vary depending on the choice of the period for which the methods are tested. The logical way to improve projections is to examine the explanations behind the changes in past migration, which can then be used to determine future changes.

3.3. Explanations

Extrapolations are based on the assumption that changes in the past can be projected into the future. However, without knowing the mechanisms affecting past trends, it is difficult to assess to what extent this assumption is valid. Moreover, as discussed in the previous section, different extrapolation methods may lead to widely different projections. Therefore, it is useful to look for explanations of changes in migration by identifying the main factors affecting changes in immigration and emigration. These factors can be assessed on the basis of migration theories. Massey *et al.* (1993) and Howe and Jackson (2005) give overviews of various theoretical frameworks. Most theories focus on push factors creating migration pressure in sending countries (*e.g.*, poverty, unemployment and political turmoil) and pull factors emphasizing the importance of the attractiveness of receiving countries which give direction to migration flows. Beyond this, the frameworks focus aspects, such as differentials in wage levels between countries, social networks or the role of policies.

The lack of an overall migration theory makes it difficult to forecast migration. In fact, it is questionable whether one theory is capable of explaining all kinds of changes in migration flows through time. Not only have the levels of migration changed over time, but the types and mechanisms of migration have changed as well. In the 1960s, there existed shortages in the Western European labour market. This created opportunities for large numbers of persons from Southern European countries to migrate in search for jobs. In the late 1960s and early 1970s, the origins of labour migrants shifted to Turkey and the Maghreb area. After the rise of unemployment caused by the economic recession of 1973-1974 and the influx of post-war babyboomers in the labour market, most Western European countries imposed immigration restrictions (Jennissen, 2004). As a result, many Southern European migrants returned home. The other labour migrants who stayed brought their families over, which lead to an increase of family reunification. While immigration was relatively low in the second half of the 1970s and the first half of the 1980s, immigration rose sharply during the second half of the 1980s. One of the main factors for this was the collapse of communism in Eastern Europe. A large number of ethnic Germans from Poland, the Soviet Union and Romania entered West Germany. Another cause of the rise in immigration

was the increase in the number of asylum seekers. In the second half of the 1990s, asylum migration decreased because of the end of the war in Bosnia-Herzegovina and stricter asylum policies. In short, different types of migration were predominant in different periods.
Rather than selecting one theory, one could instead focus on the main types of immigration and emigration, as they tend to be affected by different factors and change in different ways in successive periods. Here, the discussion focuses on labour migration, family-related migration and asylum seekers. Labour migration is primarily affected by the situation in the labour market (*e.g.* wage rates and unemployment rate). Marriage migration is affected by the choice of partners of the resident migrant population and, thus, by networks. Migration of asylum seekers is affected by political turmoil in sending countries and asylum policies in receiving countries.

Forecasts of migration can be based on explanations by identifying quantitative explanatory models for different types of migration. One problem in estimating quantitative explanatory models is the lack of time series data on different categories of migration. One way of dealing with this problem is to identify specific migration flows distinguished by, *e.g.*, country of origin or country of birth for which data are available and which can be considered to represent a particular type of migration. For example, the immigration flow of EU citizens to the Netherlands is comprised mostly of labour migrants, whereas the corresponding flow from Turkey and Morocco is comprised mostly of family migrants. Thus, one may expect the size of immigration of EU citizens to the Netherlands to depend on the situation of the labour market in the Netherlands.

The annual number of EU immigrants to the Netherlands during the 1977-2003 period can be explained by a regression model that includes the number of unemployed persons and a linear trend (De Beer, 2004), specified as:

$$IMEU_t = \underset{(0.007)}{-0.020}\ UN_t + \underset{(36)}{599}\ T_t + \underset{(577)}{17309} + e_t \quad R^2 = 0.92 \tag{5}$$

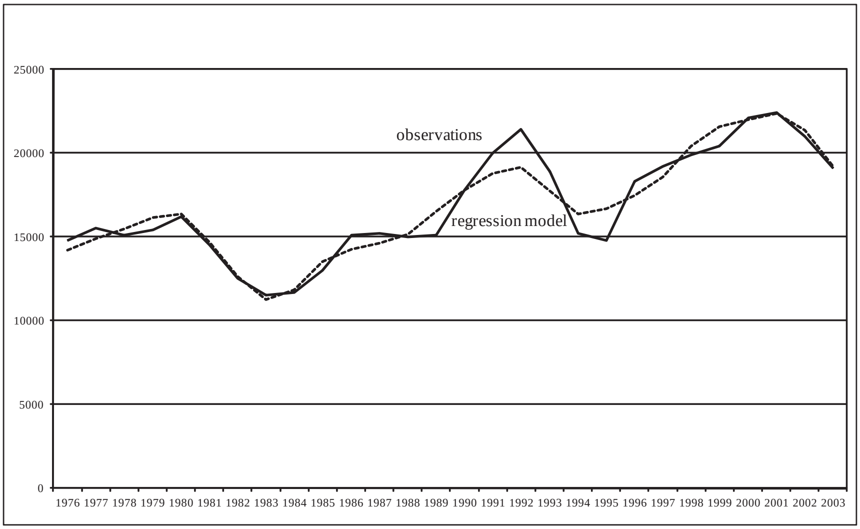
where $IMEU_t$ is the number of EU immigrants in year t, UN_t is the number of unemployed persons and T_t is a linear trend term. According to this model, a decrease in the number of unemployed persons by 100 thousand leads to an increase in the number of EU immigrants by two thousand. As shown in *figure 3.4*, the model is capable of accurately capturing fluctuations in the number of immigrants. This model suggests that the decline in immigration in the last years of the observation period is temporary and should, therefore, not be projected into the future. The estimated model indicates that, apart from short-run fluctuations due to the business cycle, there is a positive long-run trend. Brunborg and Cappelen (2010) use a similar model for projecting migration to Norway. Their model includes income and unemployment in Norway as explanatory variables as well as lagged immigration.

Since time series data for different types of migration are often lacking, expert opinions may be included in the forecasting model to obtain more accurate results. The next three sections discuss how different factors affect the main types of immigration and emigration and how they may be used to estimate future changes in migration flows.

3.4. Types of immigration

In identifying categories of immigration, it is useful to distinguish between nationals and foreigners. The size of national immigration is related to the size of national emigration in previous years. This relationship depends on the percentage of nationals who return after a stay abroad for, say, at least one year and on the length of their stay abroad. On the basis of Dutch data,

Figure 3.4. Immigration of EU citizens to the Netherlands, 1976-2003: Observed and fitted values



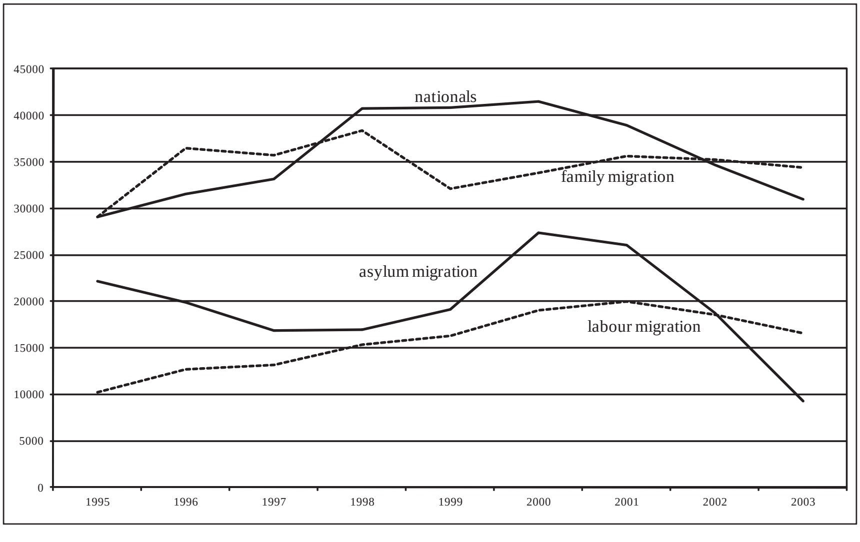
it has been estimated that one-half of all nationals who emigrated in 1995 had returned within eight years and that 60 percent return in the long run (Nicolaas, 2004). Thus, long-run forecasts of national immigration levels are equal to 60 percent of the projected national emigration levels.

In the Netherlands, the levels of labour migration, family migration and asylum seekers have changed in different ways (see *figure 3.5*). Students, retired persons and other types of immigrants are not included here. Their numbers tend to be considerably smaller than the above categories (*i.e.*, for the Netherlands). Methods for assessing and projecting the size of illegal migration are beyond the scope of this chapter.

3.4.1. Labour migration

In making assumptions about changes in the size of labour migration, one should distinguish short- and long-run developments and skill levels. In the short run, changes in the size of labour migration depend largely on the business cycle. In the previous section, it was shown that the number of EU immigrants to the Netherlands can be explained by the size of unemployment in the Netherlands. Since it is very difficult to project the course of business cycles in the future, this type of immigration cannot be projected with accuracy. However, to the extent that upturns and downturns follow each other, the business cycle does not affect the total flow of immigrants in the

Figure 3.5. Main types of immigration to the Netherlands, 1995-2003



long run. For projections, the business cycle can be used to assess recent changes in the size of immigration. For example, if immigration declined in recent years, and this decline was due to an economic downturn, it may be expected that the future level of immigration will be higher than the current level.

For long-run forecasts of labour migration, the main question is whether the ageing of the labour force in Europe will lead to shortages in the labour market and whether these shortages will lead to 'replacement migration'. The ageing process, caused by low levels of fertility and mortality, can be partially offset by increases in labour force participation rates or immigration. In 2000, the UN Population Division published a report which contained calculations on the levels of immigration needed to counteract the process of ageing (United Nations, 2000). The estimates depend on a number of assumptions, such as the rate of growth in productivity and the rate of growth in GDP. Johansson and Rauhut (2005) present various calculations on the total number of migrants in the EU in the period 2000-2050 that would be needed to stabilise (1) the size of the population, (2) the number of persons in the working ages and (3) the ratio of the working population to elderly population. In addition, they assess the effect of different rates of growth in productivity. Their calculations are based on net migration numbers. As shown in *table 3.3*, in order to stabilise the number of people in the working ages (15-64 years) in the EU25, the annual size of net migration would need to be around 2.5 million. If the rate of growth of productivity would be

Table 3.3. Average annual net migration and population size in the EU25, 2000 to 2050

	Net migration (thousands)			Population size (millions)		
	2000	2025	2050	2000	2025	2050
Constant population size	747	1 934	2 706	452	452	452
Constant population 15-64 years	747	2 677	2 422	452	467	480
Constant ratio population 15-64 yrs/65 yrs	747	10 412	15 040	452	650	940

Source: Johansson and Rauhut (2005).

one percent higher, the annual number of immigrants would be about 100 thousand less. Thus, the effect of the rate of productivity is relatively small. The table also shows that stabilisation of the elderly dependency ratio would require unrealistically high numbers of migrants, *e.g.*, 15 million immigrants *per year* around 2050. This would lead to a doubling of total population size by 2050.

Obviously, these kinds of calculations only give a general sense of the possible sizes of future migration. They do not take into account, *e.g.*, changes in the demand for labour, changes in the labour force participation or differences in the qualification structure of labour supply and demand. Moreover, the calculations cannot be directly used for making forecasts of labour immigration, as they refer to total net migration. Only if labour force participation rates of other types of immigrants would be the same as those of labour migrants, these calculations can be applied to total immigration. Otherwise, total immigration would need to be higher in order to achieve the same effect on the size of the labour force. Furthermore, since the emigration rate of labour migrants tends to be relatively high, the total number of immigrants will have to be considerably higher than the size of net migration. For example, if 50 percent of the labour migrants will return after some time, the total number of immigrants will have to be twice the size of net migration in order to have the same effect on the size of the labour force in the longer run. Nevertheless, despite these difficulties in assessing the future size of labour migration, it seems plausible to assume that the ageing of the labour force will cause the structural level of future labour migration to be higher than it used to be in the last decades.

3.4.2. Family related migration

Four types of family-related migration can be identified. First, a labour migrant may enter a country with family. Second, a labour migrant may bring in family some time after entering a country, Similarly, a refugee may be allowed to bring in family if the asylum request is granted. Third, migrants may marry a partner living abroad. And, fourth, nationals may marry a partner from abroad. Generally, family-related migration is only allowed under certain conditions and the rules differ by country.

For migration forecasts, assumptions about the future size of the first two categories can be related to forecasts of labour and asylum migration. In the Netherlands, about one fourth the size of total labour migration are family members accompanying labour migrants. During the last decade, the number of migrants that arrived because of family reunification was about one third of the total number of labour and asylum migrants, taking into account some time lag between the arrival of the labour or asylum migrant and corresponding family members.

The number of marriage migrants can be forecasted based on assumptions about the choice of partners by migrants. This may differ strongly between origins of migrants. For example, in the Netherlands, about two thirds of the Turkish and Moroccan migrants tend to marry a partner from the country of origin. Even a large proportion of their children born in the Netherlands (the so-called ‘second generation’) tend to marry a partner from their parents’ country of origin. For example, over 50 percent of second generation Moroccans and over 60 percent of second generation Turks marry someone from Morocco and Turkey, respectively.

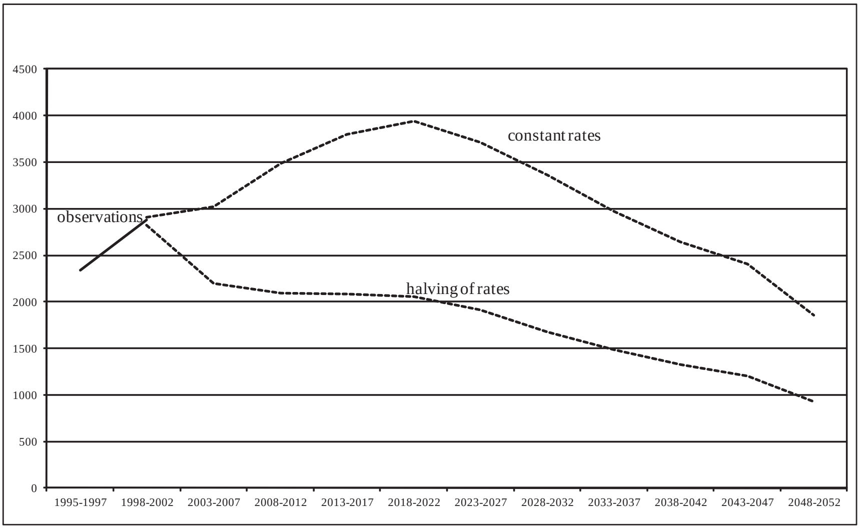
Alders (2005) developed a model for projecting the number of young Moroccans and Turks without a partner residing in the Netherlands on the basis of the current population structure (by age, sex and household position). Assuming that 95 percent of all Moroccans and Turks will eventually have a partner, the number of persons who will find a partner can be calculated. On the basis of assumptions about the percentage of these young Moroccans and Turks who will marry with a partner from the country of origin and at what age, he calculates the number of marriage migrants that can be expected in the next decades. If it is assumed that the rates of marriages with partners from abroad will remain constant, the annual number of marriage migrants would grow for some 20 years (see *figure 3.6*). If, however, it is assumed that this rate will decline gradually (a more realistic assumption), the annual number of marriage migrants is expected to decline. As illustrated in figure 3.6, if it is assumed that the percentage of young (mostly second generation) migrants marry a partner from the country of origin will halve, the annual number of marriage migrants will be considerably lower.

In addition to the marriage behaviour of migrants, one should also take into account marriages of nationals to foreigners. These numbers are considerably lower than those of migrants. In the Netherlands, they consist of about ten percent of the total number of marriage migrants. This type of migration is so small that it hardly affects changes in the total size of migration. Therefore, for migration forecasts, one can simply assume constant rates of these migrants over time.

3.4.3. Asylum seekers

For making forecasts of the total number of asylum seekers in each European country, it is useful to distinguish between changes in the total flow of

Figure 3.6. Observed and estimated marriage migrants from Turkey to the Netherlands, 1995-1997 to 2048-2052



asylum seekers to Europe and changes in the distribution of asylum seekers within Europe. Whereas the total number of asylum seekers to Europe is mainly determined by the situation in the countries of origin, the distribution among European countries is, to an important extent, affected by differences in asylum policies across European countries. In the first half of the 1990s, the total number of asylum seekers entering the EU-countries rose sharply from 400 thousand in 1990 to 675 thousand in 1992, and then fell back to 275 thousand in 1995. Since 1996, the fluctuations have been considerably smaller, increasing from 234 thousand in 1996 to 391 thousand in 2000 and subsequently declining to around 250 thousand in 2004. The average annual change declined from 135 thousand in the years 1991-1995 to 38 thousand in the years 1996-2004.

The effect of changes in the distribution of asylum seekers over the EU countries can be estimated by calculating how much the number of asylum seekers in country *i* in year *t* would have changed if the total number of asylum seekers entering the EU would not have changed compared with year *t-1*. This is the distribution effect. One alternative method is described in Van Wissen and Jennissen (2008), in which the substitution effects are estimated between all pairs of countries, rather than total distribution effects for each country separately.

The effect of changes in the total inflow of asylum seekers entering the EU on the number of asylum seekers moving to country *i* in year *t* (*i.e.*, 'generation effect') can be estimated by calculating how much the number of asylum seekers in country *i* in year *t* would have changed if the fraction of the total inflow of asylum seekers moving to country *i* would not have changed compared with year *t-1*.
In formulas:

$$D_{i,t} = A_{t-1}\,\Delta f_{i,t} \tag{6}$$

$$G_{i,t} = f_{i,t-1}\,\Delta A_{t} \tag{7}$$

where $D_{i,t}$ is the distribution effect for country *i* in year *t*, A_{t-1} is the total number of asylum seekers moving to the EU in year *t-1*, $f_{i,t}$ is the fraction of the total number of asylum seekers moving to country *i*, and $\Delta x_t = x_t - x_{t-1}$. It should be noted that these effects added together do not explain the change in the number of asylum seekers between year *t-1* and *t* completely. There is also an interaction effect; changes in the total inflow and the distribution can either reinforce each other or offset each other. However, the interaction effects are relatively small; they accounted for only three percent of the annual changes in the number of asylum seekers in the EU15 countries during the 1991-2004 period.

The average generation and distribution effects for the EU15 countries during the 1991-2004 years are set out in *table 3.4.* These estimates differ from those by Van Wissen and Jennissen (2008), as they estimated substitution effects rather than distribution effects. One benefit of their approach is that it provides more detailed estimates, *i.e.*, substitution between all pairs of countries. Their estimates, however, are based on assumptions about unobserved patterns. Table 3.4 shows that in most countries the distribution effects exceeded the generation effects. The exception is Germany. As more than half of the total number of asylum seekers in the early 1990s moved to Germany, the sharp changes in the total inflow in this period strongly affected changes in the number of asylum seekers in Germany. In the Netherlands, Sweden and the UK the distribution effects are considerably higher than the generation effect. Increases in the number of asylum seekers in these countries went together with decreases in other countries. This suggests a substitution effect, which can be caused by the fact that the asylum procedure in a certain country becomes stricter than in another country. For example, a decrease in the recognition rate in one country may lead asylum seekers to prefer to submit an asylum request in another country. There is a strong

Table 3.4. Average annual change in asylum seekers due to generation and distribution effects in EU15, 1991-2004

	Generation	Distribution
Austria	3 023	4 911
Belgium	3 135	3 942
Denmark	1 425	2 419
Finaland	375	519
France	6 215	10 295
Germany	34 729	33 242
Greece	589	1 176
Ireland	459	345
Italy	1 391	5 166
Luxembourg	97	129
Netherlands	5 033	15 213
Portugal	109	461
Spain	1 489	2 925
Sweden	5 272	12 853
UK	9 173	20 150
Average EU15	4 834	7 583

negative correlation between the distribution effects of Germany and the UK (-0.75) and Germany and the Netherlands (-0.74). This suggests that there are substitution effects between these countries.

For making assumptions about the future number of asylum seekers, separate assumptions can be made about the total inflow to the EU and the distribution between EU countries. If one assumes that there will be more co-ordination of asylum procedures in the EU, one would expect that in the short term the distribution effects will change in such a way that the flows of asylum seekers will be distributed more evenly among EU countries according to some criterion, such as the number of asylum seekers per one thousand inhabitants, and that, in the longer run, the distribution effects will become smaller when the distribution has become more even. If changes in the total flow to the EU will not exceed those in the past, one may expect fluctuations in the number of asylum seekers in separate countries in the long run to be smaller than in the period after 1990. The development in the period since 1990 clearly exhibited the effect of stricter asylum procedures in specific countries. However, as policies in other countries were less strict, the direction of the flow changed. If policies became more strict in all countries, one could expect the total inflow to the EU to become smaller. As discussed above, the ageing of the work force may lead to an increase of immigration. However, it seems likely that the EU countries will try to direct the immigration flow in order to achieve that those migrants will arrive that are qualified to occupy the jobs for which there are vacancies. Hence, some selection procedure seems likely. This means that even when the total level of migration increases, the number of asylum seekers could still decline.

3.5. Types of emigration

In making forecasts of the total size of emigration, it is useful to distinguish between return migration of foreigners and emigration of nationals. Return (e)migration of foreigners is related to foreign immigration in previous years. Thus projections of foreign emigration can be based on the immigration that occurred in preceding years. Since the patterns of foreigners emigrating to their home country differs between different types of migrants, it is again useful to distinguish between labour migrants, family-related migrants and asylum seekers.

3.5.1. Foreigners

The tendency of foreigners to return to their country of origin differs strongly between categories of migrants. Both the motive of immigration (such as labour, marriage or asylum) and the country of origin (industrialised or developing country) are important determinants. A much larger proportion of labour migrants tend to return to their home country than do marriage migrants or asylum seekers (if granted a residence permit). Immigrants from industrialised countries are more inclined to return than immigrants from developing countries. The return migration rate is higher for males than for females. The return migration rate of immigrants in adult ages are higher than those of children and older immigrants. Finally, the return migration rates of Western immigrants are higher than of non-Western immigrants. In the Netherlands, around 70 percent of male immigrants from Western countries, which are mainly labour migrants and students, return to their home country (De Jong and Nicolaas, 2005). In contrast, only about 15 percent of female immigrants from Morocco, which are mainly marriage migrants, return.

The differences imply that there are strong relationships between the size of immigration and net migration for different types of migration. For labour migration the size of immigration may be more than twice the size of net migration, whereas for marriage migration the difference between the size of immigration and net migration may be considerably smaller. Because of the rather strong relationship between the type of migrants distinguished by immigration motive (labour, asylum or family) and their demographic characteristics, forecasts of emigration can be based on distinctions by age, sex and country of origin − if data on the types of migrants are lacking.

The emigration rate of foreigners decreases with duration of stay. About one half of all emigrants leave within three years of entry. This implies that the number of emigrants is related to the number of immigrants in preceding years. For example, in the Netherlands the annual number of Western emigrants equals about two thirds of the number of immigrants three years earlier, whereas the number of non-Western emigrants equals 40 percent of the number of immigrants. Hence, if an increase in some immigration category is projected, one would expect the number of emigrants to increase with some time lag. In the long run, one could expect immigration and emigration to move in the same direction. In the short run, however, immigration and emigration may change in opposite directions, as their relationship with the business cycle differs. An economic downturn tends to lead to a decrease of immigration and an increase of emigration.

Return migration of foreigners is not always voluntary. Emigration of asylum migrants depends to an important extent on the question whether the asylum request is granted. As asylum procedures in one country become more strict, this may have an effect on both immigration and emigration numbers. First, the number of asylum seekers that are not allowed to stay and have to leave the country will increase. Secondly, the number of asylum seekers coming to that country will decline, as they will appeal for a request in another country. The migration from Africa to the Netherlands is illustrated in *figure 3.7.* Moroccan migration is excluded because the main motives are marriage migration and family reunion. For other African migrants, the main motive is asylum. In recent years, the emigration of Africans has risen for two reasons. First, the number of immigrants has risen in previous years. Second, the vast majority of asylum seekers is not allowed to stay. Forecasts of emigration of Africans can be based on the assumption that emigration will remain high in the short run as a considerable number of asylum migrants did not yet leave the country. If it is assumed that the decline of immigration of Africans will be permanent because of stricter asylum procedures, then it can be assumed that emigration will decline also in the longer run.

Finally, the age pattern of emigration of foreigners differs from that of immigration, as illustrated in *figure 3.8.* On average, emigrants are three years older than immigrants.

3.5.2. Nationals

Five main categories of emigrating nationals can be identified based on a distinction made between temporary and permanent migration. Of the temporary emigrants, there are two subcategories: students and labour migrants. Students tend to be slightly younger than labour migrants. These emigrants mainly move to other EU countries. Labour migration is inversely related to the business cycle in the home country.

Of the permanent emigrants, that is, those expecting to move for a long, indefinite period, there are three subcategories. The first are nationals marrying a partner from abroad, who choose to move to the country of their partner (those that do, tend to move to other EU countries). This category does not appear to be very large in most Western European countries. Most nationals marrying a foreign partner tend to bring their partner in to their country, particularly if they have found a partner in a non-Western country. The second category are emigrants who want to leave their country because they are not satisfied with the general situation in their home country. Most of these emigrants move to countries like Canada, Australia and New Zealand. This category represents a relatively small share of the total number of emigrants. A recent NIDI-survey shows that only two percent of the Dutch population aged 15 or over wants to emigrate (Ter Bekke *et al.,* 2005). However, only one tenth of these people have actual plans. This implies that in 2004, 20 thousand persons had serious plans and 250 thousand persons were thinking about emigrating. On the basis of these results, one would not expect a considerable increase in the annual number of emigrants. The third subcategory represents retired people who move to Southern European countries because of the warmer climate and other amenities. France, Spain and Italy are particularly popular countries of destination for these migrants. This category is as yet not very large, but may increase in the future due to the ageing of the population.

Forecasts of number of emigrants can be based on assumptions about the future values of age- and sex-specific emigration rates. *Figure 3.9* shows that emigration rates are relatively high for young children (who move together with their parents) and for men between 20 and 35 years of age. Emigration

Figure 3.7. Migration of Africans (excluding Moroccans) from and to the Netherlands, 1995-2004

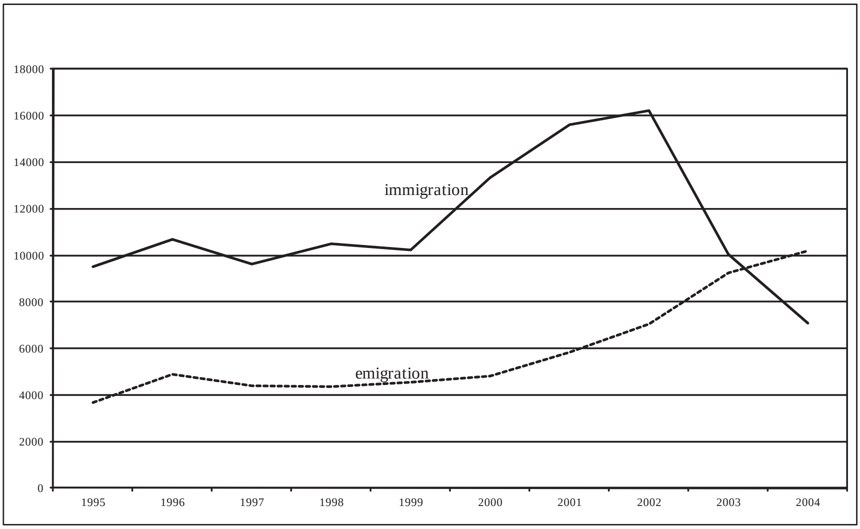


Figure 3.8. Age patterns of migration from and to the Netherlands, 2004

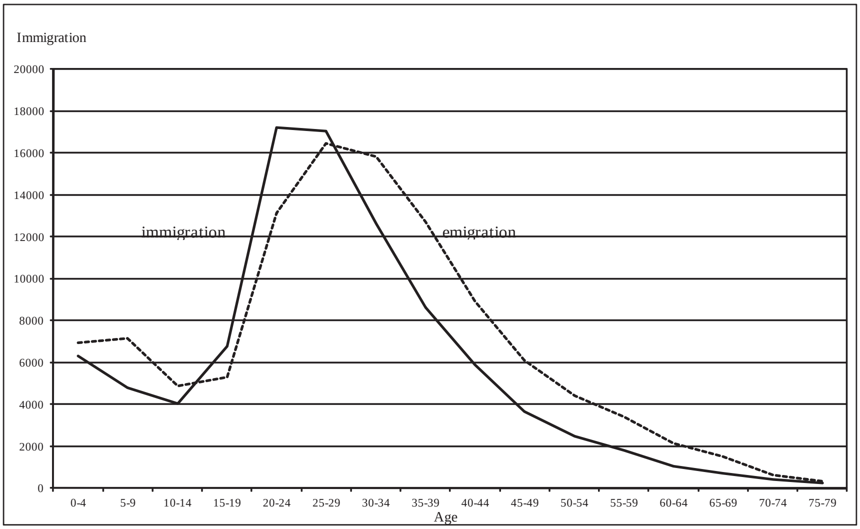
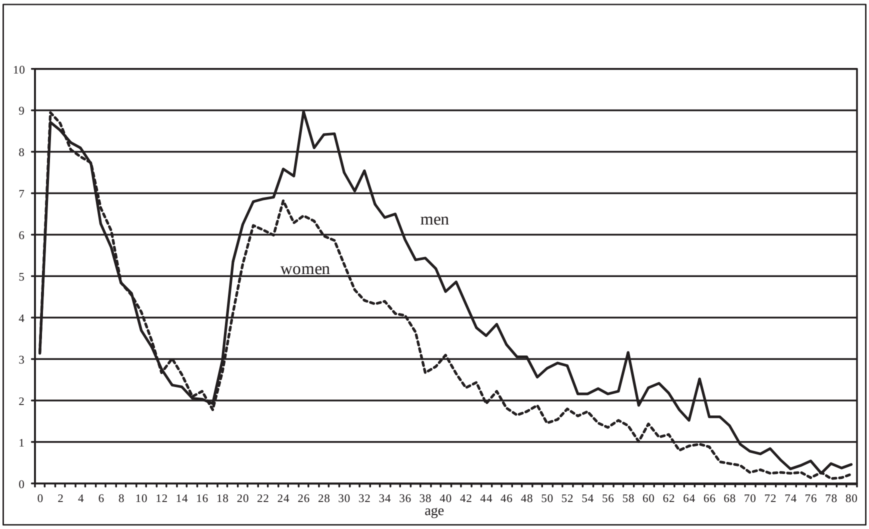


Figure 3.9. Age-specific emigration rates (per 1000) of persons born in the Netherlands, 2004



rates of women are considerably lower than those of men. If these rates are held constant over time, the changes in the number of emigrants are determined by changes in the population structure. As elderly people tend to emigrate considerably less than younger persons, ageing may be expected to have a downward effect on the size of emigration in the long run, even though there may be some increase in the number of retirement emigrants.

3.6. Assumptions on future changes in immigration and emigration

The identification of the main types of immigration and emigration and their determinants can be the basis for assumptions made on future changes. Even if no quantitative data on the separate categories of immigration and emigration are available, the distinction of types of migration is useful as a basis for argument-based forecasts of migration. In specifying assumptions on future changes in migration, it is important to take into account inter dependencies between various categories of immigration and emigration. Therefore the sequence of specifying assumptions about the separate categories is not random.

One may start with assumptions on future changes in emigration of nationals. These can be based on assumptions on the level of age- and sex-specific emigration rates. In the absence of data to estimate these rates, it could be assumed that the *numbers* of emigrants decline in the long run due to the ageing of the population, since emigration rates at older ages tend to be considerably lower than emigration rates of persons in their twenties or thirties. Subsequently, a forecast of the immigration of nationals can be based on an assumption about the percentage of emigrating nationals who return after some time. For example, as mentioned above, 60 percent of emigrating nationals from the Netherlands are expected to return. In formulas:

$$E_{N,t} = e_{N,t}\, P_{N,t} \tag{8}$$

$$I_{N,t} = r_{N,t}\, E_{N,t-k} \tag{9}$$

where $E_{N,t}$ is the number of nationals emigrating in year t, $P_{N,t}$ is the number of nationals in the population, $I_{N,t}$ is the number of returning nationals, $e_{N,t}$ and $r_{N,t}$ are emigration and return migration rates of nationals. To the extent that more detailed data are available, emigration of nationals may be related to the size of separate age groups (distinguishing different emigration rates) and immigration of nationals may be related to emigration numbers in successive years (distinguishing different immigration rates by duration since emigration).

As for making assumptions about future changes in foreign labour migration, it is useful to start with an analysis of the business cycle effect on the most recent immigration patterns. As labour immigration tends to be positively associated with the business cycle, if there is an economic upturn (downturn) at the moment the projection is made, recent immigration numbers may be higher (lower) than the structural level. Consequently, to the extent that a recent rise (fall) in immigration can be explained by the business cycle, it should not be projected linearly in the long run. Assumptions about the future development of labour migration in the long run can be based on an assessment of future shortages in the labour market caused by ageing, which may lead to an increase in labour migration. As asylum policies in an increasing number of European countries are becoming more restrictive, it can be expected that generation effects will outweigh distribution effects and that the total flows of asylum seekers to Europe will decline. The future levels of family migration are related to the choices of marriage partners in the resident population. The longer they live in a particular country in Europe, the more they will choose a partner already residing in that country. Family migration is also related to the future size of labour and asylum migration. If it is assumed that a large part of labour migrants stay only temporarily, this will have a limiting effect on the size of family related migration. Moreover, the assumed decline in the number of asylum seekers could have a downward effect on family migration. In formulas:

$$I_t = I_{N,t} + I_{NN,t,} \tag{10}$$

$$I_{NN,t} = I_{L,t} + I_{A,t} + I_{F,t} \tag{11}$$

$$I_{L,t} = p_t V_t \tag{12}$$

$$I_{A,t} = a_t P_t \tag{13}$$

$$I_{F,t} = m_{L,t} I_{L,t-j} + m_{A,t} I_{A,t-j} \tag{14}$$

where $I_{NN,t}$ is number of immigrating non-nationals, $I_{L,t}$ is number of labour migrants, $I_{A,t}$ is number of asylum migrants, $I_{F,t}$ is number of family migrants, V_t is number of vacancies and P_t is population size. It is assumed that the number of labour migrants is related to the number of vacancies, the number of asylum seekers is related to population size and the number of family migrants is related to the numbers of labour and family migration j years earlier.

Foreign emigration rates differ between categories of immigrants. As a larger part of labour migrants tend to return to their home country within a particular time period than family and asylum migrants, the assumption is that the share of labour migrants in total immigration increases and that family and asylum migration decreases. This leads to the expectation that the number of emigrants will decline in the long run. In formulas:

$$E_{NN,t} = e_{L',t} I_{L,t-i} + e_{A,t} I_{A,t-i} + e_{F,t} I_{F,t-i} \tag{15}$$

$$E_t = E_{N,t} + E_{NN,t'} \tag{16}$$

Finally, the purpose of this discussion has been to demonstrate how categories of migrants can provide a foundation for argument-based forecasts of migration.

3.7. Uncertainty

There are various reasons why forecasts of migration are uncertain. First, the quality of migration data in many countries of Europe is poor. If only net migration totals are available, forecasts based on these patterns contain less information about the causes of observed changes. Second, migration patterns tend to exhibit large fluctuations, even in the short run. For example, the size of labour migration tends to change heavily over the course of the business cycle, whereas the number of asylum seekers may change quickly due to changes in policies. Net migration of the EU25 decreased from 1.3 million in 1992 to 0.6 million in 1997 and subsequently rose to 1.9 million in 2003. Migration tends to show much stronger fluctuations than annual numbers of births and deaths. As a result, for the short run, migration is the most uncertain component of population growth. Finally, migration depends on policy changes in both the country of destination and in other countries. Changes in these policies are particularly difficult to forecast.

The degree of uncertainty of migration projections can be addressed with alternative scenarios. Usually, scenarios include a high net migration variant (*i.e.*, combining high immigration with low emigration) and a low net migration variant (*i.e.*, low immigration and high emigration), ignoring the fact that there are usually positive relationships between immigration and emigration patterns. Equation (15) shows that an increase in the number of immigrants in a given year can be expected to result in an increase in the number of emigrants in succeeding years. If these relationships are strong, the degree of uncertainty in net migration will be smaller than for immigration and emigration separately. Keep in mind that, in the short run, there may be a negative relationship between immigration and emigration, due to business cycle effects.

In general, there are three ways for assessing the degree of uncertainty of migration forecasts. First, one may look at historic forecast errors (*e.g.*, De Beer, 1997 and Keilman and Pham, 2002). Secondly, stochastic time-series models (such as discussed in section 2.2) produce forecast intervals (*e.g.*, De Beer, 1993). The problem with these first two options is that for many countries in Europe, the available time series on immigration and emigration are short. The third approach uses expert judgements to determine the width of the forecast intervals by including subjective probabilities. This approach incorporates possible explanations for why migration flows could be higher or lower than expected and whether these high and low values can be thought to be permanent or temporary. Assumptions regarding upper and lower boundaries for forecast intervals of immigration and emigration can be based on the same explanations that are underlying the central or baseline projections. For example, the upper boundary of, say, the 90 percent forecast interval of immigration can be based on the assumption that (1) the number of labour migrants will equal the number that would be required to stabilize the number of people in the working ages, (2) the percentage of young foreigners marrying a partner from abroad will not decline and (3) the recent decline in the number of asylum migrants is only temporary and will rise again to the levels observed several years ago. Accordingly, assumptions can be specified about the lower boundary of immigration and about the upper and lower boundaries of emigration. These upper and lower boundaries define the width of the forecast interval in a given forecast year. For earlier years the width of the interval will be smaller, for later years the interval will be wider.

3.8. Conclusion

This chapter has discussed the usefulness of distinguishing different types of immigration and emigration, as they are affected by different factors and hence may change in different ways. This allows one to make argument-based forecasts rather than simply extrapolating changes observed in the past. First, one may distinguish migration of nationals and foreigners. Emigration of nationals can be projected on the basis of assumptions on the future values of age- and sex-specific emigration rates. Immigration of nationals can be projected on the basis of an assumption about the percentage of emigrants who will return. Second, for projections of immigration of foreigners, three main categories of migrants can be distinguished: labour, asylum and family migration. For each of these categories, assumptions on future changes can be formulated. The future size of labour migration depends on the effect of ageing on the labour market. The number of asylum seekers depends on the question to what extent policies within the European Union will be co-ordinated. This is particularly important, because a larger part of changes in the number of asylum seekers in individual European countries were due to changes in the distribution of asylum seekers over European countries rather than to changes in the total inflow to the EU. The number of family migrants depends on the tendency of labour and asylum migrants to marry a partner from the country of origin.

Foreign emigration rates differ between the three main categories of immigrants. A relatively high proportion of labour migrants tend to return to the country of origin within a limited number of years. As a relatively large proportion of asylum requests are not granted, many asylum migrants are required to leave. However, those who are allowed to stay tend to remain for a relatively long period. Marriage migrants also tend to stay for a long time.

Finally, because of the poor quality of data and the sharp fluctuations in migration time series, forecasts of net migration tend to be rather uncertain. The degree of uncertainty can be assessed by looking at the size of errors of forecasts made in the past. Another approach is to estimate the width of the forecast interval on the basis of a time-series model. Both approaches assume that the uncertainty of future migration can be assessed on the basis of past developments. However, as the size of the various categories of migrants tends to change in different ways, one may question the validity of this assumption. Hence it is useful to follow an argument-based approach in which the uncertainty of future migration is assessed by looking for reasons why immigration and emigration could be higher or lower than expected.

4. An explanatory model for projecting regional fertility differences in the Netherlands

Abstract
Current differences in the level of the total fertility rate (TFR) between Dutch municipalities are smaller than they were in the 1970s and 1980s. Nevertheless there are still considerable differences. Small municipalities have higher TFRs than large cities. This chapter aims to answer the question whether these differences will decline further until differences between large and small cities will have disappeared. For that purpose we develop a regression model of regional differences in the TFR including demographic, socioeconomic and cultural variables. Using the estimation results we decompose differences in fertility between large and small cities into the contribution of differences in levels of the determinants versus differences in the relationships between the determinants and fertility. The results show that differences in the cultural variables have a larger effect on differences in the TFR than the demographic and socioeconomic variables. As cultural differences do not tend to change quickly, they will not lead to quick changes in regional differences in the TFR. The demographic differences are not expected to lead to strong changes either, as the two demographic variables (the household structure and the ethnic structure) have opposite effects. As the effect of the socioeconomic variable is caused by differences in the magnitude of the regression coefficient rather than by differences in the value of this variable, even if the differences in this variable would disappear, this would still not lead to convergence of the TFR. Thus the chapter concludes that differences in the TFR between large and small cities are not likely to diminish quickly.

4.1. Introduction

Despite the small size of the Netherlands there are considerable regional differences in fertility rates. Whereas the average value of the Total Fertility Rate (TFR) equals 1.8, the levels of the TFR of the almost 500 municipalities range from 1.3 to 3.2. For making regional population forecasts assumptions need to be made about the future regional differences in the level of fertility, in addition to assumptions about migration and mortality. These assumptions may be based on projections based on observed differences. However, without having an explanation for the regional differences, it is difficult to decide whether changes observed in the past are likely to continue in the future and, if so, to what extent. In order to assess whether or not differences may be persistent, this chapter examines which factors explain regional differences in fertility in the Netherlands. The chapter focuses on differences in the level of TFR between small and large cities.

Three types of explanations are examined. First, differences in the TFR between municipalities may be explained by differences in the demographic structure of the population as well as by socioeconomic and cultural differences. Second, the relationship between these determinants and fertility may differ across municipalities. Third, the level of fertility of municipalities in specific regions may systematically differ from that of municipalities in other regions, apart from the differences that can be attributed to these determinants. The relative importance of each of these three types of explanations is assessed by means of specifying a regression model. The model is estimated on the basis of data that are obtained from Statline, the electronic database of Statistics Netherlands. By means of estimating the model both for all municipalities and for small and large cities separately, the model can be used to decompose differences in fertility between large and small cities into differences in the values of the explanatory variables and differences in the values of the regression coefficients. On the basis of assumptions on possible future changes in the determinants of regional fertility differences we will discuss whether the three types of explanations are likely to lead to a decline of fertility differences between large and small cities or whether differences may be expected to be persistent.

4.2. Explanations of regional fertility differences

Most studies of regional differences in fertility focus on the total fertility rate (TFR). One important reason for using this indicator is that it is not affected by differences in the age and sex structure. One problem in using the TFR as a measure of fertility is that it is affected by changes in the age at childbearing. Hence for analyzing changes in fertility on the national level an indicator of cohort fertility may be used. However, for an analysis of the level of fertility in small regional areas cohort fertility is a less useful measure than at the national level, since a relatively large part of the population moves between different municipalities during the reproductive ages. Thus a cohort fertility indicator for a given municipality does not measure the fertility behavior of ‘real’ cohorts living in that municipality. It would be affected heavily by migration flows in the past. Hence cohort measures of fertility do not seem to be very useful for analyzing fertility differences between municipalities.

In explaining and interpreting differences in fertility between regions one should be careful because of the danger of ecological fallacy. Regional differences cannot simply be interpreted as differences between individuals living in different regions. Differences across regions can be caused by differences in the composition of the population. Duchêne *et al.* (2004) make a distinction between differences in the structure of the population and differences due to different characteristics of the regions. The structure of the population affects the level of fertility, because the level of fertility differs between subcategories of the population. For example, fertility rates for married women aged around 25 of ethnic origin are higher than fertility rates for young, native women living alone. Hence a municipality in which the former group is relatively large and the latter group is relatively small will have a higher TFR than other municipalities. Since the level of the TFR is not affected by the age and sex structure, age and sex do not have to be included in an explanatory model for the TFR. Obviously other effects of the structure of the population on fertility, such as marital status and ethnicity, might also be accounted for by means of standardizing, but that would require very detailed data on both fertility and the structure of the population which are usually not available at a low regional level.

Boyle (2003) and Sandberg and Westerberg (2005) note that there are only few recent studies on regional differences in fertility and that most studies focus on cross-country comparisons. One notable exception is Hank (2001, 2002), who distinguishes two categories of regional characteristics that affect fertility behavior: economic opportunities and constraints on the one hand and social structure and culture on the other. First, fertility behavior is affected by constraints imposed by the regional living conditions (*e.g.* Courgeau and Baccaini, 1998). Hank (2001) mentions the degree of urbanization (reflecting the “general opportunity of an individual’s residence”), the local labor market, the availability of child care, the occupational structure and regional unemployment. Duchêne *et al.* (2004) add the housing market. Second, the social environment affects fertility behavior because of regional differences in attitudes towards the family and children.

Most economic studies on fertility refer to the ‘new home economics’ theory of Becker (*e.g.* Becker, 1960, 1991). Becker argues, that 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 raising the children than men, the income that a woman could earn if she participated in the labour market has an impact on fertility. Fahey and Spéder (2004) note that when Becker formulated his theory on the economics of fertility, there was a negative relationship between female employment and fertility across OECD countries. However since the 1980s the relationship has turned the other way around and become strongly positive. Engelhardt *et al.* (2004) and Del Boca (2002) argue that the change of the sign in the cross-country correlation can be explained by the fact that the 'costs' of children do not only depend on the female wage level, but on institutions determining the ability of women to combine children and work, *e.g.* opportunities for part-time employment and availability of child care. Sandberg and Westerberg (2005) conclude that high labor income in a region may imply good economic conditions which in turn may encourage young people to start a family. This is in line with the results shown by Hoem (2000), that there is a positive relationship between employment at the municipal level and fertility in Sweden. Sandberg and Westerberg (2005) assume that poor economic conditions are discouraging. Hence they expect that high local unemployment has a negative impact on fertility. Kravdal (2002) argues that unemployment does not only affect the level of fertility of those currently unemployed but that high local unemployment rates may depress wages generally. Moreover, high unemployment in the neighbourhood strengthens people's doubts about having another child as people may consider the risk of experiencing unemployment in the future as relatively high. Gauthier and Hatzius (1997) state that high unemployment has a discouraging effect on women in permanent jobs, since the risk of not being re-employed on the same terms as before childbirth will be too high. Several empirical studies on regional fertility show a negative relationship between unemployment and fertility: Naz (2000) and Kravdal (2002) for Norway, Johansson (2000) for Sweden and Del Bono (2002) for Great Britain and Italy.

Whereas economic explanations of differences in fertility are based on the assumption that fertility behavior depends on weighing costs and benefits of having children, cultural explanations emphasise the role of values and norms as to the 'ideal' family size. In analyzng the decline of fertility to below-replacement levels in many European countries, Lesthaeghe and Van de Kaa introduced the concept of the 'second demographic transition' (Lesthaeghe and Van de Kaa, 1986 and Van de Kaa, 1987). They explain the decline of fertility by the rise of values fostering individual autonomy, secularism, postmaterialism and emancipation in addition to economic factors, such as female labour force participation and housing conditions (Lesthaeghe and Surkyn, 2002 and Surkyn and Lesthaeghe, 2004). The concept of the second demographic transition is based on the assumption that shifts in values are similar across countries: 'Post-material' values emphasizing individualism are gaining ground at the expense of more conservative values emphasizng duty (Van de Kaa, 2001). Coleman (2004) questions, however, whether liberating forces would lead to convergence, as people may not necessarily be liberated in the same direction. Billari and Wilson (2001) show that preferences regarding family formation differ according to cultural context and differences between European countries are stable. Hofstede (1981) claims that cultural differences between countries are very stable through time. There is only a convergence of superficial aspects of culture (*e.g.* consumption patterns, amusement), but not of the fundamental values. Accordingly one may expect regional cultural differences within the same country to be persistent. Reher (1998) shows that differences in norms on family size between European countries have been persistent. They have deep historical roots and they are not diminishing in any fundamental way.

From this discussion of the literature we conclude that a model for explaining regional differences in fertility should include demographic variables reflecting differences in the structure of the population, socioeconomic variables reflecting differences in opportunities and constraints and cultural variables that reflect differences in values. The question to what extent differences in fertility between large and small municipalities are likely to change in the future depends on the question whether the differences in the determinants are likely to change and on the magnitude of the effect of the separate determinants on fertility differences.

4.3. Method

For making assumptions about future differences in regional fertility it is important to assess which causes of differences in fertility tend to be permanent and which causes may be temporary. First, differences in the TFR are caused by differences in the demographic structure between municipalities, particularly differences in the proportions of women of ethnic origin and of married women in the childbearing ages. These differences may change due to migration. Secondly, differences in the TFR can be explained by socioeconomic and cultural differences between municipalities. Billari and Wilson (2001) state that whereas economic forces have led to converging trends in Europe, cultural factors have generated diverse family trends. Thirdly, the level of the TFR of municipalities in specific regions may differ from that in other regions, even if differences due to demographic, socioeconomic and cultural variables are accounted for. By means of examining whether these differences were also observed in the past one may conclude whether these differences are likely to be persistent.

In order to assess the size of the effects of these sources of variation a model is developed in which regional differences in the TFR are explained in two steps. First an explanatory model is specified which includes variables that reflect demographic, socioeconomic and cultural differences between municipalities. In the second step systematic regional patterns in the TFR that cannot be explained by these variables are identified on the basis of an analysis of the residuals of the model specified in the first step and regional dummy variables are added to the model. In the first step the following model is specified:

$$TFR_i = b_0 + \sum_j b_j x_{i,j} + r_i \tag{1}$$

where TFR_i is the total fertility rate in municipality i, $x_{i,j}$ are the explanatory variables and r_i are regional differences in the TFR that cannot be explained by the variables included in the model, with $E(r_i) = 0$. TFR, x and r refer to year t; a subscript indicating the year t is left out for the sake of readability. It can be expected that r exhibits spatial autocorrelation, as municipalities within the same region may show similar differences in fertility that cannot be explained fully by the variables included in model (1). *Moran's I* coefficient is the most commonly used coefficient in spatial autocorrelation analyses (*e.g*. Diniz-Filho *et al*., 2003). If there is spatial autocorrelation (*i.e.* if Moran's I is close to –1 or +1), estimation of the coefficients b_0 and b_j of (1) by OLS would lead to underestimating the standard errors. Moran's I measures the overall pattern of spatial autocorrelation within a given distance class. However, even if the value of Moran's I is close to zero, there still may be systematic patterns in the residuals in some specific regions, which do not lead to a high absolute value of Moran's I if there are no systematic patterns in other regions. Therefore it is useful to examine whether there are regions in which the residuals r_i indicate that the TFRs of the municipalities within that region are systematically lower or higher than would be expected on the basis of the values of the explanatory variables. These systematic differences can be modeled by:

$$r_i = \sum_k c_k D_{i,k} + \varepsilon_i, \tag{2}$$

where $D_{i,k} = 1$ if municipality i belongs to region k and $D_{i,k} = 0$ otherwise and ε_i is an error term with $E(\varepsilon_i \varepsilon_k) = 0$ for $i \neq k$ and $E(\varepsilon_i)^2 = \sigma_\varepsilon^2$.
The term $\sum_k c_k D_{i,k}$ describes the systematic regional differences in the TFR that cannot be explained by model (1), whereas the error term describes the random variations.

Combining (1) and (2) yields:

$$TFR_i = b_0 + \sum_j b_j x_{i,j} + \sum_k c_k D_{i,k} + \varepsilon_i, \tag{3}$$

If the error term is serially uncorrelated, the parameters can be estimated by OLS. One benefit of modeling spatial correlation by means of including dummy variables rather than introducing a spatial lag or error model is that the dummy variables allow to account for differences in the degree of autocorrelation across regions. Even if over-all autocorrelation is relatively small, autocorrelation between municipalities in specific regions may be relatively high. Introducing dummy variables for the latter regions provides information on deviations in the TFR that can be attributed to characteristics of specific regions that cannot be accounted for by the demographic, socioeconomic and cultural variables included in the model.

By means of estimating equation (3) both for all municipalities and for large and small cities separately, the regression model can be used to decompose differences in fertility into the effect of differences in the values of the determinants and the effect of differences in the values of the regression coefficients. The contribution of differences in determinants can be calculated by multiplying the estimated values of the regression coefficients in the model estimated for all municipalities by the average values of the explanatory variables in large and small cities respectively and calculating the difference of both products for each explanatory variable. The contribution of the differences in the values of the regression coefficients is calculated by multiplying the average value of the explanatory variables for all cities by the regression coefficients estimated for small and large cities respectively and calculating the differences.

4.4. Data

As discussed in the previous section, for the explanatory model (1) three categories of variables are specified. As the data are obtained from Statline, the electronic database of Statistics Netherlands that can be found on http://statline.cbs.nl, the choice of variables depends on the availability of data in this database. Statline contains regional data on population, households, labour, income, social security, housing and elections.

Demographic variables

These variables reflect differences in the household and ethnic structure of the population. As noted in the previous section changes in the age and sex structure do not have to be included in the model as the TFR is not affected by those changes. It can be expected that the level of the TFR is affected by the household structure, since the level of fertility of couples is considerably higher than that of people living alone. In addition, the level of the TFR is expected to depend on the size of ethnic groups, as women from a non-Western origin tend to have more children than native women. Thus two demographic variables are included in the model:

- Household structure: This variable is measured by the percentage of women aged 20-40 years living alone. This age group is selected because the major part of fertility is realized within this age group.
- Ethnic structure: Measured by the percentage of women aged 15-30 years with a foreign, non-Western background, more specifically women with a Turkish or Moroccan background. The age group is younger than that of the household variable, because Turkish and Moroccan women tend to have their children at a younger age than native women. Turks and Moroccans make up two of the largest four ethnic groups in the Netherlands. As the other two large groups, Surinamese and Antilleans, do not have higher fertility than the average Dutch level, this variable is restricted to Turkish and Moroccan women.

Socioeconomic variables

Socioeconomic variables are included in the model in order to reflect the assumption that the level of fertility depends on economic constraints and opportunities. The housing market may have an effect on couples' childbearing decisions. The availability of houses may attract couples from other municipalities, thus leading to selective migration of couples who want to have children. In particular, areas in which relatively many new houses are built tend to attract couples in the family building stage of life. In addition the level of fertility is assumed to be related to wealth. As raising children is expensive it is assumed that couples with a low income and especially couples in which one or both partners do not have a job, tend to have less children than the ideal family size. This assumption corresponds with the empirical finding discussed in the previous section that cross-country studies show a positive relationship between income and the TFR and various regional studies show a negative relationship between unemployment and fertility. Thus it is expected that the TFR is low in municipalities in which a relatively large proportion of the population does not have paid work. Hence the following variables are included in the model:

- New houses: The number of newly built houses as a percentage of the stock of houses. As it is assumed that young couples first move to their new house and then have children, the percentage of new houses in the two years preceding the year for which the TFR is to be explained is included in the model.
- The percentage of the population with a low income. This is measured by the percentage of persons receiving the minimum wage.
- The percentage of the population receiving social benefits, either because of unemployment, disability or absence of other means of income.

Cultural variables

One problem in identifying cultural differences between municipalities is that they are difficult to measure directly. Surveys which include questions on values do not have enough observations for analyses at the level of municipalities. For that reason the impact of cultural influences is assessed indirectly by means of specifying indicators assumed to reflect the effects of cultural differences on fertility. In the Netherlands, as in most other Western countries, the effect of religion on the level of fertility nowadays is much smaller than it used to be some decades ago. Nevertheless, there is still some effect, as orthodox Calvinist couples tend to have much higher fertility than the average population (Sobotka and Adigüzel, 2002). This leads to relatively high values of the TFR in the so-called Bible Belt, which extends from the South and Western part of the Netherlands in a North and Eastern direction. In addition, many studies have shown that in rural areas the level of fertility tends to be higher than in urban areas. Norms have a stronger impact in rural areas as social control and direct social influence play a more important role in rural than in urban areas. As cultural differences tend to be persistent over a long period of time, the effect of unobserved cultural differences on the TFR can be assessed by examining to what extent differences in the TFR have been long lasting. For that reason, the differences in the TFR between each municipality and the average level some decades ago is included in the model. Hence the following three variables assumed to represent cultural differences are included in the model:

- Religion: Since there are no accurate data for small municipalities of the percentage of the population affiliated with orthodox Calvinist churches, an indirect measure is used, *viz.* the percentage of persons who voted for orthodox Calvinist parties during the elections of the Dutch Lower House in 2002. Similarly, Brunetta and Rotondi (1989) use election results of the Christian Democrats as an indicator of the importance of Catholic culture in a province.

- Urbanization rate: The degree of urbanization is measured by the number of addresses per squared kilometer. Five classes of urbanization rate are distinguished ranging from very low urbanization rate (less than 500 addresses per km^2) to very high urbanization rate (more than 2,500 addresses per km^2). Four dummy variables are included in the model representing the levels of urbanization, ranging from very low to high urbanization rates.
- Non-specified cultural differences: Differences of the TFR from the average level in 1969 are regarded as a proxy for long-lasting differences in fertility. In the Netherlands the TFR changed dramatically in the years 1969-1975. On the national level the TFR dropped from 2.75 in 1969 to 1.66 in 1975. One major cause of this fall was the strong decline in the age at childbearing. As the change in the timing of fertility also affected the level of the TFR in subsequent years, it was decided to include the TFR in the last year preceding this unstable period in the model rather than the TFR in 1975. Thus the difference between the TFR of each municipality in 1969 and the average level is included in the model.

On the basis of the expected signs of the regression coefficients, it is assumed that the TFR is high in municipalities where a high percentage of women is living with a partner, a high percentage of women in the reproductive ages has a non-Western background, the percentage of new houses is high, there are low percentages of persons with a low income and persons receiving social benefits, a high percentage of the population belongs to the orthodox Calvinists, the urbanisation rate is low, and the level of fertility has been high in the past.

An analysis of the residuals of model (1) shows to what extent there are systematic regional patterns that cannot be accounted for by the explanatory variables included in the model. On the basis of the classification developed by Eurostat, three levels of regional aggregations of municipalities are examined:

a. NUTS I level: The Netherlands is divided into four parts: North (consisting of 68 municipalities), East (103 municipalities), West (207 municipalities), and South (118 municipalities). These regions are separated by geographical boundaries.

b. NUTS II level: The Netherlands consists of 12 provinces. These regions have political boundaries. The number of municipalities per province ranges from 6 to 92.
c. NUTS III level: 40 so-called COROP regions are distinguished. These are socioeconomic regions. Each region is part of one province. The number of municipalities per COROP region ranges from 2 to 33.

After assessing in which regions there are systematic differences in the TFR of the municipalities belonging to that region which cannot be accounted for by the explanatory variables, dummy variables for model (2) are specified. In order to limit the number of variables in the model, a hierarchical procedure is followed, *i.e.* first it is examined whether there are significant deviations at the NUTS I level, subsequently at the NUTS II level and finally at the NUTS III level.

The analyses are based on data for all 496 municipalities of the Netherlands (this was the number of municipalities on 1 January 2002). Population size of the municipalities ranges from 1000 inhabitants to over 700,000 inhabitants. The model is estimated on the basis of data for the year 2002. As the TFR for many small municipalities shows relatively large random fluctuations from one year to the other, it was decided to calculate the average value of the TFR for three successive years (2000, 2001 and 2002). Whereas for almost 60 percent of the municipalities the TFR ranges from 1.6 to 2.0, 15 percent of municipalities has a level of the TFR above replacement level (2.1) and 6 percent has a TFR lower than 1.5. The (unweighed) average value of the TFR equals 1.8, and the standard deviation equals .26. The TFR is low in both the most Southern and most Northern provinces (1.6 on average), which are characterized by rather poor economic conditions, and in the urbanized Western provinces (1.7). The TFR is high in the new province of Flevoland (2.0) and also in the rural Eastern provinces (1.9). A large part of Flevoland was reclaimed from the IJsselmeer lake. It consists of three polders, the last of which was created in the 1960s. Its biggest city, Almere, received its first inhabitants in 1976. Now it has 170 thousand inhabitants.

Table 4.1 shows the mean values and standard deviations of the TFR and the explanatory variables, separately for small and large municipalities. The table shows that the TFR is higher in small municipalities than in larger ones. On the basis of the hypotheses on the signs of the coefficients discussed above it can be assumed that the relatively high level of the TFR in small municipalities can be explained by the relatively low percentage of women living alone, the high percentage of orthodox Calvinists and the high level

Table 4.1. Descriptive sample statistics

	<25 000 inhabitants		>25 000 inhabitants		all municipalities	
	mean	std.dev.	mean	std.dev.	mean	std.dev.
TFR	1.88	.23	1.78	.19	1.84	.22
% Women living alone	7.21	3.10	11.63	6.92	8.83	5.31
% Moroccan and Turkish women	1.31	2.01	4.53	4.14	2.49	3.35
% New houses	4.51	3.43	5.50	4.71	4.87	3.97
% Persons with low income	7.17	1.86	8.27	2.38	7.58	2.13
% Persons receiving social benefits	11.96	3.15	14.58	3.37	12.92	3.47
% Orthodox Calvinists	5.26	8.21	4.38	5.52	4.94	7.34
Very low urbanisation (dummy)[1]	.47		.10		.33	
TFR in the past (deviation from average)	.12	.60	-.21	.42	0.00	.56
N	314		182		496	

[1] Standard deviation is not given, as this is a binary variable.
Source: Statline (www.cbs.nl).

of fertility in the past. However, these effects are counterbalanced by the low percentage of people with a non-Western foreign background, the low percentage of new houses and the low percentage of persons receiving social benefits. Thus a multivariate analysis is needed to quantify the size of these different effects on the fertility differences.

4.5. Results

Most regression coefficients of the explanatory variables turn out to differ significantly from zero and have the expected sign. The regression coefficient of the income variable does not differ significantly from zero. Hence this variable is not included in the model. Furthermore, three of the four dummies representing different degrees of urbanization do not differ significantly from zero. Thus only the coefficient of the dummy representing very low urbanization is included in the model. Moran's I is calculated by estimating the spatial autocorrelation of the values of the TFRs and the residuals of municipalities within the same regions at the NUTS III level. Moran's I of the TFR equals .24 and that of the residuals of the model including the demographic, socioeconomic and cultural variables equals .14. Thus there is no strong spatial autocorrelation. However, for two regions at the NUTS II and six at the NUTS III level the residuals turn out to be systematically positive or negative. For that reason eight regional dummies are added to the model. After including the regional dummies Moran's I equals .05, indicating that there is no autocorrelation left in the residuals. In five regions the TFR is higher than would be expected on the basis of the demographic, socioeconomic and cultural explanatory variables, whereas three regions turn out to have a relatively low TFR. The TFR is especially high in the relatively new province of Flevoland. This province attracts relatively many young couples who move from Amsterdam, as this province provides many dwellings with gardens which are considered to be attractive for rearing children. Moreover, this province includes one ‘old' municipality belonging to the Bible Belt, Urk, with very high fertility which cannot be completely accounted for by the explanatory variables (we will come back to this later).

The model is estimated separately for the 182 municipalities with 25 thousand and more inhabitants and the 314 municipalities with less than 25 thousand inhabitants. *Table 4.2* shows the estimated regression coefficients, their standard errors and the t-statistics. The model turns out to explain 78 percent of the variance of the TFR for the large municipalities and 61 percent of that for small municipalities. Taking all municipalities together the model explains 67 percent of the variance. The main part of the explained variance can be attributed to the demographic, cultural and socioeconomic explanatory variables. These variables explain 62 percent of the variance of the TFR for all municipalities.

By means of combining information from table 4.1 on the mean values of the explanatory variables and the values of the regression coefficients shown in table 4.2 one can explain the higher value of the TFR in the small cities. In cities with less than 25,000 inhabitants the average value of the TFR equals 1.88 and in the larger cities the TFR equals 1.78. This difference can be decomposed into the contribution of differences in the values of the explanatory variables between large and small cities versus differences in the

Table 4.2. Estimation results on the determinants of the TFR

Explanatory variables	<25 000 inhabitants			>25 000 inhabitants			all municipalities		
	b	s.e.	t	b	s.e.	t	b	s.e.	t
Intercept	1.988	.49	40.8	1.921	.044	43.9	1.969	.031	62.9
Demographic variables									
% Women living alone	-.12	.003	-4.1	-.014	.001	-11.3	-.013	.001	-9.7
% Moroccan and Turkish women	.009	.005	1.9	.006	.002	2.7	.008	.002	3.7
Socioeconomic variables									
% New houses	.004	.003	1.5	.004	.002	2.6	.004	.002	2.5
% Persons receiving social benefits	-.012	.003	-3.7	-.005	.003	-1.5	-.009	.002	-4.3
Cultural variables									
% Orthodox Calvinists	.012	.001	10.1	.012	.001	8.1	.013	.001	13.6
Very low urbanisation (dummy)	.039	.020	1.9	.060	.025	2.4	.041	.015	2.8
TFR in the past (deviation from average)	.093	.018	5.1	.055	.024	2.3	.088	.014	6.4

	c	s.e.	t	c	s.e.	t	c	s.e.	t
NUTS II regions									
Overijssel	.069	.048	1.4	.055	.027	2.1	.056	.027	2.1
Flevoland	.260	.168	1.5	.195	.070	2.8	.209	.080	2.6
NUTS III regions									
South West Friesland	.081	.063	1.3	.231	.095	2.4	.107	.051	2.1
South East Friesland[1]				.107	.041	2.6	.115	.054	2.1
South West Drenthe	.220	.105	2.1	.121	.068	1.8	.177	.065	2.7
Groot-Rijnmond	-.148	.037	-4.0	-.141	.032	-4.4	-.144	.026	-5.6
West Noord Brabant	-.151	.068	-2.2	-.074	.035	-2.1	-.103	.037	-2.8
South Limburg	-.099	.045	-2.2	-.169	.044	-3.8	-.111	.033	-3.4
N	314			182			496		
R square	.608			.783			.667		

[1] No coefficient for small municipalities is shown because there are no small municipalities in this region.

values of the regression coefficients. The contribution of both differences is shown in *table 4.3*.

Looking at the differences in the explanatory variables between small and large cities it turns out that the two demographic variables have opposite effects. The percentage of women living alone accounts for a difference in the TFR of .06. This can be calculated as follows. In small municipalities 7.2 percent of women aged between 20 and 40 years live alone compared

Table 4.3. Difference in TFR between small and large cities

	Contribution of differences in mean values of explanatory variables	Contribution of differences in regression coefficients	Total contribution to difference in TFR between small and large cities
Demographic variables			
% Women living alone	.06	.02	.08
% Moroccan and Turkish women	-.03	.01	-.02
Total effect	.03	.03	.06
Socioeconomic variables			
% New houses	.00	.00	.00
% Persons receiving social benefits	.02	-.09	-.07
Total effect	.02	-.09	-.07
Cultural variables			
% Orthodox Calvinists	.01	.00	.01
Very low urbanisation (dummy)	.02	-.01	.01
TFR in the past (deviation from average)	.03	.00	.03
Total effect	.06	-.01	.05
Intercept	.00	.07	.07
Total	.11	-.01	.10

with 11.6 percent in large cities (see table 4.1). As the regression coefficient of this variable in the model estimated for all municipalities equals -.013 this variable accounts for a difference of -.013 x (7.2 – 11.6) = .06 in the TFR between small and large cities. The percentage of women with a non-Western background has an opposite effect. As the percentage of women with a Turkish or Moroccan background is lower in small cities than in large cities, whereas this variable has a positive effect on the level of the TFR, this variable has a downward effect on the TFR for small cities. The size of this effect on the difference between small and large municipalities equals -.03. Thus taken together, the two demographic variables explain .03 of the total difference in the TFR. Similarly the effects of the different values of the cultural and socioeconomic variables can be calculated. The two socioeconomic variables explain a difference in the TFR of .02 (.00 by new houses and .02 by persons receiving social benefits) and the three variables representing cultural differences explain a difference of .06 of the TFR (.01 by religion, .02 by urbanization and .03 by past differences in the TFR).

As to the differences in the estimated values of the regression coefficients, the main difference between small and large cities concerns the percentage of persons receiving social benefits. The percentage of people receiving social benefits is higher in large cities than in small cities. As the regression coefficient is negative, this partly seems to explain the lower fertility in large cities. However, this effect is offset by the fact that the (absolute) value of the coefficient is larger in small cities than in large cities. Thus, even though in small cities the percentage of people receiving social benefits is lower than in large cities, this variable has a larger impact on the TFR in small cities. Thus in small cities with many people receiving social benefits the TFR is relatively low compared with other small cities. As to the difference between small and large cities, the large (absolute) value of the coefficient in small cities implies that this variable has a negative impact on the level of the TFR in small cities. This negative effect exceeds the positive effect due to the lower percentage of persons receiving social benefits in small cities. The opposite is true for the percentage of women living alone. As the (absolute) value of the regression coefficient is higher for large cities this enlarges the effect of the higher percentage of women living alone in large cities. The difference in the values of the intercept for small and large cities implies that part of the difference in the TFR cannot be accounted for by the explanatory variables. The different values of the regression coefficients of the regional dummies indicate that in four regions the difference in the TFR between small and large municipalities is larger than the difference in the intercept indicates.

If we take both effects together, it turns out that the .10 difference in the TFR between large and small municipalities is made up of a difference by .06 that can be explained by the demographic variables, .05 by the cultural variables and .07 by not-specified differences (accounted for by differences in the intercept between large and small cities), whereas the socioeconomic variables have an opposite effect of .07.

In all eight regions included in the model the TFR in the past already differed in the same direction as the signs of the regression coefficients indicate. These differences are not completely explained by the variable *TFR in the past*. This indicates that the reduction of the TFR during the last decades has not been similar in all regions. For all municipalities taken together the model implies that current differences in the TFR equal less than ten percent of the differences in 1969 (the regression coefficient equals .087, see table 4.2). However, for the eight regions included in the model the current differences in the TFR are about one half of the past differences. This indicates that in those regions the difference of the TFR with the national average has declined much more slowly than in other regions.

4.6. Implications for forecasting

In the previous section we showed to what extent the difference of the TFR between large and small municipalities can be explained by differences in the values of demographic, socioeconomic and cultural variables and by differences in the magnitude of the regression coefficients. On the basis of assumptions on the possible future direction of change in differences in these variables one may conclude to what extent future differences in the TFR are expected to decline or to be persistent.

Future changes in the demographic structure depend on the current age and sex structure and on future changes in migration and household formation (changes in mortality hardly play a role in explaining changes in the number of women in the childbearing ages and changes in fertility have an effect in the long run only). As to changes in socioeconomic regional differences Cuadrado-Roura (2001) shows that after a period of regional economic convergence in the European Union, this process has almost completely ended. Fingleton (1999) claims that there is only weak evidence that EU regions are converging, requiring more than two centuries for economic convergence to be achieved. As to cultural differences Lesthaeghe and Neels (2002) suggest, on the basis of an analysis of spatial differences in fertility that are linked to cultural variables between regions in Belgium, France and Switzerland, that these are rather stable across time.

The demographic variables lead to a lower TFR in large municipalities due to the higher percentage of single women. Thus if the difference in this percentage between municipalities will become smaller this will lead to smaller differences in the TFR. However, the other demographic variable (*viz.* ethnicity) has an opposite effect. Since this variable has an upward effect on the level of the TFR, convergence of the percentage of foreign women (due to a stronger increase of this percentage in small municipalities than in large municipalities) would lead to diverging trends in the level of the TFR. As the coefficient of the household variable exceeds that of the ethnicity variable one may expect that the former effect will be larger than the latter (assuming that the change in the percentage of ethnic women is not considerably larger than the change in the percentage of women living alone). Consequently, if demographic differences between large and small cities would become smaller, this can be expected to lead to some convergence in the TFR, although the total effect will be only moderate due to the effects in opposing directions. However, one may question whether it is likely that demographic differences between small and large cities will become smaller. Selective migration may cause differences in the population structure to be persistent. If couples wanting to raise a family move to non-urban regions whereas singles move to large cities, the differences in the household structure between large and small cities may not tend to decline. Similarly, if new immigrants move to cities where already many ethnic groups are living, the differences in the ethnic structure may not become smaller either. Furthermore one may expect the level of fertility of ethnic groups to decline in the future due to the integration of ethnic groups in society. This would lead to a lower value of the regression coefficient of the ethnic variable. Consequently the TFR in large cities would decline, causing a bigger difference in the TFR between large and small cities. Summing up, it can be concluded that demographic changes are not expected to lead to smaller differences in the TFR between large and small cities.

The main effect of the socioeconomic variables is that of the percentage of people receiving social benefits. In the previous section it was shown that the effect of this variable on the difference in the TFR between small and large cities is not so much caused by differences in the value of this explanatory variable but rather by the difference in the size of the effect (*i.e.* the value of the regression coefficient). This is due to the fact that this variable has a larger impact on the TFR of small municipalities than that of large ones. This implies that if the difference in the percentage of people receiving social benefits between small and large cities would reduce and the value of the regression coefficient would not change, this would have only moderate effect on the difference in TFR.

The three cultural variables have effects in the same direction. However, as cultural differences do not tend to change quickly, they may not lead to a strong convergence of the TFR in the near future. But in the long run one might expect this variable to lead to some convergence of the TFR. For example, if the percentage of orthodox Calvinists in a small municipality will drop by ten percent, the TFR is expected to decrease by .12. In addition, the effect of the level of the TFR in the past may lead to a decrease in future differences in the TFR as there has been some convergence in the TFR during the past decades. In 1969 the TFR in small cities was .3 higher than in large cities, whereas around 2000 the difference was .1. However, according to the model the effect of this decrease on the future difference of the TFR between small and large cities will only be moderate, since the regression coefficient is relatively small. For the next 30 years the effect of the past decrease in the difference in the TFR will be a reduction of the average TFR for small cities by .02.

Finally the model includes a number of parameters that take account of other effects than those of the explanatory variables. First, the intercept differs between small and large cities. This implies that the relatively high fertility in small cities cannot be completely explained by the demographic, socioeconomic and cultural variables. In the absence of a clear explanation it is obviously difficult to argue whether a future reduction of this difference can be expected. True, as remarked above, in the last decades we have seen a reduction of the difference in the TFR, but that does not necessarily imply that a further reduction should be expected. Note that the effect of the past reduction in the TFR on the future level is already accounted for as the differences of the TFR in the past are included in the model as explanatory variable. Moreover, as Sobotka and Adigüzel (2002) show, regional variation in the TFR declined in the 1970s and 1980s, but has hardly changed in the 1990s. Second, the model includes regional dummies indicating that three regions have lower fertility and five regions have higher fertility than would be expected on the basis of the values of the explanatory variables. As discussed in the previous section, the difference of the TFR in these regions with the national average has declined considerably more slowly than in other regions. This suggests that these differences may be rather persistent, even though the size of the differences has diminished during the last decades.

Summing up these arguments it can be concluded that even though some convergence of the TFR may be expected, it seems likely that it will be only slowly and moderately, and therefore differences in the TFR between large and small cities are likely to be rather persistent.

4.7. Conclusions

Even though there are considerable differences in the level of the TFR between Dutch municipalities, current differences are smaller than they were in the 1970s and 1980s. This chapter is aimed to answer the question whether differences will decline further until convergence will be reached or whether differences between municipalities may be expected to be persistent. In order to answer this question we developed a model explaining differences in the TFR between municipalities. The model includes demographic, socioeconomic and cultural explanatory variables. The demographic variables reflect the effect of differences in the structure of the population on the level of the TFR, the socioeconomic variables reflect differences in opportunities and constraints, and the cultural variables reflect the effect of differences in values. Since these variables are not capable of explaining all systematic regional variations in the TFR, regional dummies are added to the model. In 2 of the 12 regions at the NUTS II level fertility turns out to be higher than would be expected on the basis of the values of the explanatory variables for the municipalities in those regions; moreover in 3 of the 40 regions at the NUTS III level fertility is relatively high, whereas in another 3 regions fertility is low.

The model explains two thirds of the variance of the TFR in the almost 500 municipalities of the Netherlands. Differences in the TFR between large and small cities can be attributed to both differences in the determinants and to differences in the relationship between the determinants and fertility. In order to assess the size of these differences, the model is estimated separately for small and large municipalities. In small municipalities the TFR is .1 higher than in large municipalities. Looking at the difference that can be explained by differences in the values of the explanatory variables the two demographic variables (the household structure and the ethnic structure) turn out to have opposite effects. The differences in the cultural variables turn out to have a larger effect than the other two types of variables. As to the differences in the regression coefficients, the main difference between small and large cities concerns a socioeconomic variable. The percentage of the population receiving social benefits has a much larger impact in small cities than in larger ones. If we take both types of differences together, it turns out that the .10 difference in the TFR between large and small municipalities is made up of a difference by .06 that can be explained by the demographic variables, .05 by the cultural variables and .07 by not-specified differences, whereas the socioeconomic variables have an opposite effect of .07. In all eight regions included in the model the TFR in the past already differed in the same direction. In those regions the difference of the TFR with the national average has declined much more slowly than in other regions.

Since the two demographic variables included in the model have opposite effects on the difference in the TFR between small and large municipalities, even if the demographic variables would converge, this would not lead to a complete convergence of the TFR. Moreover, due to selective migration one may question whether strong convergence of the demographic variables is likely. Thus demographic trends cannot be expected to lead to strong convergence of the TFR. The effect of the main socioeconomic variable is not so much caused by differences in the value of the explanatory variable but rather by the difference in the size of the regression coefficient. If the difference in the value of this variable between small and large cities would reduce and the value of the regression coefficient would not change, this would have only moderate effect on the difference in TFR. As cultural differences do not tend to change quickly, they may not lead to a strong convergence of the TFR in the near future either, but in the long run one might expect this variable to lead to some convergence of the TFR. In addition, one should take into account the differences in the TFR that cannot be explained by the selected variables. As in the three regions with relatively low fertility and the five regions with relatively high fertility the difference of the TFR with the national average has declined considerably less than in other regions, these differences may be assumed to be rather persistent in the future. In conclusion, even though some convergence of the TFR may be expected, it is not likely to be quick and strong and thus differences in the TFR between small and large cities may be expected to be rather persistent.

5. A new relational method for smoothing and projecting age-specific fertility rates: TOPALS

Abstract

Age-specific fertility rates can be smoothed using parametric models or splines. Alternatively a relational model can be used which relates the age profile to be fitted or projected to a standard age schedule. This chapter introduces TOPALS (tool for projecting age patterns using linear splines), a new relational method that is less dependent on the choice of the standard age schedule than previous methods. TOPALS models the relationship between the age-specific fertility rates to be fitted and the standard age schedule by a linear spline. This chapter uses TOPALS for smoothing fertility age profiles for 30 European countries. The use of TOPALS to create scenarios of the future level and age pattern of fertility is illustrated by applying the method to project future fertility rates for six European countries.

5.1. Introduction

In order to make population projections, assumptions need to be based on the future values of age-specific fertility rates in addition to assumptions about mortality and migration. Due to random fluctuations in fertility rates over time, assumptions based on extrapolation from past changes in each age-specific fertility rate tend to result in erratic age patterns. Moreover such a procedure does not take into account the fact that changes in fertility rates which are caused by changes in the timing of fertility are temporary. Postponement of fertility will first lead to a decline in age-specific fertility rates at young ages, then some time later to an increase at older ages. After a certain period the decline at young ages will come to an end, then some time later the increase at older ages will stop. Thus past trends will not continue forever. For these reasons assumptions about future fertility may be based on a parameterized model age schedule rather than by projecting individual age-specific fertility rates separately. The parameters of the model schedule may reflect the level and timing of fertility. This means that the model schedule can be used to make assumptions about the extent to which both the level and the timing of fertility may change. A large number of model age schedules for fertility have been developed. One reason why so many models have been developed is that most of them do not accurately describe age patterns of fertility for all countries in all periods.

There are two main criteria for assessing the usefulness of a method for smoothing age profiles: the accuracy of the fit of the model to the data and the possibility of interpreting the values of the parameters. There is a trade-off. As many model age schedules turn out not to provide a very accurate description of the age pattern for all ages, the model may be adjusted by the introduction of additional parameters. However, this may hamper the interpretation. Non-parametric methods, such as splines, are more flexible than parametric models and are therefore capable of describing all kinds of age patterns. However, they lack interpretable parameters. One alternative approach is to use a relational method: that is, a method in which the age pattern is related to one standard age schedule. The function specifying this relationship indicates the way in which the age pattern under study differs from the standard age schedule.

The idea of modelling deviations from a standard age schedule was developed by Brass (1974). Brass assumes a linear relationship between a double logarithmic transformation of the age pattern to be fitted and a double logarithmic transformation of a standard age schedule. One problem with this approach is that the values of the parameters lack a clear demographic interpretation. Another problem with using only two parameters (slope and intercept) is that the accuracy of the fit depends heavily on the choice of the standard age schedule. If in one or more age intervals there is some deviation between the age pattern to be fitted and the standard age schedules, two parameters are not enough. This article introduces the new relational method TOPALS (tool for projecting age patterns using linear splines), which is more flexible than the Brass method, produces a better fit, and is easier to interpret.

TOPALS is capable of describing all kinds of age curves. The parameters can be interpreted easily. TOPALS models the age pattern of the ratios of the age-specific rates to be fitted and the standard age schedule by a linear spline. The standard age pattern may be the average age pattern of a group of countries (*e.g.*, the EU average). For making projections, this may be useful if one assumes future convergence of age-specific fertility rates of different countries to an average pattern. It is also possible to use the age pattern of another country as standard. This may be a ‘forerunner’ country and one may assume that the age-specific fertility rates of different countries will move in the direction of that of the forerunner country. Alternatively the standard age pattern may be a model age schedule, such as the Hadwiger, Beta, or Gamma function. TOPALS can be used for making projections of future changes in age-specific rates by specifying assumptions about changes in the values of the ratios of the age pattern to be projected and the standard age schedule for successive age intervals. For example, if the standard age schedule is that of a forerunner country, one can assume that the future values of the rate ratios for other countries will move towards one.

The second section of this chapter presents a short overview of previous studies on modelling age patterns. We distinguish between parametric models, splines and relational methods. Section 5.3 describes TOPALS. Section 5.4 applies TOPALS in order to smooth age patterns of fertility for 30 European countries. The results of TOPALS are compared with those for six other methods. Section 5.5 describes how TOPALS may be used for creating scenarios of future values of age-specific fertility rates. Section 5.6 concludes the chapter and discusses some other possible applications of TOPALS.

5.2. Methods for fitting age-specific fertility rates

There is an extensive literature on parametric age schedules describing the typical age pattern of fertility (*e.g.*, Hoem *et al.* 1981; Rogers 1986; Booth 2006; Peristera and Kostaki 2007). Even though the general age pattern of fertility has been similar across many countries for many decades, there are important differences as well. As a consequence most model age schedules do not describe the fertility rates at each part of the age range accurately for all countries in all periods. For that reason various authors have proposed variations to the ‘traditional’ models such as the Beta, Gamma and Hadwiger models. One alternative is to use non-parametric models such as splines. They provide an accurate fit for all kinds of age pattern. However, since they do not include parameters that can be interpreted, these methods are less useful for making assumptions about future fertility as an input for population projections. One alternative is to specify a spline in such a way that the parameters can be interpreted (Schmertmann, 2003). Another is to develop a relational model in which the age pattern of a given country is related to a standard age schedule. The relationship specifies the way in which the age pattern differs from the standard age schedule.

5.2.1. Parametric models

This section describes the three most frequently used parametric models for fitting fertility age patterns: the Hadwiger, Beta and Gamma models (Hoem *et al.*, 1981 and Booth 2006). In addition a recently developed parametric method will be discussed (Peristera and Kostaki, 2007).

One of the earliest models proposed in the literature is the Hadwiger function (Yntema, 1969; Gilje, 1972 and Hoem *et al.*, 1981). This function is described by:

$$f(x) = \frac{ab}{c}\left(\frac{c}{x}\right)^{\frac{3}{2}} \exp\{-b^2(\frac{c}{x}+\frac{x}{c}-2)\} \tag{1}$$

where *f(x)* is the fertility rate at age *x* of the mother and *a, b,* and *c* are the three parameters to be estimated. Parameter *a* is associated with the total level of fertility, parameter *b* determines the height of the curve, and parameter *c* is related to the mean age at motherhood. Even though the parameters have a demographic interpretation, they indicate a direction of change only and their actual values are not directly interpretable. A higher value of *a* indicates that the total fertility level is higher. But the value of *a* does not equal the value of the total fertility rate (TFR). For example, in fitting the Hadwiger function to age-specific fertility rates in 30 European countries in 2008, we find that *a* ranges from 0.75 in Hungary (where the TFR equals 1.35) to 1.24 in Iceland (where the TFR equals 2.15). A linear regression of the values of *a* and the TFR shows that TFR equals 1.76**a*. The value of *c* turns out to be very close to the mean age at childbearing. The value of *ab/c* is related to the modal age-specific fertility rate (Chandola, Coleman, and Hiorns 1999), but this does not make the value of *b* itself easily interpretable and so is not very helpful in making assumptions about fertility.

In several European countries, such as the United Kingdom, Ireland, and Spain, fertility at young ages is higher than would be expected according to the Hadwiger function. For that reason Chandola, Coleman and Hiorns (1999) propose an extension of the Hadwiger function for describing the bulge in fertility at young ages. They assume that the relatively high fertility level at young ages reflects heterogeneity in the population related to the educational level and social status of the mothers, as well as to ethnic differences in the timing and level of fertility. They distinguish two subpopulations with a different timing and level of fertility. They describe this pattern by a mixture model: That is, they replace the right-hand side of equation (1) by the weighted sum of two similar terms which describe the age patterns of the two subpopulations.

The Gamma function is given by:

$$f(x) = R\frac{1}{\Gamma(b)c^b}(x-d)^{b-1}\exp\{-(\frac{x-d}{c})\} \text{ for } x > d \tag{2}$$

where R determines the level of fertility and d the minimum age at childbearing. The Gamma function is equivalent to the Pearson Type III model which was applied by George et al. (2004) to Canadian data. Hoem *et al*. (1981) show how the parameters b and c are related to the mode, mean and variance of the function but not in a simple, linear way and so they do not have a direct demographic interpretation.

The Beta function is given by:

$$f_x = R\frac{\Gamma(A+B)}{\Gamma(A)\Gamma(B)}(\beta-\alpha)^{-(A+B-1)}(x-a)^{A-1}(\beta-x)^{B-1} \text{ for } \alpha<x<\beta \tag{3}$$

where R determines the level of fertility. The Beta function is equivalent to the Pearson Type I curve proposed by Romaniuk (1973) and Mitra and Romaniuk (1973). Hoem *et al*., (1981) state that α and β represent lower and upper age limits of fertility, but Peristera and Kostaki (2007) show that in several cases the value of β far exceeds the maximum age. Hoem *et al.* (1981) show that A and B are related to the mean and variance, but not in a simple, easily interpretable way.

Peristera and Kostaki (2007) note that the form of the fertility curve has changed in recent years in various countries, as did Chandola, Coleman and Hiorns (1999) before them. Peristera and Kostaki (2007) propose a flexible model that describes both the standard and the distorted age-specific fertility pattern in countries such as the United Kingdom, Ireland, and Spain. Their basic model resembles the normal distribution but is asymmetrical, as the spread before and after the peak differs:

$$f(x) = c_1 \exp\{-(\frac{x-\mu}{\sigma(x)})^2\} \tag{4}$$

where $\sigma(x) = \sigma_{11}$ if $x \le \mu$ and $\sigma(x) = \sigma_{12}$ if $x > \mu$ and c_1, μ, σ_{11} and σ_{12} are the parameters to be estimated. The parameter c_1 is associated with the total fertility rate, μ is the modal age of fertility and σ_{11} and σ_{12} reflect the spread of the distribution before and after its peak respectively. In order to fit fertility curves with high fertility at a young age Peristera and Kostaki (2007) add a second term:

$$f(x) = c_1 \exp\{-(\frac{x-\mu_1}{\sigma_1})^2\} + c_2 \exp\{-(\frac{x-\mu_2}{\sigma_2})^2\} \tag{5}$$

where c_1 and c_2 reflect the level of fertility at the first and second peak respectively, μ_1 and μ_2 are related to the mean age of the two subpopulations and σ_1 and σ_2 reflect the spread around the two humps. Even though the parameters are related to the level, mean age, and spread of the fertility curve, the actual values of the parameters are difficult to interpret. For example, in fitting model (4) to fertility data of various European countries the estimate of c_1 varies from 0.09 in Italy to 0.15 in Denmark. Although there is a positive correlation between c_1 and the TFR, the correlation is certainly not perfect. If model (4) is estimated for 30 European countries, the value of c_1 explains not more than 80 percent of the variance in the TFR across European countries. Gayawan *et al.* (2010) propose the Adjusted Error Model for modelling age-specific fertility rates in African countries. This model is very similar to equation (5): they assume $c_1 = c_2$ and they add an intercept to the model.

One conclusion that applies for all parametric models is that even though the values of the parameters are related to the level, mean age, and variance of the functions, the values are not equal to well-known demographic indicators such as the TFR and the mean age at childbearing. This hampers interpretation of the parameters and limits their usefulness for creating demographic scenarios. Moreover one simple model with a limited number of parameters does not describe adequately the variety of age patterns of fertility across countries in different periods. Therefore several complex models including more parameters are needed.

5.2.2. Splines

Instead of specifying a statistical model, one may use a non-parametric model for smoothing age patterns of fertility. The structure of a non-parametric model is not specified a priori but is determined from the data. Non-parametric does not mean that the model does not include parameters, but that the number of parameters is not fixed in advance and that the parameters lack a clear statistical interpretation. There are several approaches to estimating non-parametric models. The most widely applied are local polynomial regression and smoothing splines (Fox, 2000). To the best of my knowledge, local polynomial regression is not applied to fitting fertility schedules. Quadratic and cubic splines are very flexible and so may provide an accurate fit of various types offertility curves. Brass (1960, 1975) fits third-degree polynomials to describe the age pattern of fertility. The fit of the third-degree polynomial is less accurate than that of a cubic spline, which is a piecewise cubic function (Hoem *et al.*, 1981). Cubic splines are very flexible (McNeil, Trussell and Turner, 1977 and Gilks 1986). A cubic spline can be described by:

$$f(x) = a + b(x-m) + c(x-m)^2 + \sum_{j=1}^{n} d_j (x-m-k_j)^3 D_j \qquad (6)$$

where $D_j = 0$ if $x - m \le k_j$ and $D_j = 1$ otherwise, m is the minimum age, $x \ge m$, k_j are the knots, n is the number of knots, a, b, c and d_j are the coefficients to be estimated.

A quadratic spline is a piecewise quadratic function that can be described by:

$$f(x) = a + b(x-m) + \sum_{j=1}^{n} c_j (x-m-k_j)^2 D_j \qquad (7)$$

where $D_j = 0$ if $x - m \le k_j$ and $D_j = 1$ otherwise, m is the minimum age, $x \ge m$, k_j are the knots, n is the number of knots, a, b and c_j are the coefficients to be estimated. Both a quadratic and a cubic spline can provide a good fit. In general a quadratic spline may require more knots than a cubic spline to provide an accurate fit. For both the quadratic and cubic splines, the coefficients can be estimated by Ordinary Least Squares (OLS) if the knots are fixed a priori. Otherwise an iterative estimation procedure is needed: for example, a non-linear least squares method.

Kostaki *et al*. (2009) propose a new non-parametric method: support vector machines (SVM). Even though this provides a good fit, its usefulness for projecting age-specific fertility rates seem limited. SVM models are more complex than splines. Since quadratic and cubic splines are capable of producing a good fit as well, the question is why use SVMs rather than splines.

One problem in using splines is that the values of the coefficients lack a clear interpretation. For that reason, even though splines provide an accurate fit, they are not very suitable for specifying assumptions about the values of age-specific fertility rates for population scenarios. For that reason Schmertmann (2003) proposes fitting splines by choosing as knots particular ages that can be interpreted. Schmertmann fits the age pattern of fertility by a quadratic spline including four knots, which means that the age schedule is described by five quadratic pieces. This would require 13 parameters. However, Schmertmann reduces the number of parameters by determining the knots for specific ages: The youngest age at which fertility rates are above zero, the age at which fertility reaches its peak level, and the youngest age above the peak age at which fertility falls to half its peak level. Furthermore the value of the overall level of fertility is included. Schmertmann imposes restrictions on the values of the coefficients of the quadratic spline in such a way that the resulting function describes the typical form of the fertility curve. The Schmertmann method improves the usefulness of splines for making projections. However, it is questionable whether one of the indicators, the age at which fertility falls to half its peak level, really has an obvious demographic interpretation. One general conclusion is that non-parametric models provide a better fit than parametric models. The latter are smoother and thus are not capable of describing various specific patterns at certain ages. However, the usefulness of non-parametric methods for making projections or creating scenarios is limited, as they lack easily interpretable parameters.

5.2.3. Relational methods

Fertility age schedules can be fitted by specifying how the age-specific fertility rates in a particular country deviate from some standard age schedule. Coale and Trussell (1974) modelled age-specific marital fertility rates as the product of two model age schedules: a nuptiality schedule and a marital fertility schedule. The parameters of the model indicate to what extent the age-specific fertility rates for a particular country deviate from the model age schedules. In the 1970s this model performed rather well (Hoem *et al*. 1981 and Rogers, 1986). However, since the 1980s extramarital fertility has increased in many countries. For that reason the usefulness of modelling marital fertility has decreased.

Brass (1974) presents a more general relational method in which fertility rates can be related to any fertility age schedule as long as it captures the general shape of the age pattern of fertility to be fitted. The Brass relational method is based on the assumption that the (cumulative) age pattern of fertility can be described by the Gompertz distribution. This implies that the log-log transformation of the rates to be fitted is linearly related to age. As the Gompertz distribution provides a reasonable fit except at extreme ages, Brass proposed to improve the fit by using a standard age schedule. Assuming that this standard age schedule can be described by a Gompertz distribution as well, there is a linear relationship between the log-log transformation of the rates to be fitted and that of the standard age schedule:

$$Q_x = \alpha + \beta Q_x^* \tag{8}$$

where $Q_x = -\ln(-\ln f(x))$ and $Q_x^* = -\ln(-\ln f^*(x))$; $f(x)$ are the age-specific fertility rates to be fitted and $f^*(x)$ are the fertility rates according to the standard age schedule. The parameters α and β can be estimated by OLS regression. Even though the basic assumption underlying the method is that rates can be described by a Gompertz distribution, Brass and others have shown that various types of age profile may be used as the standard age schedule, including observed rates from another country, as long as the standard schedule represents the general age pattern of the rates to be fitted (Zeng Yi *et al*., 2000). In addition to using the relational model as an instrument for making projections, the model can also be used for making estimates of age-specific rates for countries with incomplete data. The parameters α and β can be interpreted as follows: α determines the location of fertility and β the spread. Thus α indicates whether the age pattern lies to the right or left of the standard age schedule and β determines whether the age pattern is more or less dispersed than the standard.

Zeng Yi *et al*. (2000) note that even though the parameters α and β can be interpreted, in practice they do not turn out to be very useful for making projections of demographic rates for the future. The main reason is that the values α of β and are not comparable over time and regions, as they depend on the choice of the standard age schedule, and in order to have an optimal fit, the choice of standard age schedules may vary across time and space. Moreover, changes in the values of α and β lack a clear demographic interpretation. They indicate the direction of change only. For example, if the age curve is assumed to move to the right, the value of α should become larger. But it is not clear by how much. For that reason Zeng Yi *et al*. (2000) propose an alternative method for the estimation of α and β. They show that the value of β equals the ratio of the interquartile ranges of the standard age schedule and that of the age pattern to be fitted. The value of α is related to the median age. Zeng Yi *et al.* demonstrate that this model is capable of describing various unimodal age patterns: namely, fertility, first marriage, divorce, remarriage, and leaving the parental home. They use Chinese, French, Swedish, and US data. They use the model for fitting age patterns based on both a standard age pattern from an earlier period and a standard pattern from another country.

One main advantage of the method proposed by Zeng Yi *et al.* is its simplicity. Once an appropriate standard age schedule is available, one needs to estimate the values of only two parameters. One problem is that the goodness of fit depends heavily on the choice of the standard age schedule. If the age pattern to be fitted differs from the pattern of the standard age schedule in some age range, the two parameters are not sufficient to adjust the curve to produce an overall good fit. Therefore different standard age schedules may have to be used at different points in time, which makes it difficult to use this method to extrapolate changes over time.

For that reason the present chapter proposes an alternative method that is more flexible than the Brass relational method: TOPALS. Flexibility is achieved by introducing more parameters. However, this does not make the parameters harder to interpret. The parameters simply indicate the extent to which the age pattern to be fitted differs from the standard age schedule in successive age intervals. Since for different points in time the parameters may refer to the same standard age schedule they can be used for analysing changes over time and therefore become the basis for extrapolations into the future.

5.3. TOPALS

We assume that a standard age schedule of fertility rates is given. The age profile for a given country can be estimated on the basis of ratios of the age-specific fertility rates of that country and those according to the standard age schedule. The rate ratio at age x is:

$$r(x) = \frac{f(x)}{f^*(x)} \tag{9}$$

where $f^*(x)$ is the fertility rate at age x according to the standard age schedule. The age pattern of the ratios can be described by a linear spline function. This is a piecewise linear curve. The ages at which the successive linear segments are connected are called ‘knots’. The ratios at each age can be estimated by the linear spline function:

$$\hat{r}(x) = a + b_0(x-m) + \sum_{j=1}^{n} b_j(x-m-k_j)D_j \tag{10}$$

where $D_j = 0$ if $x - m \leq k_j$ and $D_j = 1$ otherwise, m is the minimum age, $x \geq m$, k_j are the knots, n is the number of knots, a and b_j are the parameters to be estimated.

This model can be estimated in several ways. The knots can be fixed a priori, *e.g.* on the basis of visual inspection of the age pattern of the rate ratios. Alternatively they can be chosen in such a way that the fit of the linear spline to the data is optimal. In the latter case a non-linear estimation method is required, *e.g.* a non-linear least squares method. If the location of the knots is fixed a priori, a and b_j can be estimated by OLS. However, these parameter values are difficult to interpret as they indicate the slopes in the successive age intervals rather than the levels at specific ages. In the first age interval $m,\ldots, k_1$ the slope equals $\hat{a} + \hat{b}_0$, in the second age interval $k_1+1,\ldots, k_2$, the slope equals $\hat{a} + \hat{b}_0 + \hat{b}_1$, etc. It is much easier to interpret the levels of the rate ratios at specific ages than the successive slopes in the age intervals. Thus from the regression estimates one can calculate the values of $\hat{r}(x)$ at the knots. These can be used as a basis for making projections. Alternatively the linear spline can be estimated in a more simple and straightforward way by assuming that at the knots the values of the spline equal the observed values. It turns out that this provides a fit that is very close to the one produced by applying OLS. Thus we assume that $\hat{r}(m)= r(m)$, $\hat{r}(k_1) = r(k_1)$, $\hat{r}(k_2) = r(k_2)$, etc. Then the values of a, b_j can be estimated by substituting the values of $\hat{r}(m)$, $\hat{r}(k_1)$, $\hat{r}(k_2)$, etcetera in (10). This yields:

$$\hat{a} = r(m)$$

$$\hat{b}_0 = \frac{r(k_1) - r(m)}{k_1 - m} \tag{11}$$

$$\hat{b}_j = \frac{r(k_{j+1}) - r(k_j)}{k_{j+1} - k_j} - \sum_{i=1}^{j} \hat{b}_{i-1}$$

For the ages above the last knot we assume that the slope equals zero. Alternatively one might assume that the slope above the last knot equals that before the last knot. Since fertility rates at high ages are small this choice hardly affects the results. In contrast, when using TOPALS to fit other schedules, such as age-specific mortality schedules this choice would make a difference. In those cases it is an empirical question which choice one would make.

The age-specific fertility rates are estimated by multiplying the ratios which are estimated by the linear spline function $\hat{r}(x)$ by the age-specific fertility rates according to the model age schedule $f^*(x)$:

$$\hat{f}(x) = \hat{r}(x) f^*(x)\,. \tag{12}$$

In the application of TOPALS in the next section we will select the knots by minimizing the sum of squared differences between $\hat{f}(x)$ and $f(x)$ by means of a non-linear least squares method. We apply a grid search where for each set of knots we calculate the values of a and b_j by solving equation (11) rather than by applying OLS.

The standard age curve can be the average of several countries (for example, the EU average), the age curve of another country (for example, a ‘forerunner’ country), the age curve of the same country in a previous year, or a model age schedule. Using the estimated rate ratios for making projections on future values of age-specific fertility rates, two alternative procedures may be followed. First, for each country one may calculate a time series of rate ratios by dividing the age-specific fertility rates for successive years by a standard age schedule which is assumed to be the ‘target’ age pattern to be reached in the long run. This may be the age schedule of a forerunner country. The time series of rate ratios shows whether and, if so, how rapidly the age-specific fertility rates of the different countries move into the direction of the fertility rates of the forerunner country. A partial adjustment model can be used to project the future values of the rate ratios. This can be considered as a quite ‘objective’ method. The choice of the ‘target’values of the age-specific fertility rates may be subjective, but the parameter of the partial adjustment model estimated for some historical period determines how rapidly the rate ratios will move towards one and thus whether the age-specific fertility rates will reach the ‘target’ values before some specific forecast horizon. Second, one may follow a more subjective approach by making assumptions about the future values of the rate ratios for selected ages based on qualitative arguments. For example, if the standard age schedule is the average pattern over a number of countries, for each separate country one may make assumptions about the extent to which one assumes that the fertility rates at different age ranges will remain different from the average or will move towards the average values. In section 5.5 we will demonstrate both procedures for using TOPALS to create scenarios.

5.4. Smoothing age-specific fertility rates

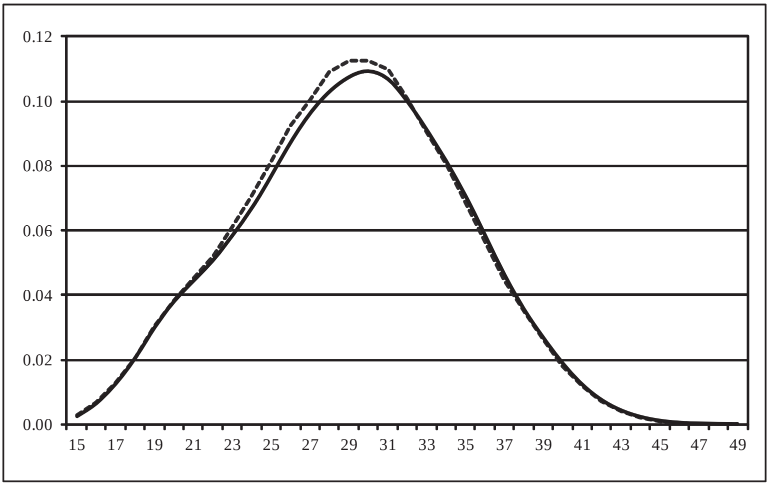
We illustrate the use of TOPALS by fitting age-specific fertility rates for 30 European countries: the 27 EU countries (Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and the United Kingdom) plus Iceland, Norway, and Switzerland. We label these 30 European countries as the EU27+3 countries. The data are obtained from the database of Eurostat, which is available online (Eurostat, 2010). At the time of writing data for 2008 were available for 27 countries. For Italy the most recent data were from 2007, for the United Kingdom from 2006, and for Belgium from 2005.

We calculated unweighted and weighted averages of age-specific fertility rates for the EU27+3 countries. Rather than using total population size of each country as weights, we calculated the weighted average by weighing the fertility rates for each country by the number of women aged 15-44 years. The TFR based on the unweighted average of the EU27+3 countries equals 1.61, while that based on the weighted average equals 1.58.

The values of the TFR range from 1.32 in Slovakia to 2.15 in Iceland. Most countries in Southern, Central and Eastern Europe have a TFR between 1.3 and 1.5, whereas most Northern and Western European countries have a TFR above 1.6. In the remainder of this chapter we will use weighted averages.

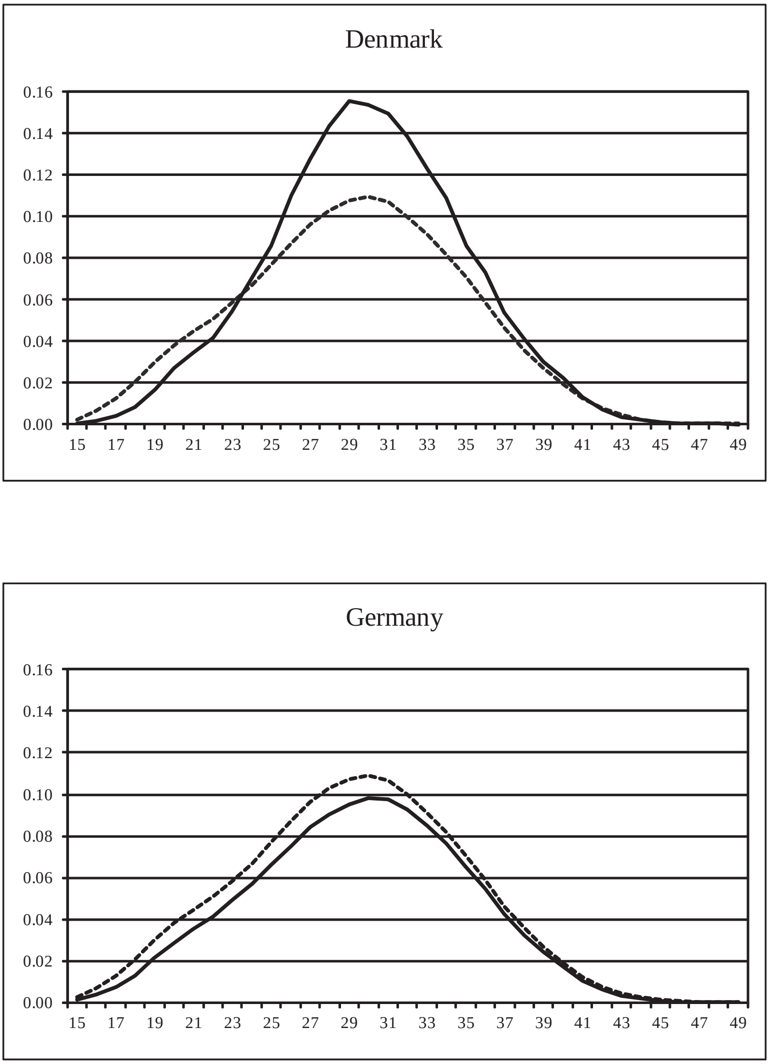
Figure 5.1 shows the unweighted and weighted age-specific fertility rates for the EU27+3 countries for the year 2008. Since several large EU countries, such as Germany and Italy, have relatively low fertility rates among women in their 20s or early 30s, the peak of the weighted average age pattern of fertility rates is lower than that of the unweighted average. *Figure 5.2* compares age-specific fertility rates of six European countries with the average of the EU27+3: Denmark, France, Germany, Italy, Poland and United Kingdom. These countries are selected as they reflect differences in the level and age pattern of fertility rates in different parts of Europe. The three Northern

Figure 5.1. Age-specific fertility rates, average of EU27+3 countries, 2008



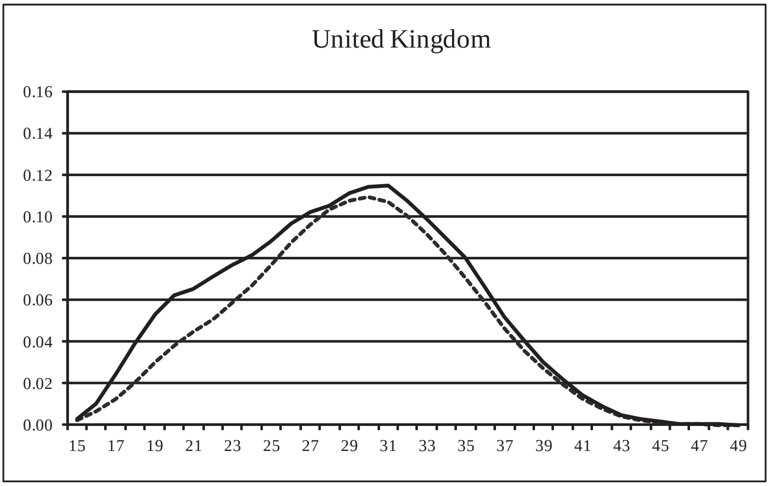
Solid line: weighted average; dotted line: unweighted average.

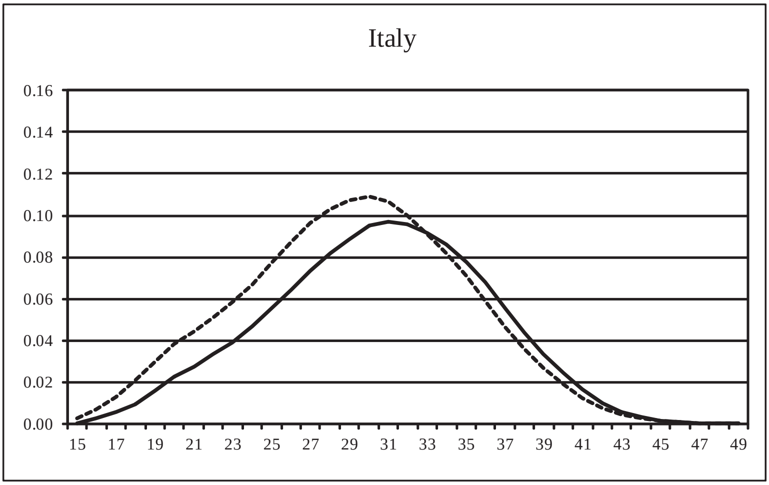
Figure 5.2. Age-specific fertility rates of six European countries, compared with the EU27+3 average, 2008



Solid line: weighted average; dotted line: unweighted average.

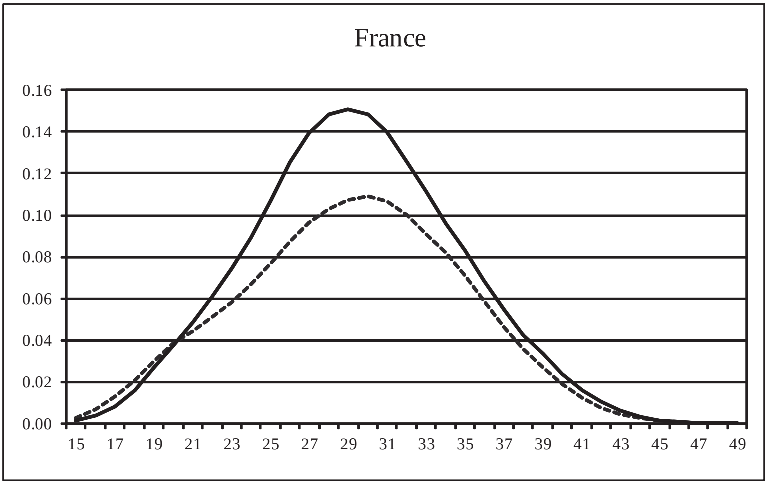
Figure 5.2. Age-specific fertility rates of six European countries, compared with the EU27+3 average, 2008 (continued)

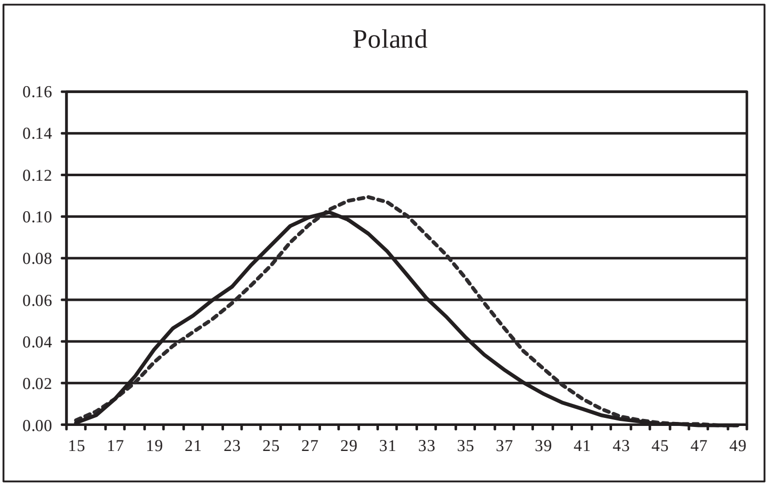




Solid line: weighted average; dotted line: unweighted average.

Figure 5.2. Age-specific fertility rates of six European countries, compared with the EU27+3 average, 2008 (end)





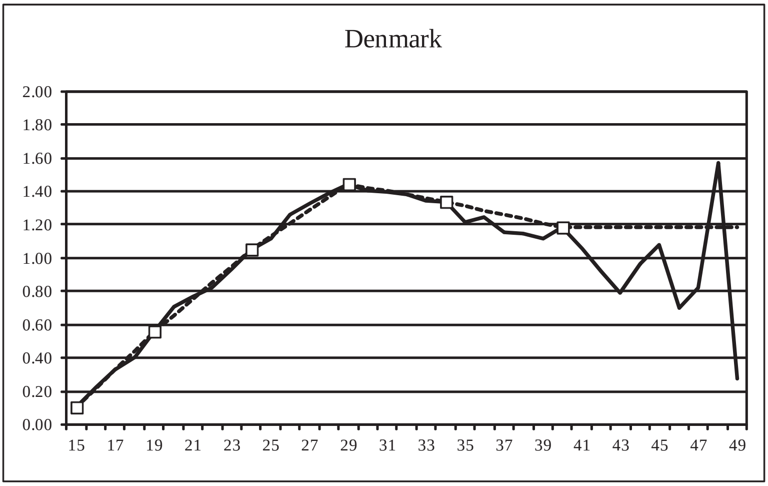
Solid line: observed values; dotted line: EU27+3 average.

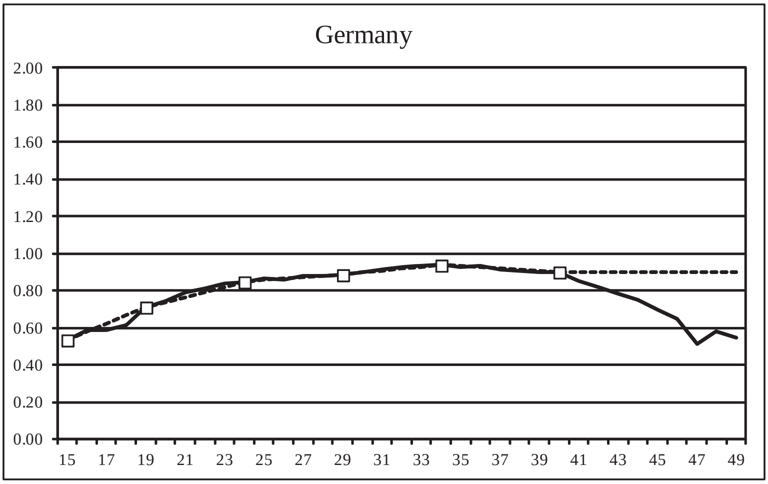
and Western countries, Denmark, the United Kingdom and France, have an above average TFR. But there are clear differences in the age pattern of fertility rates across these three countries. Whereas in Denmark fertility rates at young ages (under 23 years) are below the European average, those in the United Kingdom are significantly higher. The age patterns of Denmark and France are much more peaked than that of the United Kingdom. The peak in France is at slightly younger ages than the European average. The other three countries, Germany, Italy and Poland, have below average TFR. In Germany the age-specific fertility rates at all ages are below the European average, whereas in Italy the fertility rates at ages 32 or under are lower than the European average, but at older ages they are higher. In Poland fertility rates at younger ages are relatively high, but at ages 27 or above are very low.

Figure 5.3 shows the rate ratios: that is, the ratios of the age-specific fertility rates of these six countries to the EU27+3 average. The solid lines show clearly that the differences between the age-specific fertility rates of individual countries and the European average vary by age. Thus in addition to differences in the overall level of fertility, as reflected by differences in the TFR, there are differences in the age pattern. Usually differences in the age pattern are described by the mean age at childbearing. However, differences in the mean age do not capture all differences in the age pattern. The average age at childbearing of the EU27+3 countries equals 29.2 years. If in a country the mean age at childbearing is higher than the European average this can be caused by low fertility rates at young ages (*e.g.*, Italy), by a relatively high peak of fertility around age 30 (Denmark), or by high fertility rates at older ages (France). A low level of the mean age can be caused by high fertility rates at young ages (*e.g.*, the United Kingdom) or low fertility rates at older ages (Poland).

The linear splines shown in the figure are fitted by calculating equation (11). The knots are selected by applying a non-linear least squares method to data for the 30 European countries in this study. For all countries the same knots are selected. This makes cross-country comparisons easier. Linear splines based on minimum age 15 years and 5 knots (ages 19, 24, 29, 34 and 40 years) turn out to fit the observed rate ratios accurately. *Table 5.1* shows the values of the rate ratios at these ages for all 30 countries. The root mean square error (RMSE) equals 2.83×10^{-3} (see *table 5.2*). If four knots are selected the fit is clearly less accurate: the RMSE equals 3.20×10^{-3}. However, for some countries four knots would be sufficient. The selection of a knot at age 19 is needed to capture the relatively high level of fertility at young ages in countries such as the United Kingdom and Poland, but it

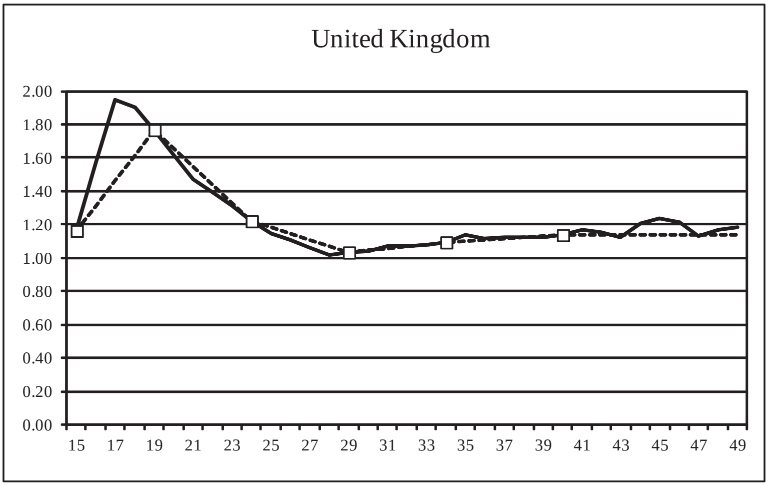
Figure 5.3. Rate ratios of age-specific fertility rates of six European countries and EU27+3 average, 2008

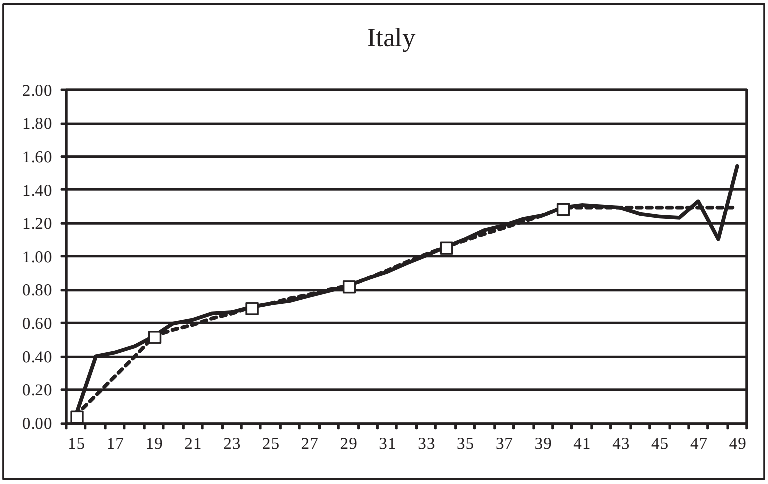




Solid line: observed values; dotted line: linear spline; squares: values at knots.

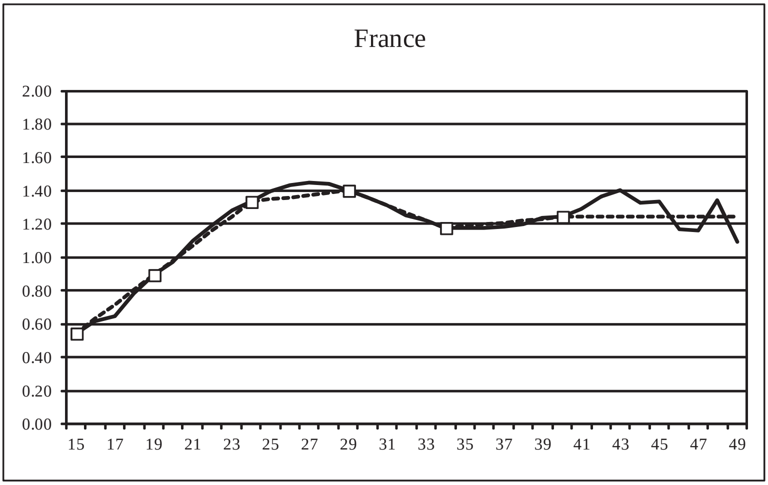
Figure 5.3. Rate ratios of age-specific fertility rates of six European countries and EU27+3 average, 2008 (continued)

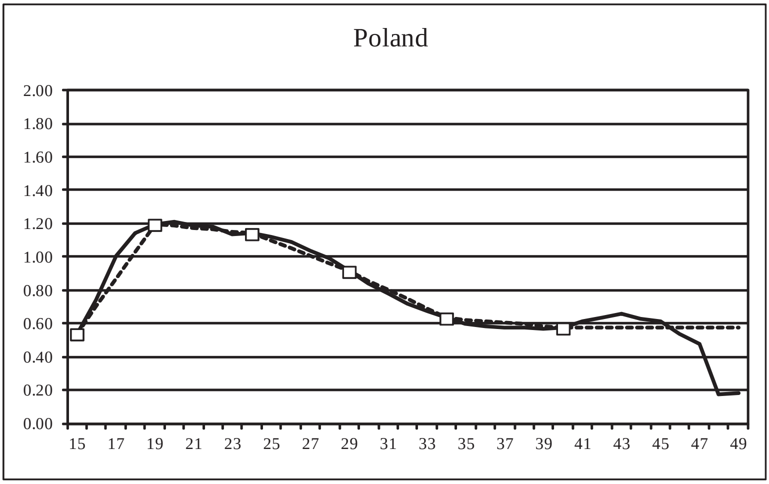




Solid line: observed values; dotted line: linear spline; squares: values at knots.

Figure 5.3. Rate ratios of age-specific fertility rates of six European countries and EU27+3 average, 2008 (end)





Solid line: observed values; dotted line: linear spline; squares: values at knots.

Table 5.1. Values of the rate ratios of age-specific fertility rates of European countries and EU27+3 average at knots, Total Fertility Rate (TFR) and mean age at childbearing (MAC), 2008

	15	19	24	29	34	40	TFR	MAC
Austria	0.57	0.85	0.99	0.91	0.85	0.81	1.41	29.0
Belgium	0.51	0.87	1.24	1.35	0.95	0.71	1.76	28.8
Bulgaria	8.52	2.03	1.30	0.76	0.53	0.37	1.47	26.1
Cyprus	0.36	0.64	0.92	1.01	0.94	0.92	1.46	29.7
Czech Republic	0.44	0.85	1.00	1.15	0.82	0.63	1.50	28.8
Denmark	0.10	0.56	1.05	1.44	1.34	1.18	1.89	29.9
Estonia	1.28	1.45	1.26	0.94	0.89	0.98	1.65	28.3
Finland	0.12	0.84	1.18	1.24	1.20	1.27	1.85	29.6
France	0.54	0.89	1.33	1.40	1.18	1.24	2.00	29.4
Germany	0.54	0.71	0.85	0.89	0.94	0.90	1.38	29.6
Greece	1.45	0.83	0.89	0.98	1.03	0.98	1.51	29.6
Hungary	2.13	1.12	0.83	0.90	0.74	0.63	1.35	28.4
Iceland	0.00	1.26	1.46	1.40	1.22	1.46	2.15	29.3
Ireland	0.58	1.35	0.95	1.03	1.74	2.25	2.10	30.7
Italy	0.04	0.52	0.69	0.83	1.06	1.29	1.37	30.5
Latvia	1.06	1.62	1.30	0.80	0.63	0.72	1.44	27.6
Lithuania	0.91	1.34	1.27	0.97	0.64	0.51	1.47	27.7
Luxembourg	0.00	0.73	0.88	1.03	1.11	1.06	1.61	30.0
Malta	1.50	1.01	0.96	1.00	0.85	0.56	1.44	28.7
Netherlands	0.25	0.43	0.93	1.30	1.36	0.97	1.77	30.2
Norway	0.17	0.83	1.32	1.35	1.26	1.05	1.96	29.4
Poland	0.54	1.20	1.14	0.91	0.63	0.57	1.39	28.0
Portugal	1.53	0.99	0.81	0.85	0.88	0.85	1.37	29.1
Romania	5.23	1.86	1.11	0.75	0.50	0.42	1.35	26.4
Slovakia	1.63	1.31	0.97	0.86	0.63	0.48	1.32	27.8
Slovenia	0.08	0.37	1.02	1.23	0.91	0.67	1.53	29.4
Spain	1.06	0.89	0.67	0.79	1.19	1.24	1.46	30.3
Sweden	0.18	0.55	1.14	1.32	1.35	1.29	1.91	30.1
Switzerland	0.19	0.39	0.78	0.97	1.19	1.15	1.48	30.5
United Kingdom	1.16	1.77	1.22	1.03	1.09	1.14	1.84	28.7
EU27 + 3 average							1.58	29.2

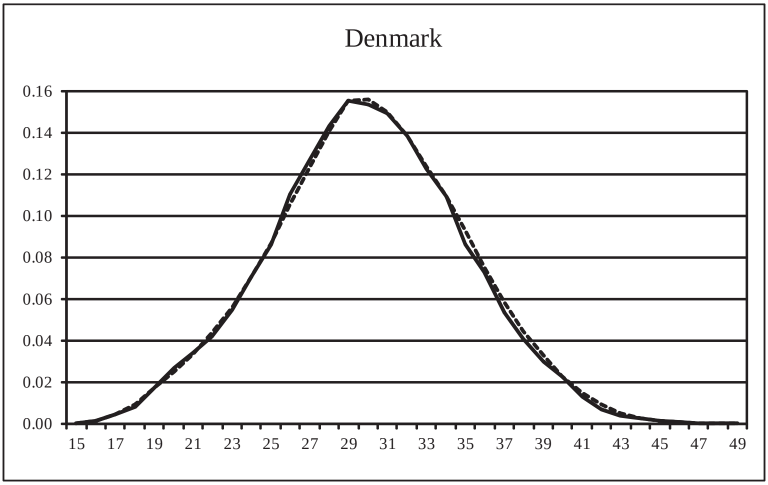
* For Italy data refer to 2007, for the United Kingdom to 2006 and for Belgium to 2005.

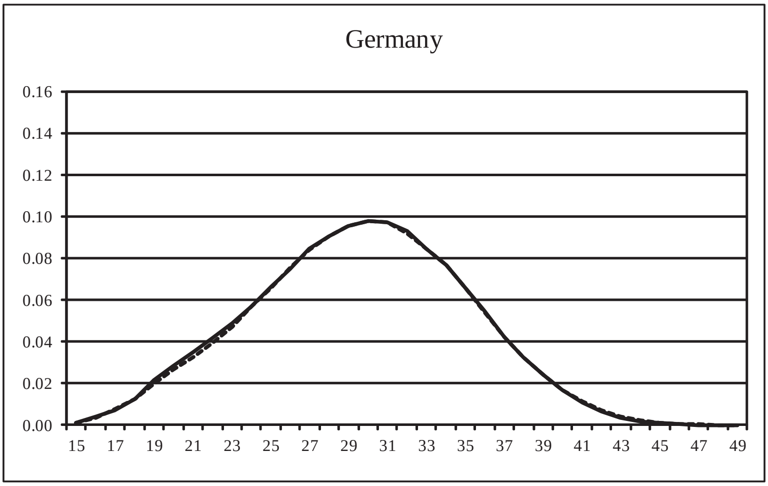
would not be needed for fitting the age schedule in, for example, Denmark. Similarly a knot at age 40 would not be needed for providing a good fit for the United Kingdom or France. Thus, when fitting a spline for a single country, the minimum number of knots needed for providing an accurate fit may be fewer than five. For 13 out of the 30 European countries the fit of the model including four knots –that is, without a knot at age 19– would be about equal to that of the model including five knots. Adding knots obviously improves the fit of the model. However, adding a sixth knot turns out to produce only a slight improvement of the fit. Adding a knot at age 27 yields a RMSE of 2.60 x 10^{-3}. Since a large number of knots with small age intervals in between hampers the usefulness of the model for creating scenarios, it was decided to apply the model including five knots.

Multiplying the linear splines shown in figure 5.3 by the (weighted) EU27+3 average fertility rates shown in figure 5.1 produces the TOPALS fertility age schedule for each country. *Figure 5.4* shows that TOPALS is capable of describing different age patterns of fertility, each based on the same standard age schedule. The model is capable of describing the relatively high fertility at young ages in the United Kingdom and Poland. The figures in Annex A shows the fit of TOPALS for the other 24 European countries. For most countries the fit is satisfactory. One exception is Bulgaria, where the fitted curve is not smooth at young ages. This is caused by the fact that the rate ratios for Bulgaria at young ages do not exhibit a linear shape between ages 15 and 19. This implies that an additional knot between these ages would be needed to obtain a smooth curve. If a knot at age 17 is added the fit improves considerably. Furthermore the fit for Belgium and Ireland can be improved by adding a knot at age 27.

It is interesting to compare the results obtained using TOPALS with those obtained using other relational and parametric models. First, we compare the results of TOPALS with those of the three most frequently applied parametric models: the Hadwiger, Gamma and Beta functions. On average the Beta function produces a better fit than the other two models. This confirms the results obtained by Peristera and Kostaki (2007). Hoem *et al*. (1981) compared a great variety of methods for smoothing fertility age patterns by applying them to Danish fertility data in the 1970s. They concluded that the Gamma and Hadwiger functions performed better than the Beta distribution. Peristera and Kostaki (2007) note that the unsatisfactory fit of the Beta model by Hoem *et al*. (1981) may be due to the fact that they had problems in obtaining least squares estimates of the coefficients of the Beta distribution. Peristera and Kostaki (2007) suggest a new parametric model that looks like

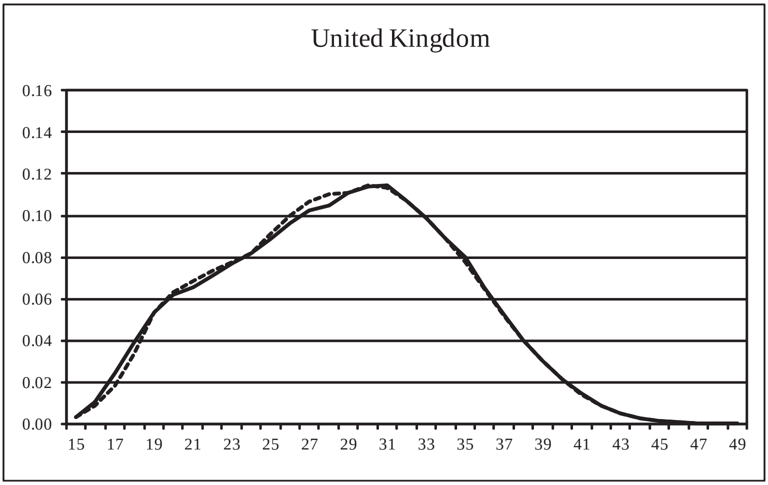
Figure 5.4. Age-specific fertility rates of six European countries and fit by TOPALS, 2008

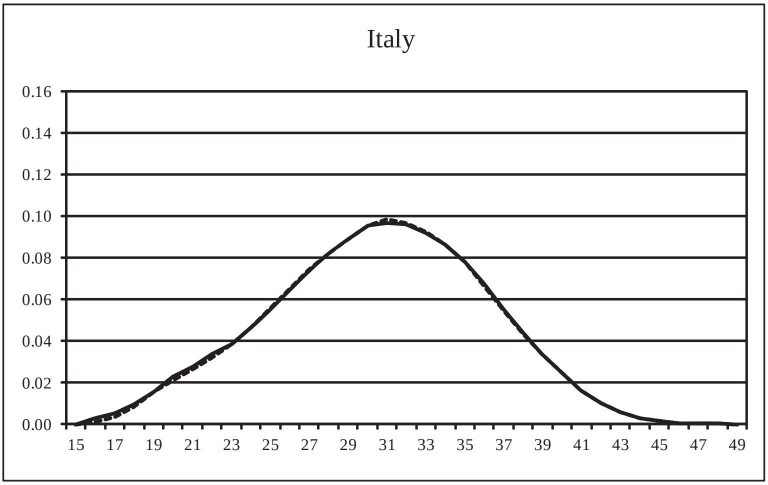




Solid line: observed values; dotted line: TOPALS.

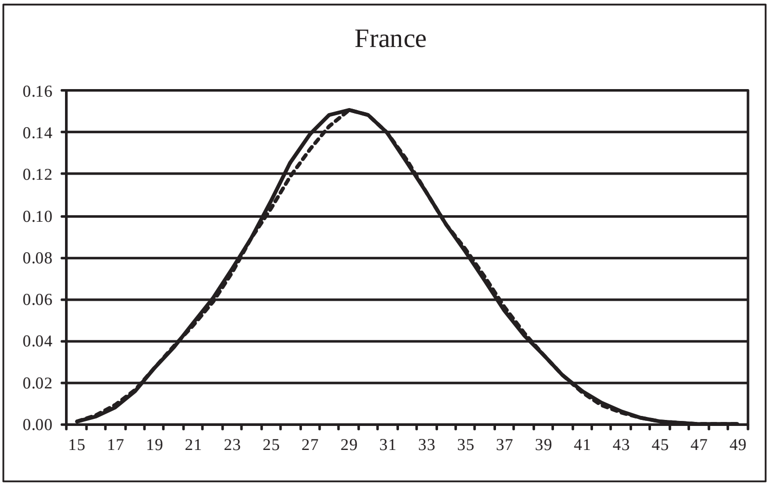
Figure 5.4. Age-specific fertility rates of six European countries and fit by TOPALS, 2008 (continued)

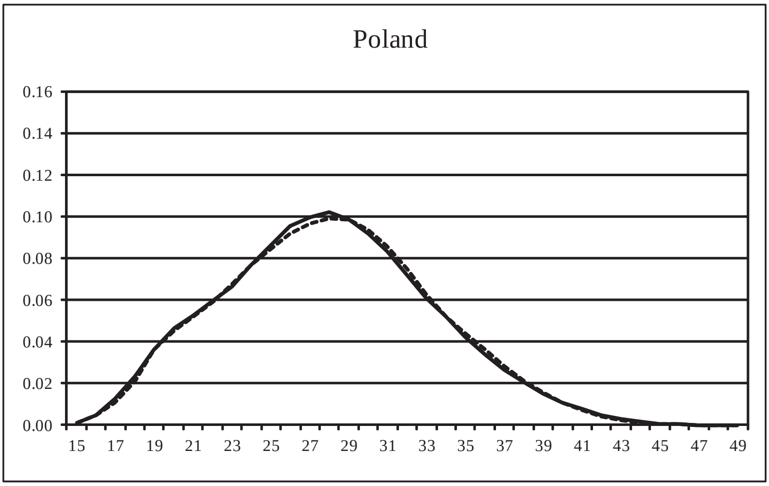




Solid line: observed values; dotted line: TOPALS.

Figure 5.4. Age-specific fertility rates of six European countries and fit by TOPALS, 2008 (end)





Solid line: observed values; dotted line: TOPALS.

Table 5.2. Goodness of fit (measured by Root mean square error) of age-specific fertility rates in 30 countries, 2008

	TOPALS	Hadwiger	Beta	Gamma	Peristera-Kostaki**	Schmertmann	Brass OLS	Brass Non linear LS
	RMSE (x 10^{-3})							
Average	**2.8**	7.9	5.9	6.4	3.2	3.6	6.6	3.5
Austria	**1.3**	5.7	3.6	3.8	2.1	2.6	3.8	1.4
Belgium	3.3	5.9	5.2	4.8	**3.2**	3.3	9.3	3.2
Bulgaria	6.1	6.2	5.7	4.2	2.5	3.6	6.0	**2.3**
Cyprus	3.7	4.9	3.8	3.9	**3.0**	3.4	8.9	3.3
Czech Republic	**2.0**	7.5	6.5	6.3	3.9	3.3	7.2	3.5
Denmark	**2.2**	6.3	4.6	6.5	2.7	3.4	4.5	**2.2**
Estonia	**3.4**	6.0	4.1	4.2	3.9	4.4	6.2	4.5
Finland	**1.9**	7.3	4.5	5.3	3.0	3.3	4.3	2.1
France	**2.1**	5.8	4.1	4.7	2.2	2.5	6.7	3.4
Germany	**0.5**	6.6	4.2	4.9	1.5	2.2	1.4	0.7
Greece	2.0	8.2	5.9	6.3	2.3	2.2	4.8	**1.8**
Hungary	**1.5**	8.8	7.6	7.0	3.8	2.9	4.2	3.4
Iceland	6.0	7.8	**5.3**	6.2	5.9	6.6	15.2	7.8
Ireland	4.0	18.6	14.3	16.2	**2.1**	6.4	15.7	8.1
Italy	**0.8**	7.0	4.2	5.6	1.5	2.2	3.0	0.9
Latvia	**2.3**	4.5	2.7	3.6	3.6	4.1	6.2	4.2
Lithuania	**2.9**	5.3	4.9	3.7	3.6	3.0	4.4	3.5
Luxembourg	4.6	8.3	6.2	6.8	4.1	3.9	6.5	**3.6**
Malta	3.9	8.3	7.4	7.0	4.7	**3.8**	5.0	4.8

	TOPALS	Hadwiger	Beta	Gamma	Peristera-Kostaki**	Schmertmann	Brass OLS	Brass Non linear LS
	RMSE (x 10^{-3})							
Netherlands	1.7	6.8	4.0	6.0	**1.1**	2.0	4.5	1.4
Norway	**1.9**	7.1	3.9	5.6	2.5	3.0	3.3	2.3
Poland	**1.5**	5.1	4.0	3.4	3.4	2.7	5.2	2.0
Portugal	**1.3**	8.8	6.8	6.9	4.1	2.9	3.1	2.1
Romania	2.5	6.5	5.7	5.1	4.1	3.9	4.7	**1.5**
Slavakia	**1.4**	7.5	6.4	5.6	3.6	2.9	6.3	2.1
Slovenia	**2.0**	5.2	4.0	4.0	3.0	3.0	6.7	2.4
Spain	2.1	11.8	9.2	10.3	**0.8**	3.3	7.5	5.0
Sweden	2.5	7.9	4.9	6.5	3.0	3.0	2.8	**2.1**
Switzerland	**1.1**	6.8	3.8	5.4	1.7	1.9	1.8	1.6
United Kingdom	2.2	11.3	8.1	8.7	**2.1**	5.9	4.9	3.2
Number of parameters	5	4	3	4	4	4	2	2

* For Italy data refer to 2007, for the United Kingdom to 2006 and for Belgium to 2005.

** Model 2 applied to Ireland, Spain and the United Kingdom.

Bold: Minimum value at each row

the normal distribution but allows the slope below and above the mode to be different. In case there is high young-age fertility, they extend the model (see Equation 5). *Table 5.2* shows that the Peristera-Kostaki model performs better than the other parametric models. Whereas the Peristera-Kostaki model provides a better fit than TOPALS for 9 of the 30 European countries, for most other countries TOPALS provides a better fit. On average the fit of TOPALS is slightly better than that of the Peristera-Kostaki method. The fit of the Peristera-Kostaki models is about the same as that of TOPALS using four rather than five knots.

Second, we compare the results of TOPALS with those of Schmertmann (2003) based on quadratic splines. Table 5.2 shows that although Schmertmann's method produces a fairly accurate fit, it does not outperform TOPALS or the Peristera-Kostaki model.

Third, we compare the results of TOPALS with the Brass relational model using the same standard age schedule. Brass suggested estimating the parameters by OLS. This minimizes the differences between the double log transformations of the estimated and the observed fertility rates. Table 5.2 shows that this does not produce an adequate fit with the observed fertility rates. The RMSE equals 6.61 x 10^{-3}. For that reason we estimated the parameters of the Brass model by non-linear least squares, which minimizes the differences between the estimated and observed age-specific fertility rates rather than the double log transformations. This improves the fit considerably: the RMSE decreases to 3.48 x 10^{-3}. The Brass model produces a very accurate fit for Denmark, Germany, and Italy. However, the method is not completely capable of describing the fertility hump at young ages in the United Kingdom, Ireland and Hungary. Table 5.2 shows that for 3 countries the Brass model clearly produces a better fit than TOPALS, whereas for 13 countries the TOPALS model clearly outperforms Brass.

Note that the goodness of fit is a necessary but not sufficient reason for selecting a model. Another important criterion is the interpretation of the parameters. The parameters of the Brass model lack an intuitive interpretation. For that reason Zeng Yi *et al*. (2000) propose a different method for determining the values of the parameters of the Brass model which can be interpreted (see Section 5.2). However, it turns out that the fit is worse than that of the Brass model estimated by non-linear least squares. The Zeng Yi procedure is very sensitive to the shape of the standard age schedule. For example, the Zeng Yi method produces a good fit for Germany and Italy, with the age pattern of fertility fairly similar to the EU27+3 average, as figure 5.2 shows (apart from a difference in the mean age at childbearing), but a poor fit for Denmark and the United Kingdom, with the age pattern having a different shape.

The parameters of TOPALS shown in table 5.1 can be interpreted easily. High values of the rate ratios at age 19 indicate relatively high fertility rates at young ages. Table 5.1 shows that 11 European countries have high fertility at young ages. Of these countries eight are Eastern European. The others are the United Kingdom, Ireland and Iceland. High values of the rate ratios at age 34 indicate that fertility at older ages is relatively high. Most countries in Northern, Western and Southern Europe have high values of the rate ratios at age 34. High values of the rate ratios at age 29 indicate that the age pattern of the fertility rates is peaked. In most cases a high value at age 29 goes together with a high value of the TFR. Ten countries have high values of the rate ratio (exceeding 1.1) at age 29: that is, around the peak age. These are mainly countries in Northern and Western Europe. Most of these countries have an above average value of the TFR. Usually age-specific fertility rates are characterized by the level of the TFR and the mean age at childbearing. However, as noted above the mean age does not capture all differences in the age pattern. Usually low values of the mean age at childbearing go together with high fertility rates at young ages. But this is not the case in Iceland and Ireland, where fertility at young ages is relatively high but the mean age at childbearing is above the European average. On the other hand, a high value of the mean age can be caused by high values of fertility rates at ages above 30 but by low values at young ages as well. For example, Slovenia and Switzerland have above average mean age due to very low fertility rates at young ages.

5.5. Scenarios

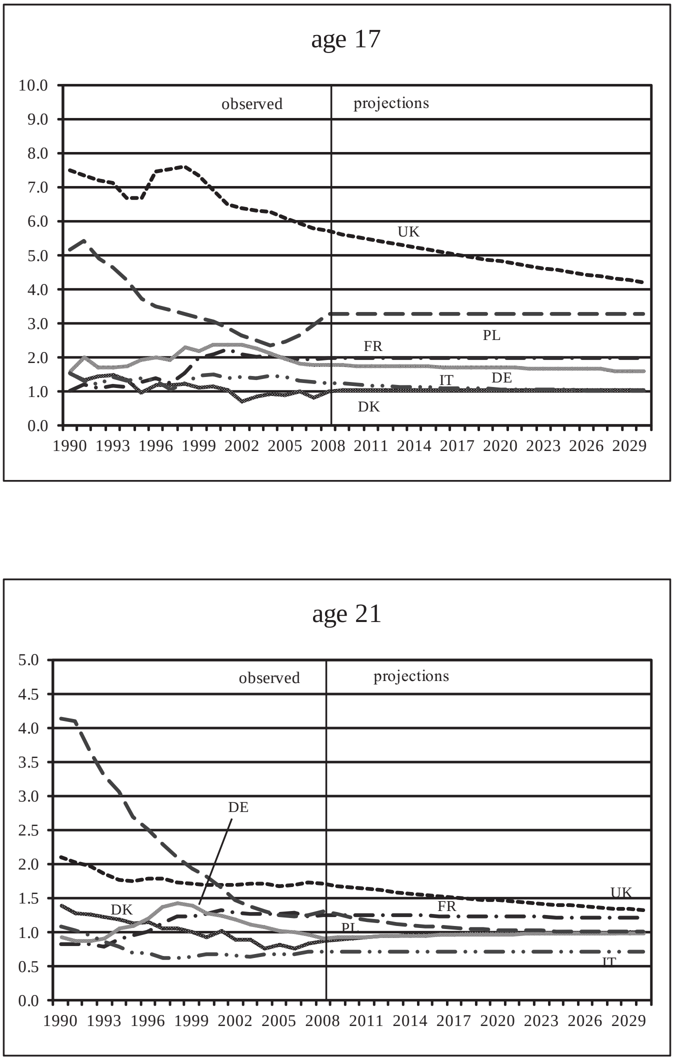
We illustrate the use of TOPALS for making scenarios by specifying assumptions about the future values of the rate ratios for the six countries discussed in the preceding section. We demonstrate two approaches. First, we compare the age-specific fertility rates of the six countries with those of one ‘forerunner’ country, Sweden. We project the rate ratios of the fertility rates of the six countries compared with those of Sweden into the future under the assumption that the six countries will move in the direction of the Swedish pattern. We estimate a partial adjustment model to assess the speed of the converging trend. The second approach uses the rate ratios compared with the EU27+3 average shown in table 5.1 as its starting point and makes assumptions about how the age patterns of the rate ratios may change in the future. The former approach is more objective, since the tendency towards the target age schedule is determined by the estimated parameter of the partial adjustment model, whereas the latter approach is based on more or less ‘subjective’ assumptions about the future values of the rate ratios.

5.5.1. Projections based on time series model

Lanzieri (2010) shows that during the last decades there has been a converging trend in fertility across EU countries, even though there have been periods of divergence. In the latest population scenarios, EUROPOP2008, Eurostat assumes that fertility will converge to levels achieved by EU member states that are considered as demographic forerunners (Giannakouris, 2008 and Lanzieri, 2009). However, it is assumed that convergence will not be reached until 2150. This implies that in the last year of the projection period, 2060, no complete convergence will be reached yet. The Eurostat scenarios are based on linear interpolation between 2008 and 2150. As a consequence, according to the Eurostat scenarios for each country, the difference in the TFR with the EU average will decline by one-third between 2008 and 2060. In 2150 the European average of the TFR is assumed to increase to 1.85, the current level of Sweden.

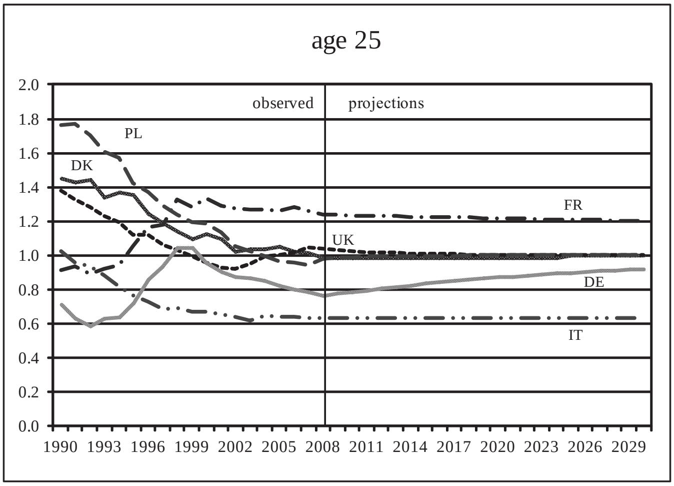
Rather than making an a priori assumption about the year when convergence will be reached, we will project the speed of convergence on the basis of an analysis of past trends. Both Eurostat and Statistics Sweden hardly expect any future change in Swedish fertility rates (apart from short-run fluctuations due, for example, to business cycles) (Statistics Sweden, 2009 and Lanzieri, 2009). Thus if we consider Sweden as the forerunner country, we can take the current Swedish fertility rates as the ‘target’ towards which the fertility rates of the other countries will move. Note that the assumption that the Swedish age-specific fertility rates will hardly change in the future implies that the current age schedule of fertility is very close to the cohort age schedule for young cohorts. To make such a convergence scenario we calculate rate ratios by dividing the fertility rates of the six countries under study by the Swedish age-specific fertility rates in 2008. Since we want to project a smooth age pattern, we use the Swedish age-specific fertility rates that are smoothed by TOPALS. In specifying linear splines, it turns out that for achieving a good fit for the six countries the knots of the rate ratios compared with the Swedish age pattern differ slightly from those compared with the EU27+3 average. The knots are at ages 17, 21, 25, 29, 34 and 40. *Figure 5.5* shows the time series of the rate ratios for the period 1990–2008 for these ages. Note that the rate ratios for all years are calculated by dividing the fertility rates of the six

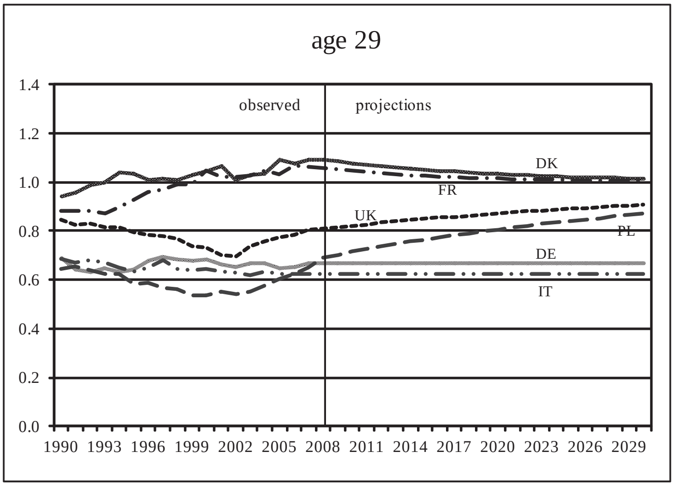
Figure 5.5. Rate ratios of age-specific fertility rates of six European countries compared with Sweden in 2008, observations 1990-2008 and projections 2009-2030



For United Kingdom and Italy: observations 1990-2007 and projections 2008-2030.

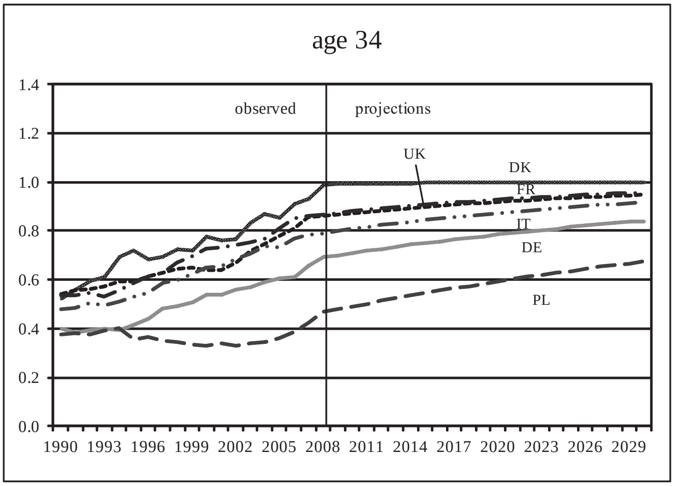
Figure 5.5. Rate ratios of age-specific fertility rates of six European countries compared with Sweden in 2008, observations 1990-2008 and projections 2009-2030 (continued)

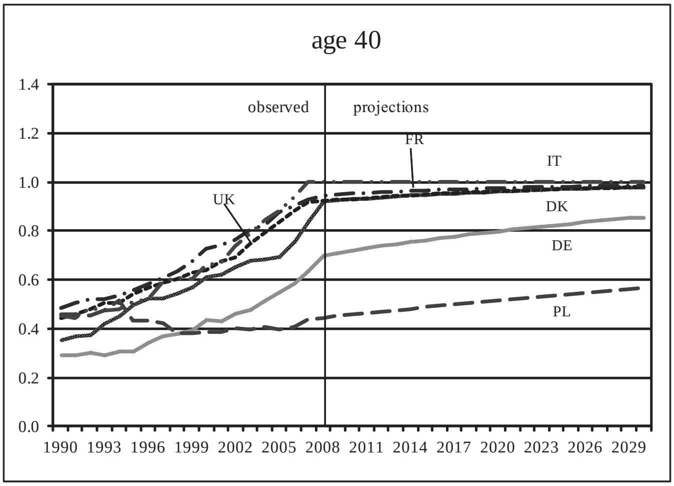




For United Kingdom and Italy: observations 1990-2007 and projections 2008-2030.

Figure 5.5. Rate ratios of age-specific fertility rates of six European countries compared with Sweden in 2008, observations 1990-2008 and projections 2009-2030 (end)





For United Kingdom and Italy: observations 1990-2007 and projections 2008-2030.

countries by the Swedish fertility rates for 2008, as these rates are considered to be the target values.

We model the time series of rate ratios as a partial adjustment model, assuming that the rate ratios move towards 1:

$$r(x)_t - 1 = \varphi[r(x)_{t-1} - 1] + e_t \tag{13}$$

where $r(x)_t$ is the rate ratio in year t, $0 \leq \varphi \leq 1$ and e_t is a random term with $E(e_t) = 0$. This model assumes that the value of $r(x)_t$ is closer to 1 than the value of $r(x)_{t-1}$. *Figure 5.6* shows how rapidly the values of $r(x)$ move towards 1 for different values of φ, starting from a value of 2 and 0.5 in 2008 respectively. The lower the value of φ, the quicker $r(x)_t$ will move towards 1. If φ is close to 1, $r(x)_t$ moves slowly to 1. If $\varphi = 1$ model (13) describes a random walk, and $r(x)_t$ does not converge to 1. Figure 5.6 shows that if $\varphi = 0.98$ the difference of the rate ratio with 1 will be halved in the year 2042, whereas if $\varphi = 0.95$ the difference with 1 will be halved in 2022.

Since $E(e_t) = 0$ projections of model (13) can be calculated by:

$$\hat{r}(x)_{t+1|t} = \varphi\ r(x)_t + 1 - \varphi \tag{14}$$

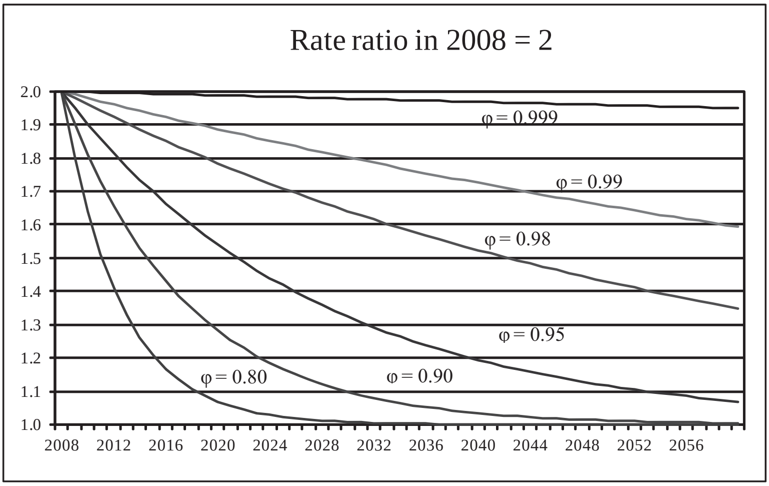
where $\hat{r}(x)_{t+1|t}$ is the projection of $r(x)_{t+1}$ based on observations up to year t. If $\varphi = 1$, the projected value of $r(x)_{t+1}$ equals the last observed value, similarly to the projections of a random walk model. Since $0 \leq \varphi \leq 1$, in the long run the projections will move to:

$$\hat{r}(x)_{t+T|t} = \varphi^T r(x)_t + 1 - \varphi^T \tag{15}$$

Thus if $\varphi < 1$ the projections will move to 1 for large T.

For each age and for each country we estimated the value of φ for the period 1990-2008 by OLS using equation (13). *Table 5.3* shows the estimated values of φ. For ages 34 and 40 the values of φ are below 1 for all countries. This indicates that there is a clear tendency towards the Swedish levels of fertility rates of women in their 30s. For Poland ages 21 and 25 have relatively low values of φ. This indicates that the fertility rates at young ages rapidly move in the direction of the Swedish levels. In contrast, for age 17 $\varphi = 1$ and thus there is no convergence towards 1. In most other countries values of φ exceed

Figure 5.6. Values of rate ratios for different values of φ



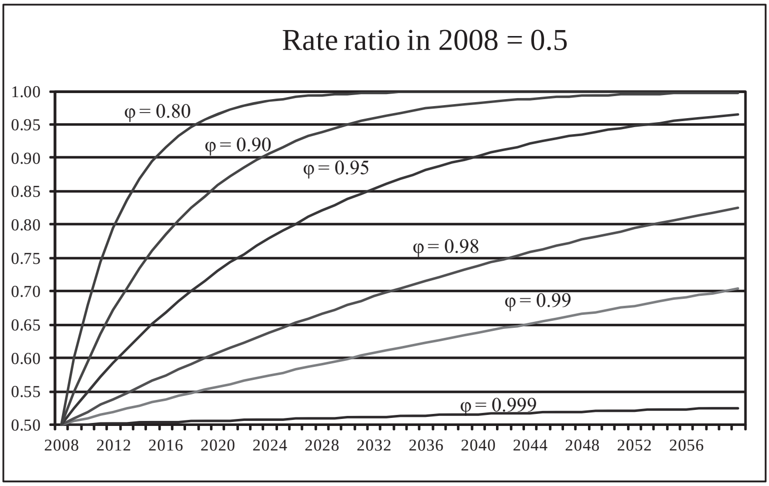


Table 5.3. Estimated values of parameter φ of partial adjustment model

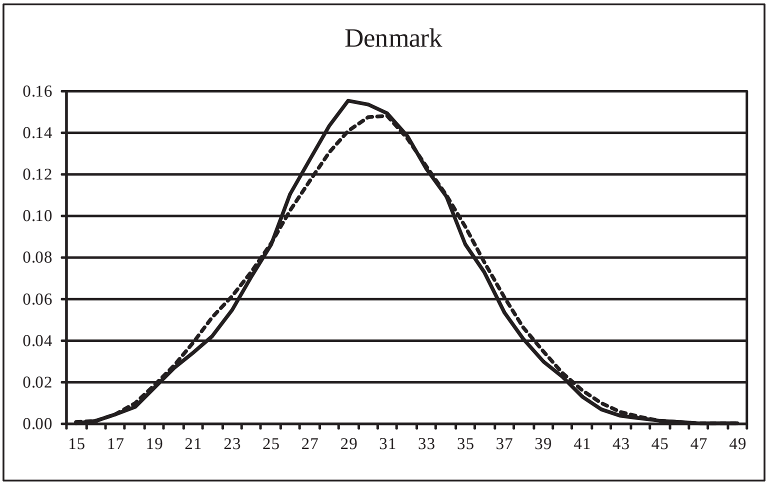
	age 17	age 21	age 25	age 29	age 34	age 40
Denmark	0.67	0.82	0.89	0.92	0.92	0.94
United Kingdom	0.98	0.97	0.82	0.97	0.96	0.94
France	1.00	0.99	0.99	0.89	0.95	0.94
Germany	0.99	0.93	0.95	1.00	0.97	0.97
Italy	0.89	1.00	1.00	1.00	0.96	0.94
Poland	1.00	0.84	0.75	0.96	0.98	0.99

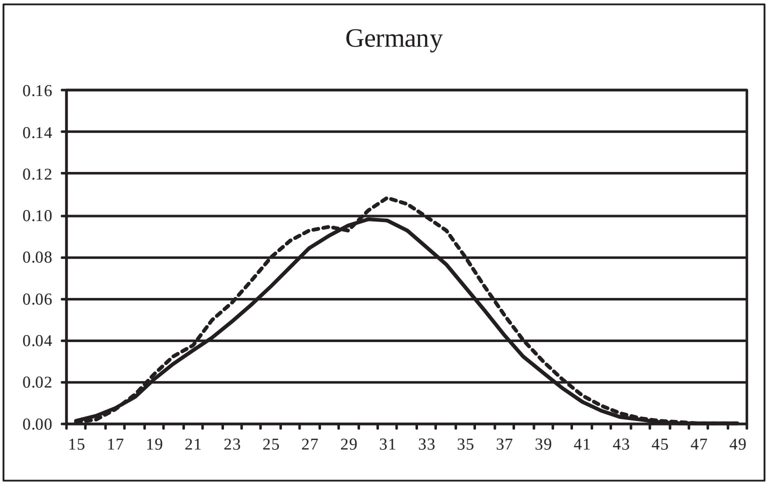
.90, indicating that for most countries it will take several decades before the current level of the Swedish age-specific fertility rates is reached.

Using the estimated values of φ, equation (14) is used to make projections of the rate ratios to 2030. Figure 5.5 shows that most rate ratios move towards 1, but that the speed varies strongly across ages and across countries. For example, for Poland the rate ratios of women in their 30s show an increase, but it will take many years before they will reach the value of 1. For Polish women at age 29 the increase towards 1 is more rapid.

The rate ratios for the year 2030 are used to make projections of the age-specific rates for the six countries. *Figure 5.7* shows that for the three low-fertility countries an increase of fertility rates in women in their late 20s and 30s is projected. Figure 5.7 shows that the projected age pattern for Denmark is less peaked than the observed pattern in 2008. This can be explained by the fact that the current age pattern for Sweden is less peaked than the Danish pattern. If this is considered implausible, one alternative would be to use the age pattern of another country as the standard age schedule: so, for example, the age pattern of fertility in the Netherlands is more peaked than that in other European countries. Below we will show an alternative scenario. Furthermore figure 5.7 shows that the projected age patterns are not smooth for all countries, particularly at young ages for the United Kingdom and Poland, and around the peak age for Germany. The reason why the curve for the United Kingdom is not very smooth is that the age pattern of fertility there at young ages differs strongly from that in Sweden. As a consequence the rate ratios do not show a linear pattern at young ages and thus a linear spline does not produce a very accurate fit. An accurate fit would require additional knots, closer to each other. For Germany

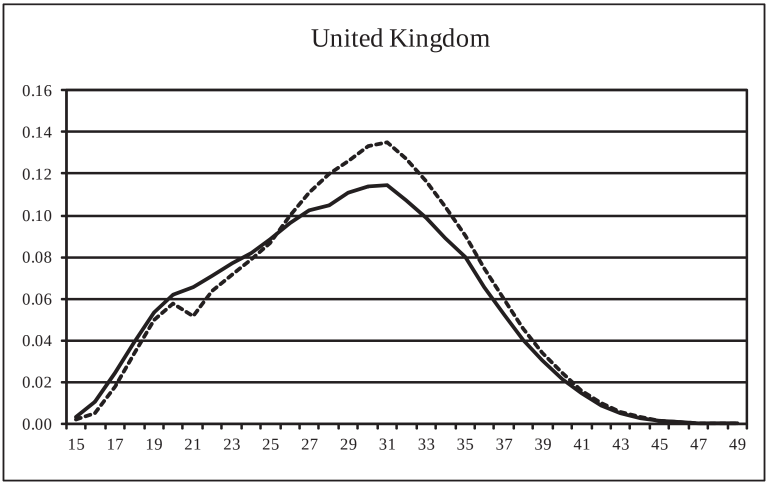
Figure 5.7. Age-specific fertility rates of six European countries, 2008, and scenario for 2030 based on projections by partial adjustment model of rate ratios compared with Swedish age-specific fertility rates

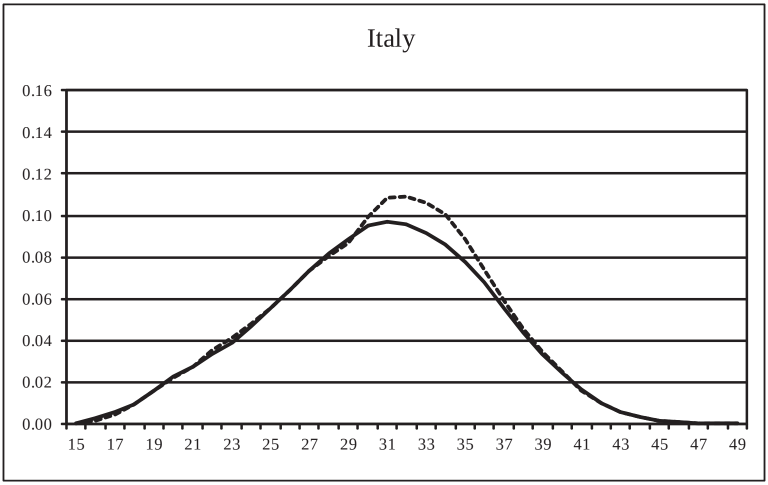




Solid line: observed values; dotted line: projections.

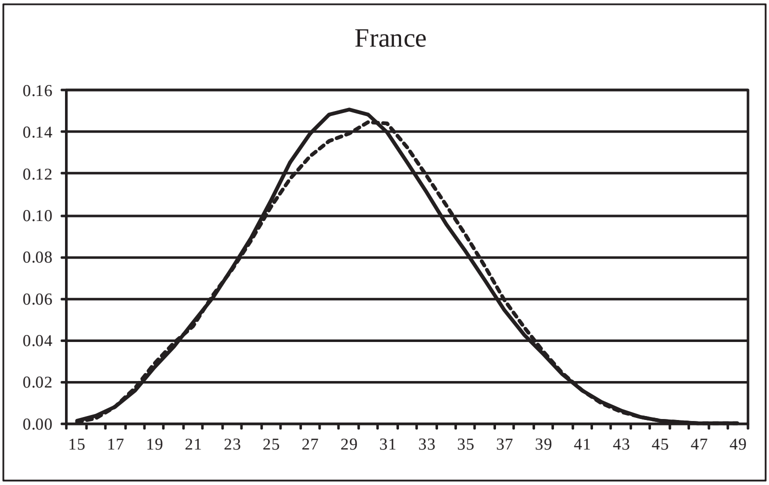
Figure 5.7. Age-specific fertility rates of six European countries, 2008, and scenario for 2030 based on projections by partial adjustment model of rate ratios compared with Swedish age-specific fertility rates (continued)

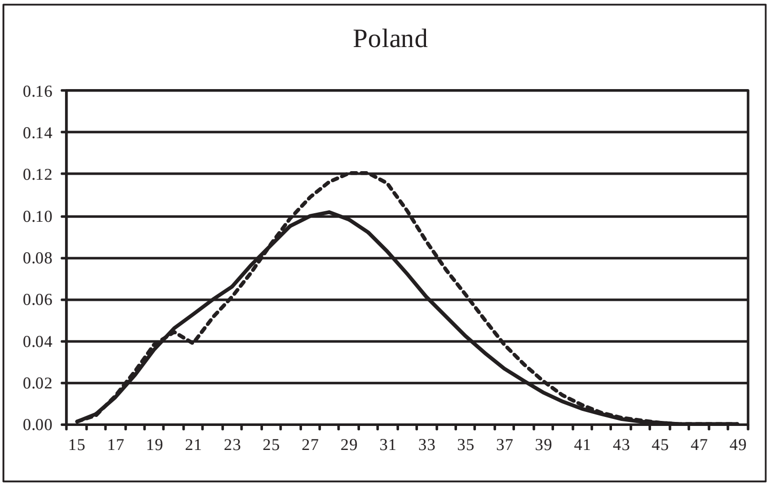




Solid line: observed values; dotted line: projections.

Figure 5.7. Age-specific fertility rates of six European countries, 2008, and scenario for 2030 based on projections by partial adjustment model of rate ratios compared with Swedish age-specific fertility rates (end)





Solid line: observed values; dotted line: projections.

the reason why the projected age pattern is not smooth around age 29 is that the projections of the rate ratios for knots next to each other differ strongly. Table 5.3 shows that the values of φ differ between ages 25, 29 and 34. If we were to replace the value of φ for age 29 by a value somewhere between the values for ages 25 and 34, this would produce a smooth age pattern. For the Polish fertility rates at young ages the same explanation applies: the values of φ for ages 17 and 21 differ markedly. Below we will show alternative scenarios that exhibit smooth age patterns.

Table 5.4 compares the values of the TFR resulting from these projected age-specific fertility rates with the EUROPOP2008 scenario given by Eurostat. Even though, mirroring our assumption, Eurostat assumes convergence towards the Swedish fertility level, our projections of the TFR exceed those of Eurostat. One explanation is that Eurostat assumes a linear change in the direction of the Swedish level, whereas the partial adjustment model that we use implies a non-linear change, as shown by figure 5.6. Another important difference between our projections and the Eurostat scenarios is that whereas Eurostat assumes Polish fertility will be lower than that of the other two low-fertility countries, Germany and Italy, our projection for Poland exceeds those for the other two countries. The main explanation is that the projected rate ratio at the peak age (29) increases more strongly for Poland than for Germany and Italy.

These scenarios show that projecting the rate ratios for the age at each knot separately may lead to age patterns that are not very smooth if rate ratios

Table 5.4. Total Fertility Rate in 2008 and scenarios for 2030

	2008	2030		
		TOPALS		EUROPOP 2008
		rate ratios compared with		
		Sweden	EU27+3 average	
Denmark	1.89	1.90	1.92	1.85
United Kingdom	1.84	1.96	1.94	1.84
France	2.00	2.01	2.09	1.96
Germany	1.38	1.57	1.55	1.42
Italy	1.37	1.45	1.56	1.46
Poland	1.39	1.62	1.57	1.36

United Kingdom and Italy: TFR in 2007 instead of 2008.

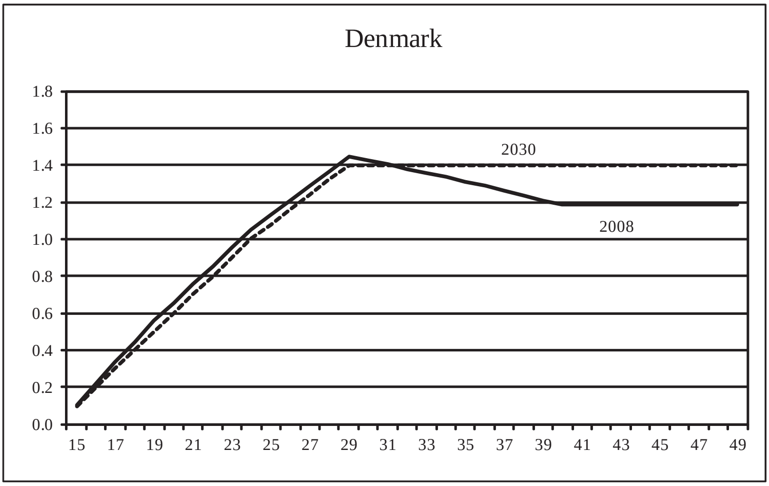
at neighbouring knots move in different directions. Instead of extrapolating values of rate ratios on the basis of time series analyses, one may make qualitative assumptions about the future values of the rate ratios assuming a smooth age pattern.

5.5.2. Scenarios based on qualitative assumptions

When specifying scenarios of future fertility it is important to identify the main determinants underlying past trends in fertility in order to assess to what extent the trends may be expected to continue. One main trend in fertility across Europe has been the postponement of fertility (Kohler, Billari and Ortega, 2002 and Frejka and Sobotka, 2008). While Northern countries seem to be at the last stage, other countries are at earlier stages (De Beer, 2006a and Frejka and Sobotka, 2008). Billari and Kohler (2004) regard cultural changes (such as secularization and individualism), the rise in the education of women, and the uncertainty during political changes in Eastern Europe as the main causes of the postponement. Goldstein (2006) argues that the biological upper age limits of fertility have not yet been reached by far and, consequently, postponement of fertility can continue for decades. Even though to some extent postponement can lead to a decline in the ultimate level of fertility of young cohorts due to the increase of infertility with age, Lanzieri (2009) notes that there is more or less a shared opinion that the catching up of postponed fertility will lead to a rise in the total fertility rate (Bongaarts, 2002; Sobotka, 2004 and De Beer, 2006a). On the basis of an analysis of recent fertility data Goldstein, Sobotka, and Jasilioniene (2009) conclude that the postponement of fertility “has begun to run its course”. As a consequence they expect a rise of the total fertility rate in the coming decades. Frejka *et al*. (2008) argue that in the foreseeable future postponement of childbearing to older ages will continue. They assume that in Northern and Western Europe fertility will be maintained close to the replacement level. In Southern, Central, and Eastern Europe they expect that some increase in fertility rates may occur, but they assume that fertility will remain well below the replacement level. Frejka and Sobotka (2008) mention various explanations of this divide, including cultural differences and differences in family policies.

Figure 5.8 shows the linear splines describing the age patterns of the rate ratios of the six countries under study in 2008. The solid lines in figure 5.8 are the same as the dotted lines in figure 5.2. Assuming that the postponement process has not yet reached its end, one would expect that the rate ratios of women in their 30s will increase. Figure 5.8 shows a possible scenario for the future values of the rate ratios. In this scenario we assume that at older

Figure 5.8. Linear splines of rate ratios of age-specific fertility rates of six European countries and EU27+3 average, 2008 and 2030



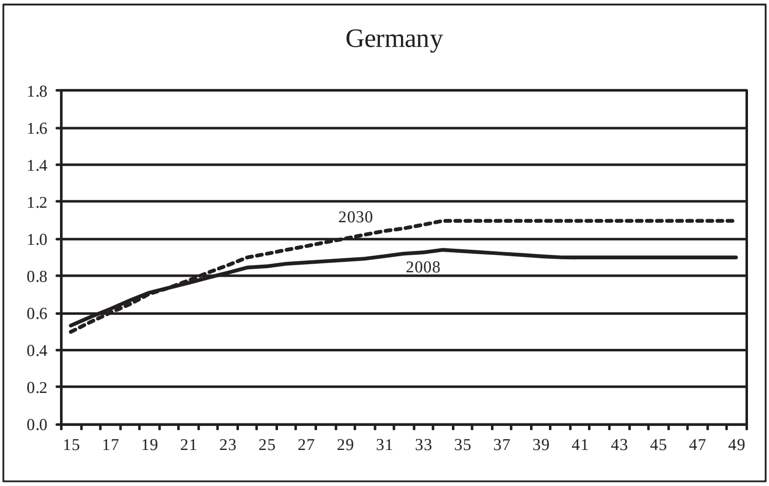
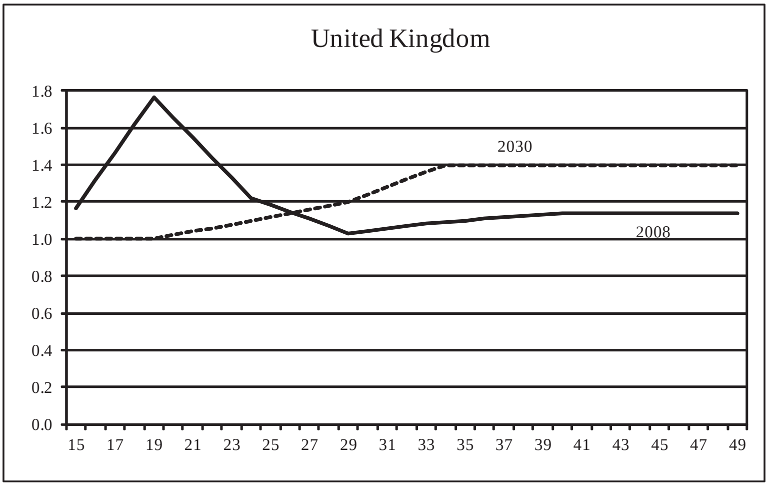


Figure 5.8. Linear splines of rate ratios of age-specific fertility rates of six European countries and EU27+3 average, 2008 and 2030 (continued)



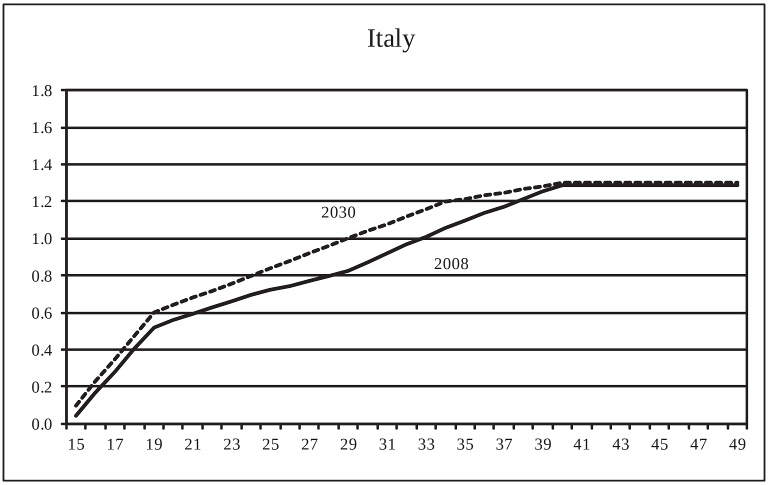
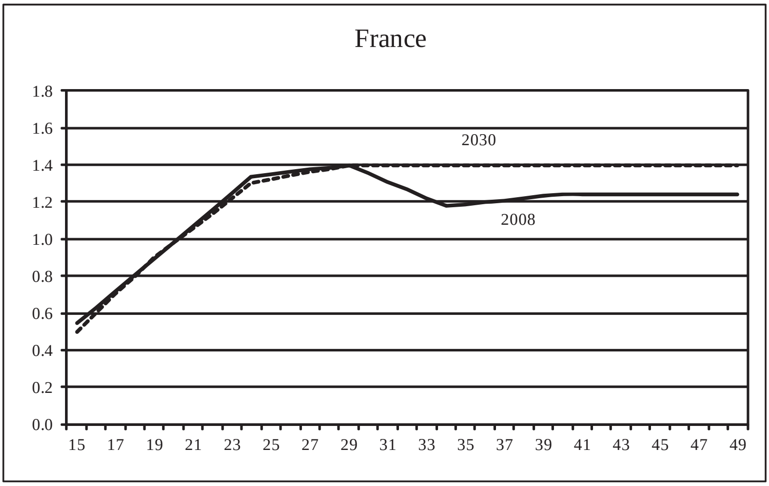
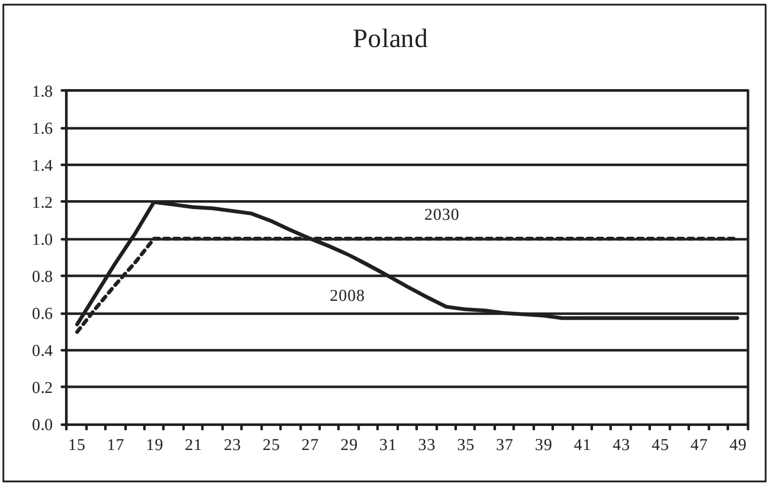


Figure 5.8. Linear splines of rate ratios of age-specific fertility rates of six European countries and EU27+3 average, 2008 and 2030 (end)





ages the fertility rates in Northern and Western Europe will remain higher than in Southern, Central, and Eastern Europe. Thus for Denmark, France, and the United Kingdom we will assume higher values of the rate ratios of women in their late 20s and 30s than for the three other countries.

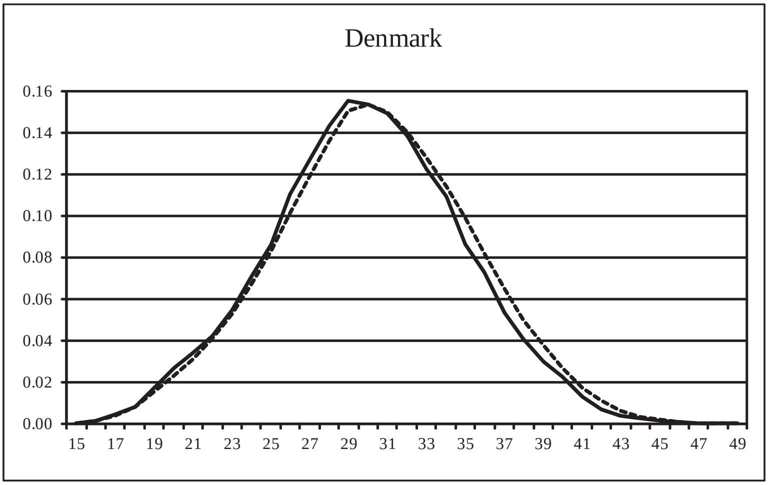
Figure 5.5 shows that the fertility rates at young ages in the United Kingdom have been declining since 1990. Therefore one plausible scenario seems to be that the high fertility rates at young ages in the United Kingdom will decline in the direction of the European average. Thus figure 5.8 shows a scenario for the United Kingdom where fertility rates at young ages will decline and older ages will increase. In the German-speaking countries in Central Europe fertility rates have been low for quite a long period. Lutz, Skirbekk, and Testa (2006) hypothesize that fertility may not rise from the current low levels due to adapted ideals of family size. This is the so-called 'low fertility trap' (Goldstein, Lutz and Testa, 2003). However, even though the total fertility rate has hardly changed during the last decade, fertility rates at ages 30 and above have been increasing in recent years. Therefore one plausible scenario may assume that for ages 29 and above there will be some movement towards the Nordic countries, but it does not seem likely that Germany will reach that level, since the gap between Northern and Central European countries has been considerable for quite some time now. Figure 5.8 shows such a scenario. In Italy fertility rates have been increasing for women in their 30s. The decrease of fertility rates among young women has stopped quite recently, so one may expect that the catching up of postponed births will continue for some time. Therefore an increase of fertility rates among women in their late 20s and early 30s may be expected. In Eastern European countries fertility levels have been low since the fall of Communism. Frejka (2008) mentions two explanations. The first is that the economic and social crises which occurred during the transition from the state socialist economies to market economies in the early 1990s were the principal causes of the decline in fertility. The second is that the diffusion of Western norms, values, and attitudes regarding family formation caused the changes in childbearing. Frejka argues that these explanations are not mutually exclusive. Figure 5.5 shows that in recent years there has been a slight increase in fertility, especially at older ages. One scenario could be to assume that the Eastern European countries will move towards the current European average, so that the rate ratios will become equal to 1. This would be in line with Frejka's assumption that young generations will adopt Western type norms, values, and attitudes regarding childbearing. That would result in a slight decline of fertility at young ages and an increase of fertility at ages 27 or over (see figure 5.8).

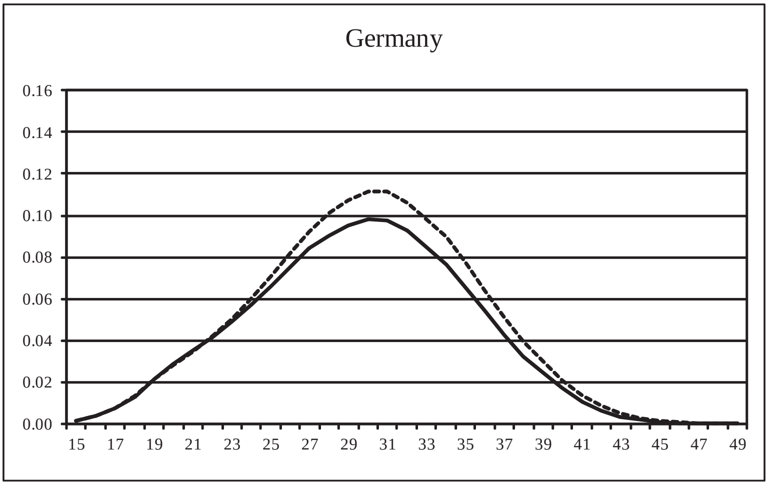
Multiplying the assumed values of the rate ratios shown in figure 5.8 by the EU27+3 average fertility age schedule produces scenarios of the age-specific fertility rates of the six countries under study. These are shown in *figure 5.9* Since the scenarios described in figure 5.8 show rather smooth age patterns of the rate ratios, the projected age-specific fertility rates shown in figure 5.9 are smoother than those shown in figure 5.7. One important difference between the scenarios shown in figure 5.9 and figure 5.7 is that the age patterns for Denmark and France are more peaked in figure 5.9. The explanation is that in figure 5.8 for these two countries we assume the same values of the rate ratios at age 29 for 2030 as in 2008, whereas the scenarios shown in figure 5.7 are based on the less peaked Swedish age pattern. The levels of the TFR implied by these scenarios are shown in table 5.4. They are close to the projections based on the assumption of convergence towards the Swedish fertility rates discussed above. The main difference is that the projected level of the TFR for Italy is higher according to the scenario shown in figure 5.9. The explanation is that the projections shown in figure 5.7 assume no increase in fertility rates for Italy up to age 29 (since the values of φ equal 1), whereas in figure 5.8 we assume some convergence towards the European average and thus an increase in rate ratios of women in their 20s. Whereas both scenarios are based on the assumption that there will be a converging tendency of fertility across European countries, the scenarios show that there will still be clear differences in the TFR of the high-fertility countries in Northern and Western Europe and the low-fertility countries in Central, Southern, and Eastern Europe. This is in line with the assumption by Frejka and Sobotka (2008) that this cross-country diversity in fertility is likely to prevail for decades to come.

5.6. Conclusion and discussion

The period TFR is determined not only by changes in the average number of children per woman among successive cohorts, but by changes in the timing of fertility as well. Since the effects of changes in the timing of fertility are temporary, we cannot simply extrapolate from recent changes in the TFR into the future. It is obvious that there are boundaries to changes in fertility rates. One solution is to adjust the level of fertility for changes in the tempo of fertility (see, *e.g.*, Frejka and Sobotka, 2008). However, there is some debate about the usefulness of such an adjustment (Van Imhoff, 2001; Schoen, 2004 and Goldstein, Sobotka and Jasilioniene, 2009). An alternative would be to make assumptions about the future values of the age-specific fertility rates rather than about the level of the TFR.

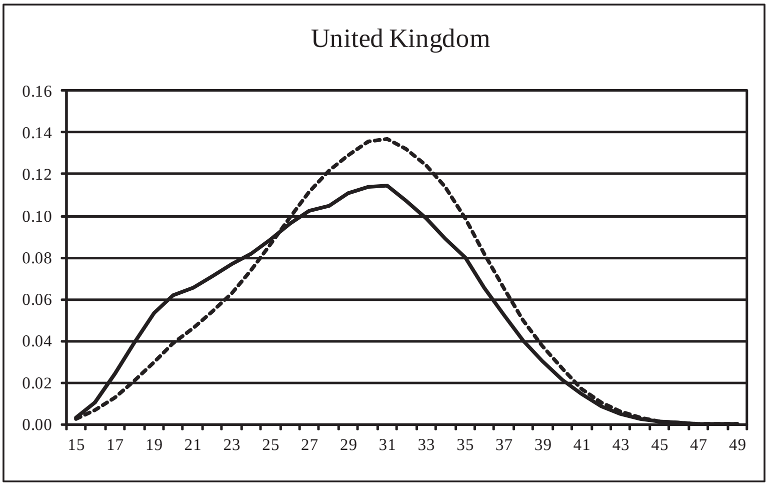
Figure 5.9. Age-specific fertility rates of six European countries, 2008, and scenario for 2030 based on assumptions about rate ratios compared with EU27+3 average

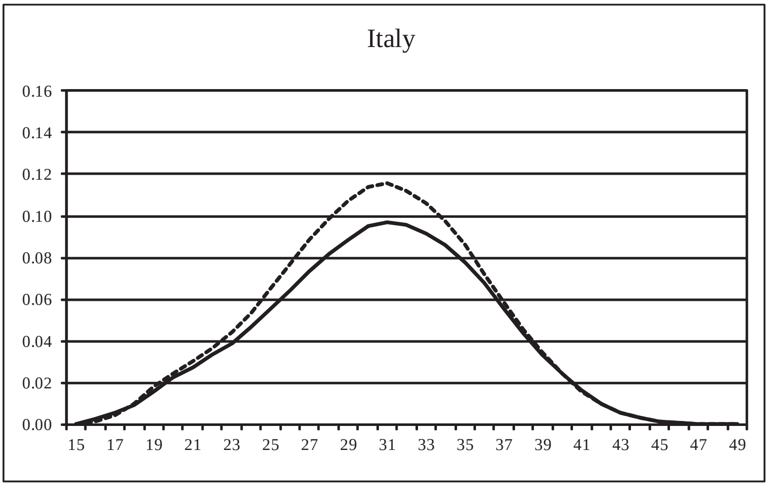




Solid line: observed values; dotted line: projections.

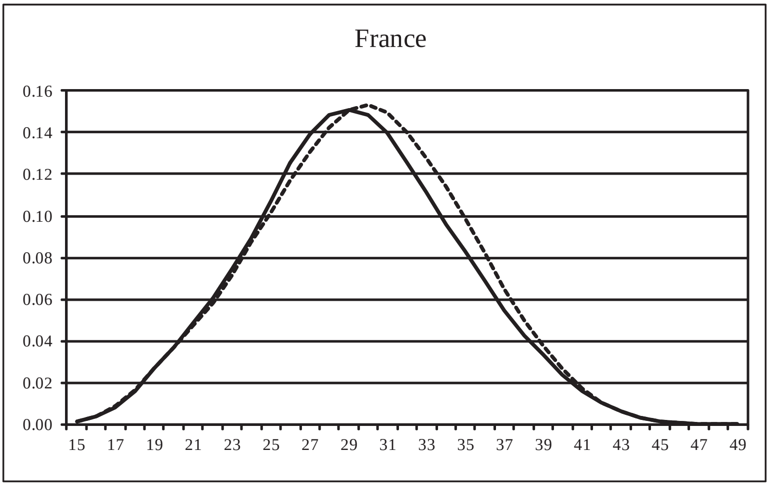
Figure 5.9. Age-specific fertility rates of six European countries, 2008, and scenario for 2030 based on assumptions about rate ratios compared with EU27+3 average (continued)

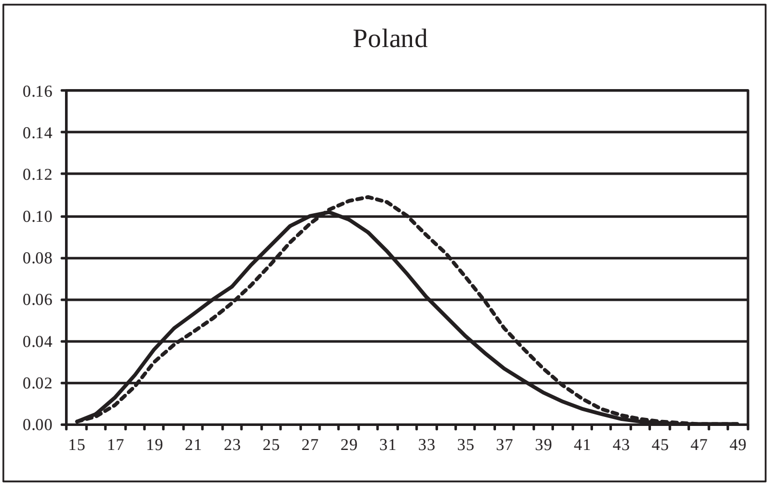




Solid line: observed values; dotted line: projections.

Figure 5.9. Age-specific fertility rates of six European countries, 2008, and scenario for 2030 based on assumptions about rate ratios compared with EU27+3 average (end)





Solid line: observed values; dotted line: projections.

For making projections of future age-specific rates it is useful to use standard age schedules rather than projecting each age-specific fertility rate separately. One approach is to make assumptions about the future values of the parameters of the standard age schedule. A disadvantage of this method is that the values of these parameters tend to be difficult to interpret. Usually they can be interpreted as indicating the direction of changes in the level, timing, and spread of fertility only, but the exact meaning of the value is not clear. Another disadvantage is that 'basic' model age schedules do not describe accurately age patterns for all age intervals in all countries at all periods. Consequently several extensions of these models have been proposed in the literature, making them more complicated. A possible way forward is to use splines. They are capable of describing all kinds of age patterns. Splines tend to provide a better fit than parametric models, since parametric models are smoother and thus do not capture deviations in specific age intervals in observed data. However, their usefulness for making projections or creating scenarios is limited, as they do not include interpretable parameters. Another approach is to use relational methods, as proposed by Brass in the 1970s. These models use a standard age schedule, but instead of making assumptions about the future values of the parameters of the age schedule, they make assumptions about the way in which the age pattern to be projected may differ from the standard age schedule. A big advantage of this approach is that the function describing the relationship between the age pattern to be fitted or projected and the standard age schedule is much simpler than the function describing the standard age pattern. One drawback of the Brass method, however, is that it includes two parameters that are difficult to interpret and is therefore not so useful in creating scenarios.

This chapter introduces a new relational method for fitting and projecting age patterns. TOPALS includes more parameters than the Brass model and is consequently more flexible and accurate. By using a linear spline function, TOPALS is flexible in two respects. First, it can describe all kinds of age curve. Second, the user can choose the desired level of goodness of fit and degree of smoothness. In general there is a trade-off: the smoother the age curve, the less accurate the fit. The flexibility of TOPALS implies that the method is less sensitive to the choice of standard age schedule than the Brass model. Therefore one standard curve may be appropriate for describing age-specific fertility rates in countries with different age patterns. However, this does not imply that the choice of the standard age schedule is arbitrary. The more strongly the age pattern of the rates to be fitted differs from the standard age schedule, the more knots are needed to provide an accurate fit. If many knots are needed one may consider choosing another standard age schedule. Visual inspection of the graph of the age pattern of the rate ratios shows immediately whether or not many knots will be needed. If the curve of the rate ratios is significantly non-linear in a particular age interval this implies that many knots will be needed for obtaining a good fit for that age interval. In selecting an appropriate standard age schedule, the main criterion is not the fit with the data but the interpretation of the values of the rate ratios. If the standard age schedule does not have a clear geographical or other relationship with the age-specific rates to be fitted or projected, the values of the rate ratios lack a clear interpretation and thus their usefulness in creating scenarios is limited.

TOPALS can be used for making scenarios in two ways, one of which is more objective, while the other is more subjective. First, one may specify a 'target' age schedule of fertility, describing the values of the age-specific fertility rates that are expected to be reached in the long run. This may be the current fertility age schedule of a forerunner country. Second, one may make an assumption about the fertility age schedule of a young cohort. In the latter case the period age-specific fertility rates are assumed to move in the direction of a cohort age schedule. The time series of the rate ratios of the age-specific fertility rates to be projected and the target age schedule shows whether the fertility rates are moving towards the target. A partial adjustment model can be estimated to determine how quickly the fertility rates will move in the direction of the target values. One benefit of using a partial adjustment model is that the forecaster does not need to specify a priori in which year the target value will be reached. In addition, if the average fertility rates over a number of countries are used as the standard age schedule the rate ratios will show how the current age pattern and level of fertility of each country differ from the average pattern. The forecaster can specify qualitative assumptions about the way in which future differences in the age pattern of fertility may differ from the current pattern. If one assumes convergence across countries, it follows that the future values of the rate ratios will move towards a value around 1, whereas if one assumes that certain country-specific characteristics of fertility are persistent (*e.g.*, relatively low or high fertility at the youngest or oldest ages) it follows that rate ratios for specific age intervals will remain constant.

The use of TOPALS for making assumptions about fertility implies that one is making assumptions about the levels of age-specific fertility rates and that the value of the TFR is an outcome, whereas in many countries it is common practice in making projections to first make an assumption about the future values of the TFR and then calculate age-specific rates in line with this assumption. The illustrations in this chapter show that when using TOPALS to create scenarios one may assume that the changes in fertility rates differ across ages. Rather than assuming only that the total level of fertility will increase or that fertility will be postponed, it is possible to make separate assumptions for different ages. Thus TOPALS makes it possible to create scenarios in which the shape of the age schedule changes. This allows the forecaster to make a distinction between a rise in the mean age at childbearing due to a decrease in fertility rates at very young ages and a rise at older ages caused by the catching up of postponed births. The use of TOPALS makes it possible to estimate whether the fertility age pattern will be more peaked in the future or not. For example, the Netherlands has a relatively high mean age at childbearing but not very high fertility rates at the oldest ages. The reason is that the fertility age pattern in the Netherlands is more peaked than the European average. The Dutch case shows that postponement of fertility will not necessarily result in a postponement of fertility to the oldest ages, which would imply an increase in the number of couples with problems of infecundity. Thus one alternative to the scenarios discussed in the previous section would be to use a combination of Swedish and Dutch fertility rates as standard age schedule: namely, the Dutch age-specific fertility rates multiplied by the ratio of the Swedish and Dutch TFRs. Then one could create a scenario which assumes that other countries will move towards the peaked age pattern of the Netherlands and the relatively high level of Sweden.

TOPALS can be used in combination with other smoothing methods. For example, one may use a cubic spline for producing a smooth age schedule of fertility for a given country in a given year and use TOPALS to make assumptions about future changes in this age schedule. In this case one is trying to explain not the form of the smooth age schedule but rather future changes compared with this standard age schedule. Alternatively TOPALS can be used in combination with a simple parametric model to describe deviations in the age pattern of fertility for a particular country in comparison with this simple model. Applied in this way, TOPALS may increase the usefulness of simple parametric models without making them more complicated. For example, the relatively high fertility rates at young ages in the United Kingdom and Ireland can be described by using the Hadwiger function as standard age schedule, with TOPALS to describe the high fertility rates at young ages.

This chapter shows how TOPALS can be applied for smoothing and projecting age-specific fertility rates for EU countries. TOPALS may be used to fit and project age-specific fertility rates in other parts of the world as well: for example, using data from the Human Fertility Database (2010), developed by the Max Planck Institute for Demographic Research and the Vienna Institute of Demography. For countries with missing or less reliable or detailed data, TOPALS can be used to estimate single-year age-specific fertility rates: for example, when only data for five-year age groups are available. Brass's relational model is widely used for this purpose. Using TOPALS, one may choose the age-specific fertility rates of a neighbouring country or a parametric model as the standard age schedule. Furthermore TOPALS may be used for regional population projections. It is common practice to make assumptions about regional differences in fertility rates compared with the national average. Thus one may calculate rate ratios of age-specific fertility rates at regional and at national levels and make assumptions about future changes in these rate ratios.

One of the implications of the flexibility of TOPALS is that it can also be used for describing age-specific rates other than fertility. For example, the next chapter shows that TOPALS can be used for fitting and projecting age-specific mortality rates. One final possible application of TOPALS is to create scenarios of age-specific rates taking into account the effects of covariates. For example, when the effect of the level of educational attainment varies between age groups, TOPALS can be used to create scenarios of the rates for different education categories (De Beer, Van der Gaag and Willekens, 2007).

6. Smoothing and projecting age-specific probabilities of death by TOPALS

Abstract

TOPALS is a relational model that can be used for smoothing and projecting age-specific rates or probabilities. This chapter shows how TOPALS can be used for projecting age-specific death probabilities for 26 European countries. Projected death probabilities for Japanese women in the year 2060 are used as standard age schedule. A partial adjustment model is used to assess to what extent death probabilities of European countries will move in the direction of the Japanese level. Three alternative scenarios are calculated.

6.1. Introduction

This chapter shows how the new relational method TOPALS (Tool for projecting age-specific rates using linear splines) can be used for smoothing and projecting age-specific probabilities of death. TOPALS is both easy and flexible. The calculations are easy and the values of the parameters can be interpreted easily. TOPALS can be used to describe different age schedules. Chapter 5 demonstrates how TOPALS can be used to smooth and project age-specific fertility rates. This chapter demonstrates how TOPALS can be used to smooth and project age-specific probabilities of death for 26 European countries on the basis of data from the Human Mortality Database (2010). TOPALS can be applied to make different types of scenarios. It can be used to calculate baseline scenarios which can be regarded as a projection of past trends. But it can be used to make alternative scenarios as well.

Projections of life expectancy at birth are usually based on extrapolations of time series, either of life expectancy itself or of the underlying age-specific probabilities of death or of parameters of a model describing the age pattern of mortality (Bongaarts, 2006 and Tabeau *et al.*, 2001). Even though there is general agreement that life expectancy will continue to grow, there is less agreement on the extent of the increase (Bongaarts, 2006; Garssen, 2006 and De Beer, 2006). Oeppen and Vaupel (2002) note that the ‘best practice’ life expectancy has been increasing linearly by 2.5 years per decade during the past 150 years. They expect that this trend will continue in the coming decades. Bongaarts (2006) agrees that life expectancy at birth will continue to increase but assumes that the average increase will be 1.5 years per decade.

Bongaarts argues that the strong decline in death rates at young ages observed in the last 50 years cannot continue since the death probabilities have already reached very low levels. Olshansky and Carnes (1994) argue that a linear projection of life expectancy is very optimistic as this can only be achieved if the decline in age-specific probabilities of death will accelerate. Olshansky *et al*. (2005) and Stewart, Cutler and Rosen (2009) argue that an increase in obesity may reduce the increase in life expectancy. In the past decades there have been different underlying changes in age-specific death probabilities. Whereas improvements in life expectancy in the first half of the 20st century were mainly caused by a strong decline in infant mortality, improvement in recent decades is mainly due to a decline of mortality at older ages. Even though there has been a linear increase in best practice life expectancy, underlying declines in age-specific death probabilities have not been linear. Thus it is not obvious that life expectancy will continue to increase linearly in the future. Moreover life expectancy in individual countries has not shown a linear increase over very long periods (Lee, 2006).

This chapter describes how TOPALS projects age-specific death probabilities by modelling a movement towards levels of death probabilities that are consistent with a linear extrapolation of best practice life expectancy. TOPALS uses a linear spline to model the ratios between the age-specific rates or probabilities to be fitted or projected and a smooth standard age schedule. For smoothing age-specific death probabilities a smooth standard age schedule is used. This can be a model age schedule or the average age pattern of mortality across a number of countries. This chapter uses the average age-specific death probabilities of 15 Northern, Western and Southern European countries as standard. Using TOPALS for making projections of age-specific probabilities, one needs to make projections for selected ages (the so-called knots) only. This article uses TOPALS for making three scenarios. Each of the scenarios uses the same standard age schedule. In line with Oeppen and Vaupel's suggestion that the best practice life expectancy of Japanese women may continue to increase linearly, we make a projection of age-specific death probabilities of Japanese women for 2060 assuming a linear increase in life expectancy. The scenarios differ by the assumptions to what extent the probabilities of death of European countries are assumed to move into the direction of the best practice level. This can be described by a partial adjustment model which includes one parameter.

The second section of this chapter gives a brief discussion of methods that can be used for smoothing age-specific death probabilities, and the third section discusses methods for projecting life expectancy. The fourth section describes TOPALS. The fifth section shows how TOPALS can be used for smoothing age-specific probabilities of death. Section 6.6 describes the use of TOPALS for making three alternative scenarios for 26 European countries. The final section summarizes and discusses the results.

6.2. Methods for smoothing age-specific death probabilities

The Gompertz model of mortality implies that mortality rates increase exponentially with age. The mortality rate equals the number of deaths at age *x* divided by the number of person-years at risk at age *x*. Mortality rates can be estimated from population statistics. The death probability is the probability that a person who has reached age *x* will die before reaching age *x+1*. Age-specific death probabilities can be derived from mortality rates. For example, assuming a uniform distribution of exposure in *x*, $q(x) = m(x)/[1+½m(x)]$ where *q(x)* is the death probability at age *x* and *m(x)* is the mortality rate. For one-year intervals the values of mortality rates and death probabilities are close, but death probabilities are always smaller than mortality rates. For individual countries age-specific death probabilities show a rather irregular pattern. Therefore for analysing changes over time and making projections it is useful to smooth the age pattern. Different methods may be used for this purpose. Parameterisation of mortality by age has a long history (Booth, 2006). Nowadays the most widely used method is the Heligman-Pollard model (Heligman and Pollard, 1980):

$$\frac{q(x)}{1-q(x)} = A^{(x+B)^C} + D\exp(-E(\log(\frac{x}{F}))^2) + GH^x \tag{1}$$

where the left-hand side represents the odds that an individual aged *x* will die before age *x+1*. This model includes eight parameters to be estimated. All these parameters have a positive value. The first term of the right hand side reflects infant and early childhood mortality, the second term the accident hump and the third term the exponential increase in mortality at later ages. One problem in using this model for projection purposes is that even though the three terms have an interpretation, the individual eight parameters lack a direct demographic interpretation. Another problem is that the parameter values are interdependent (McNown *et al.*, 1995). Booth (2006) concludes that the Heligman-Pollard model is not very useful for forecasting. Even the eight parameters in the Heligman-Pollard model seem not to be sufficient to describe age-patterns of mortality for all countries. For that reason Kostaki

(1992) presented a nine-parameter version. Carriere (1992) proposed a four-term model with eleven parameters.

Instead of making the model more complicated by adding parameters one alternative procedure is to estimate a relational model. This implies that one uses a smooth age pattern and specifies a simple model that describes how the age-specific rates to be smoothed differ from the standard age schedule. Brass (1971) presents a relational method in which age-patterns of mortality are related to each other by a linear relationship between the logits of the survivorship probabilities:

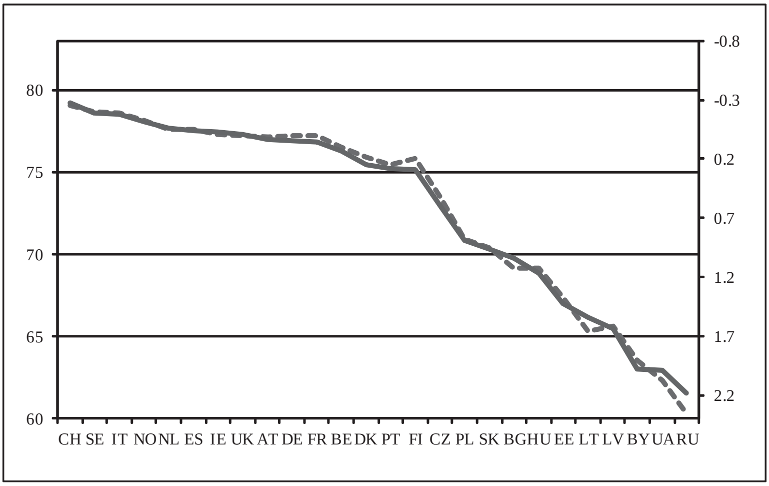
$$\ln \frac{1-l(x)}{l(x)} = \alpha + \beta \ln \frac{1-l^*(x)}{l^*(x)} \tag{2}$$

where *l(x)* is the life table survival probability and *l*(x)* is the standard survival probability. The parameters α and β can be estimated by OLS regression. The parameter α is related to the life expectancy at birth (Brass, 1971; 1974). *Figure 6.1* shows estimated values of α and life expectancy at birth for 26 European countries. The values of α are estimated using the average death probabilities of 15 Northern, Western and Southern European countries as standard (see section 6.5 for details). The figure shows a strong inverse relationship between α and life expectancy at birth. Life expectancy at birth explains 99.5 percent of the variance in α for men and 97.4 percent for women. However in contrast with life expectancy the value of α has no direct demographic interpretation. The relationship between the value of α and life expectancy differs between men and women. For example, for men life expectancy of 73 years corresponds with a value of α of 0.6, and for women with a value of α of 1.3. The value of β reflects the steepness with which the survival probabilities change with age (Brass, 1971).

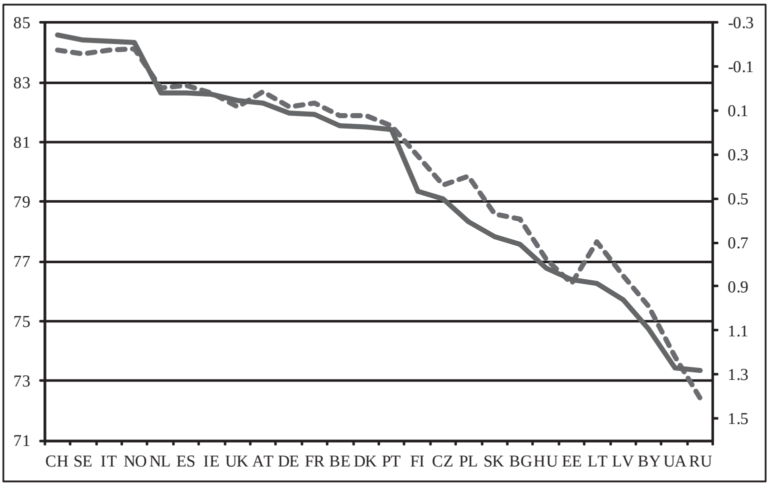
If $\beta = 1$, the steepness of the age schedule of death probabilities is the same as that of the standard schedule. If $\beta > 1$ the steepness of the age curve to be fitted is larger than that of the standard schedule. For the 26 European countries the estimated value of β ranges from .93 to 1.22. If $\alpha = 0$ and $\beta > 1$, the age-specific death probabilities for ages up to 70 years for men and up to 77 years for women are lower than those according to the average age schedule and for older ages higher. Thus for making projections, if one assumes that life expectancy at birth will increase and that this will be mainly caused by a decline of mortality at older ages, one should assume that the value of α will increase and that of β will decrease. However, it would be

Figure 6.1. Estimated values of α and life expectancy at birth for 26 European countries.

Males



Females



Solid line: α (right axis, reverse order); dotted line: life expectancy at birth (left axis).

difficult to determine to which values α and β will change as these values do not have a direct demographic interpretation.

6.3. Methods for projecting life expectancy

Life expectancy at birth can be projected on the basis of a time series of life expectancy or on the basis of a time series of age-specific death probabilities (Bongaarts, 2006). Since 1981 Japanese women have had the highest life expectancy at birth. Oeppen and Vaupel (2002) label this as the 'best practice' life expectancy. *Figure 6.2* shows that since the early 1980s the development of life expectancy at birth of Japanese women is close to linear. Thus one may project life expectancy at birth of Japanese women by a random walk model with drift:

$$e_{0,t} = e_{0,t-1} + c + u_t \qquad (3)$$

where $e_{0,t}$ = life expectancy at birth in year t, c is a constant term ('drift') and u_t is a random term, with $E(u_t) = 0$. In 2008 life expectancy of Japanese women equaled 86 years. For the period 1978-2008, the estimate of c equals 0.26. This implies that life expectancy at birth of Japanese women has increased by one year in each four years period. This corresponds with Oeppen and Vaupel's estimate. Using equation (3) to project life expectancy of birth of Japanese women leads to a projected value of 99.6 years in 2060 and 110 years in 2100. Changes in the level of life expectancy at birth are caused by changes in the underlying age-specific probabilities of death. If the logarithms of age-specific probabilities of death decrease in a linear way, life expectancy will increase less than linearly.

The Lee-Carter method has become the most widely applied model for making projections of age-specific probabilities of death (Booth, 2006). Lee and Carter (1992) decompose the level of mortality rates into age-dependent and time-dependent components. Since the time series of age-specific mortality rates and death probabilities show similar developments over time, the Lee-Carter model can be used to project changes in death probabilities as well:

$$\ln q(x)_t = a(x) + b(x)k_t + e(x)_t \qquad (4)$$

where $q(x)_t$ is the probability of death at age x in year t, $a(x)$ describes the average age pattern, k_t describes the change in probabilities of death over time, $b(x)$ determines how the change varies by age, and $e(x)_t$ is a random term with $E(e(x)_t) = 0$. Lee and Carter (1992) assume that $\sum_x b(x) = 1$ and $\sum_t k_t = 0$.

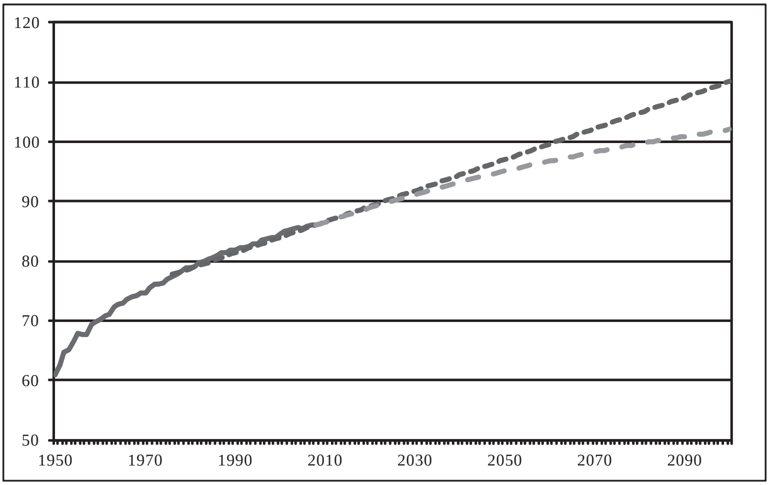
These normalizations make it possible to obtain unique least squares estimates of the values of $a(x)$, $b(x)$ and k_t. For this purpose Singular Value Decomposition (SDV) is applied, but linear regression produces similar results (Lee and Carter, 1992). Since $a(x)$ and $b(x)$ are time invariant, future values of $q(x)_t$ can be projected by projecting k_t. The Box-Jenkins method is used to identify an ARIMA model for projecting k_t. In almost all applications k_t is projected by a random walk with drift model (Booth, 2006):

$$k_t = k_{t-1} + d + u_t \tag{5}$$

where d = the drift parameter and u is a random term with $E(u_t) = 0$ and $E(u_t u_{t+j}) = 0$ for $j \neq 0$. Thus k_{t+T} can be projected by:

$$\hat{k}_{t+T|t} = k_t + T\hat{d}. \tag{6}$$

Figure 6.2. Life expectancy at birth of Japanese women, 1950-2100



Solid line: observed values, 1950-2008. Dotted line: Random walk with drift (fitted values, 1978-2008; projected values, 2009-2100). Dashed line: Lee-Carter model (projected values 2009-2100).

where $\hat{k}_{t+T|t}$ is the projection of k_{t+T} based on observations up to and including t and $\hat{d}$ is the estimate of the drift. It can easily be shown that this implies that for each age x the logarithm of the probability of death can be projected by a random walk with drift model.

From (4) it can be derived that:

$$\ln q(x)_t - \ln q(x)_{t-1} = b(x)(k_t - k_{t-1}) + e(x)_t - e(x)_{t-1} \tag{7}$$

This can be rewritten as:

$$\ln q(x)_t = \ln q(x)_{t-1} + b(x)d + v(x)_t - e(x)_{t-1} \tag{8}$$

where $v(x)_t = b(x)u_t + e(x)_t$. Since $E[v(x)_t] = 0$, the probability of death at age x can be projected by:

$$\ln \hat{q}(x)_{t+T|t} = \ln q(x)_t + Tb(x)\hat{d} - e(x)_t. \tag{9}$$

This implies that the projected change in the logarithm of the probability of death is linear. Applying the Lee-Carter model to the time series of probabilities of death of Japanese women for the period 1978-2008 leads to a projection of life expectancy at birth in 2060 of 97.1 years and a value of 102.0 years in 2100. Figure 6.2 shows that in the long run the projections of the Lee-Carter model are lower than those of linear projections of the time series of life expectancy. This demonstrates that a linear projection of logged death probabilities leads to a lower projection of life expectancy than a linear projection of life expectancy itself.

The Lee-Carter model does not produce a smooth age pattern, since the projected changes in the death probabilities differ across ages. For that reason Renshaw and Haberman (2003) suggest to smooth the projected age-specific death probabilities by a cubic spline. One alternative method is to project the age pattern of death probabilities rather than individual age-specific death probabilities. One method is to project the parameters of a model age schedule, such as the Heligman-Pollard model. However, as we noted in section 6.2 this raises two problems. The values of the individual eight parameters have no direct demographic interpretation and the parameters cannot be projected independently from each other. Another method is to apply a relational model. In section 6.2 we discussed the Brass relational model. By making assumptions about the future values of α and β one can use this model for making projections of age-specific death probabilities. Brass (1974) suggests to project α and β on the basis of past trends. One problem, however, is that if death probabilities across time are related to the same standard age schedule the fit of the model may vary across time. Thus one may question to what extent changes in α and β over time accurately describe changes in the age pattern of mortality. The next section describes the new relational model TOPALS that is less sensitive to the choice of the standard age schedule and thus is better capable of describing changes over time which makes it more suitable for projecting age-specific death probabilities.

6.4. TOPALS

We assume that a standard age schedule of probabilities of death is given. The age profile for a given country can be estimated on the basis of ratios of the age-specific probabilities of death of that country and those according to the standard age schedule. The risk ratio at age x is equal to:

$$r(x) = \frac{q(x)}{q^{*}(x)} \tag{10}$$

where $q^*(x)$ is the probability of death at age x according to the standard age schedule. The age pattern of the risk ratios can be described by a linear spline function. This is a piecewise linear curve. The ages at which the successive linear segments are connected are called ‘knots’. The risk ratios at each age can be estimated by the linear spline function:

$$\hat{r}(x) = a + \sum_{j=1}^{n} b_j (x - k_j) D_j \tag{11}$$

where $D_j = 0$ if $x \le k_j$, and $D_j = 1$ otherwise, k_j are the knots, a and b_j are the parameters to be estimated. The knots can be chosen in such a way that the fit of the linear spline to the data is optimal, *e.g.* by applying a non-linear least squares method. However, this would result in different knots for different countries. Since we want to make cross-country comparisons we decided to fix the location of the knots a priori at the same ages for each country. We use data from the Human Mortality Database. They refer to ages 0 up to and including 109. We decided to fix the knots at ages 20, 30, 40, …, 100, 109. Since age-specific probabilities of death for ages 0-20 show an irregular pattern, we assume the risk ratios for these ages to be equal to the average of the risk ratios for this age group, *i.e.* the slope of the spline is assumed to equal zero for ages 0-20. The values of a and b_j can be estimated by OLS. A simpler procedure is to assume that the values of the spline at the knots equal the observed values. It turns out that this provides a fit that is very close to the one produced by applying OLS. Thus we assume that

$$\hat{r}(k_1)=\sum_{x=0}^{20}\frac{r(x)}{21},\ \hat{r}(k_2)=r(k_2),\ \hat{r}(k_3)=r(k_3),...,\hat{r}(k_{n+1})=r(k_{n+1}).$$

Then the values of a, b_j can be estimated by substituting the values of $\hat{r}(k_1)$, $\hat{r}(k_2)$, etc. in (11). This yields:

$$\hat{a}=r(k_1);\ \hat{b}_1=\frac{r(k_2)-r(k_1)}{k_2-k_1};\hat{b}_j=\frac{r(k_{j+1})-r(k_j)}{k_{j+1}-k_j}-\sum_{i=1}^{j}\hat{b}_{i-1} \qquad (12)$$

The age-specific probabilities of death are estimated by multiplying the ratios which are estimated by the linear spline function $\hat{r}(x)$ by the age-specific death probabilities according to the model age schedule $q^*(x)$:

$$q(x)=\hat{r}(x)q^*(x). \qquad (13)$$

For smoothing age-specific probabilities of death the standard age curve can be the average of several countries, *e.g.* the EU average, the age curve of another country or a model age schedule. For projections the standard age schedule can be the age schedule of a ‘forerunner’ country or some age pattern that may be expected to be reached in the long run. Oeppen and Vaupel (2002) and Bongaarts (2006) argue that there is no evidence of approaching limits to longevity. Therefore we do not assume that a certain limit will be reached in a given year. Instead we assume that death probabilities will move towards the ‘best practice’ level in the long run. We estimate the speed with which the probabilities of death move into the direction of the target values using a partial adjustment model. This model assumes that the speed of the movement towards the target level will decline when the target level will be approached. This is in line with Oeppen and Vaupel’s finding that “rapid progress in catch-up periods typically is followed by a slower rise” (Oeppen and Vaupel, 2002). Lee (2006) finds that countries tend to converge toward the life expectancy leader and that they converge more than proportionally with the size of the gap between their life expectancy and record life expectancy. We model the time series of risk ratios as a partial adjustment model assuming that the risk ratios move towards 1:

$$r(x)_t-1=\varphi(x)[r(x)_{t-1}-1]+e(x)_t \qquad (14)$$

where $r(x)_t$ is the risk ratio in year t, $0 \leq \varphi(x) \leq 1$ and $e(x)_t$ is a random term with $E[e(x)_t] = 0$. This model assumes that the value of $r(x)_t$ is closer to 1 than the value of $r(x)_{t-1}$. The lower the value of $\varphi(x)$, the quicker $r(x)_t$ will move towards 1. If $\varphi(x)$ is close to 1, $r(x)_t$ moves slowly to 1. If $\varphi(x)= 1$ model (14) describes a random walk, and $r(x)_t$ does not converge to 1. The reason for assuming that $\varphi(x) \leq 1$ is that if $\varphi(x) > 1$ the risk ratios would move away from 1. If the probabilities of death are higher than those according to the standard schedule, (14) implies that the death probabilities are projected to decrease, whereas if the death probabilities are smaller than those according to the standard schedule, the model projects an increase.

Since $E[e(x)_t] = 0$ projections of model (14) can be calculated by:

$$\hat{r}(x)_{t+k|t} = \varphi(x)\ \hat{r}(x)_{t+k-1|t} + 1 - \varphi(x) \tag{15}$$

where $\hat{r}(x)_{t+k|t}$ is the projection of $r(x)_{t+k}$ based on observations up to year t. Since $0 \leq \varphi(x) \leq 1$, the projections equal:

$$\hat{r}(x)_{t+T|t} = \varphi(x)^T r(x)_t + 1 - \varphi(x)^T \tag{16}$$

Thus if $\varphi(x) < 1$ the projections will move to 1 for large T.
The future values of the age-specific probabilities of death can be projected by:

$$\hat{q}(x)_{t+T|t} = \hat{r}(x)_{t+T|t} q^*(x)\,. \tag{17}$$

Obviously the values of $r(x)_t$ depend on the choice of $q^*(x)$. If the target values $q^*(x)$ are very low, the values of $r(x)_t$ will be high. The estimate of the value of $\varphi(x)$ depends on the level of $r(x)_t$. If the values of $r(x)$ are high, *i.e.* the distance to the target value is large, the value of $\varphi(x)$ will be high and it will take more time to reach the target value. Thus if one assumes extremely low probabilities of death as target values, they will be reached in the very long run only. Different scenarios can be specified by different values of the target pattern. If very low target values would be assumed, the probabilities of death may decline to a much lower level. However, these low levels would only be reached in the very distant future. For the next 50 years or so that would not lead to quite different scenarios. For specifying alternative scenarios for the next 50 or 100 years the values of $\varphi(x)$ are more important. Assuming a low level of $\varphi(x)$ implies that the low target levels of age-specific probabilities of death will be reached quicker.

Instead of applying the partial adjustment model (14) the risk ratios can be projected by using a random walk with drift model for the logarithms of the risk ratios:

$$\ln r(x)_t = \ln r(x)_{t-1} + c(x) + e(x)_t \tag{18}$$

Since:

$$\ln r(x)_t = \ln q(x)_t - \ln q^*(x)_t \tag{19}$$

we can derive the model for projecting probabilities of death by substituting (19) into (18):

$$\ln q(x)_t = \ln q(x)_{t-1} + c(x) + e(x)_t \tag{20}$$

A comparison with the projections produced by the Lee-Carter model (8) shows that the projections are similar as both project a linear change of the logarithms of the probabilities of death. However, the projections based on (20) do not equal those based on (8) for two reasons. First, using TOPALS the random walk model is used for projecting probabilities of death at the knots only and the projections for ages in between are obtained by the linear spline of risk ratios (11). Secondly, the estimate of the drift *c(x)* in equation (20) does not equal the estimate of the drift *b(x)d* in equation (8). The drift in equation (20) is estimated for each knot separately by:

$$\hat{c}(x)_t = (\ln q(x)_t - \ln q(x)_{t-L}) / L \tag{21}$$

In equation (8) the drift includes an age-specific component *b(x)* that does not change over time and a component *d* that is estimated from the time series k_t:

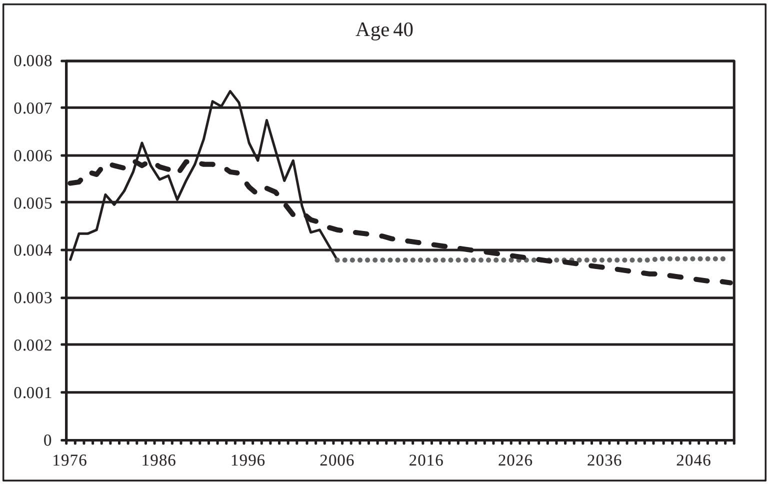
$$\hat{d} = (\hat{k}_t - \hat{k}_{t-L}) / L \tag{22}$$

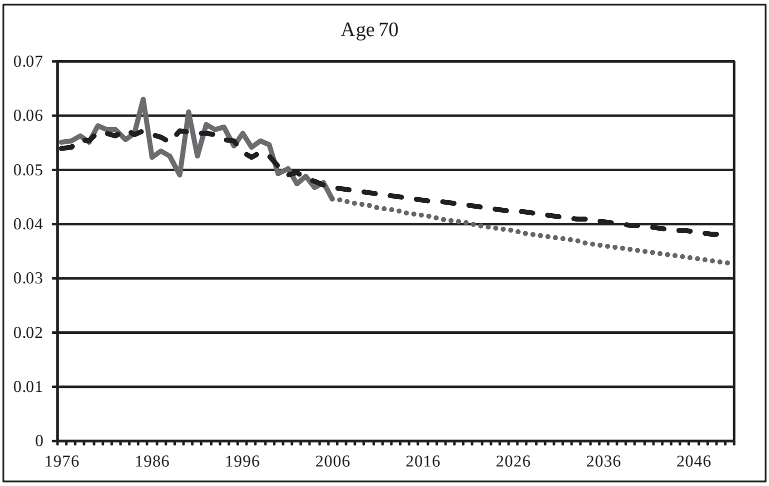
The estimates of k_t and k_{t-L} are based on a summation across all ages as can be seen from rewriting (4). Since $\sum_x b(x) = 1$ and $\sum_x e(x)_t = 0$ k_t can be derived from (4) as follows:

$$k_t = \sum_x \ln q(x)_t - \sum_x a(x) \tag{23}$$

As a consequence the time series of k_t is more stable than the time series of $q(x)_t$ for each age separately. *Figure 6.3* illustrates the difference between

Figure 6.3. Projections of death probabilities of Hungarian men, ages 40 and 70





Solid line: observed values; dashed line: Lee-Carter model; dotted line: Random walk model.

projections based on equations (21) and (22). The figure shows projections of the death probability of Hungarian men at ages 40 and 70. The observation period is 1976-2006. The dotted lines show the projections that are based on a random walk with drift model where the drift is estimated by equation (21). The dashed lines show the Lee-Carter projections based on equation (22). The latter projections extrapolate the fitted time series rather than the observed time series. Note that for both ages the Lee-Carter model describes a similar development across time apart from differences in the levels of the death probabilities between both ages. Since for age 70 the average decline in the fitted time series in the period 1976-2006 is smaller than the average decline between the last and first point of the observed time series, the Lee-Carter model projects a smaller decrease than the random walk model (20). For age 40 the opposite is true.

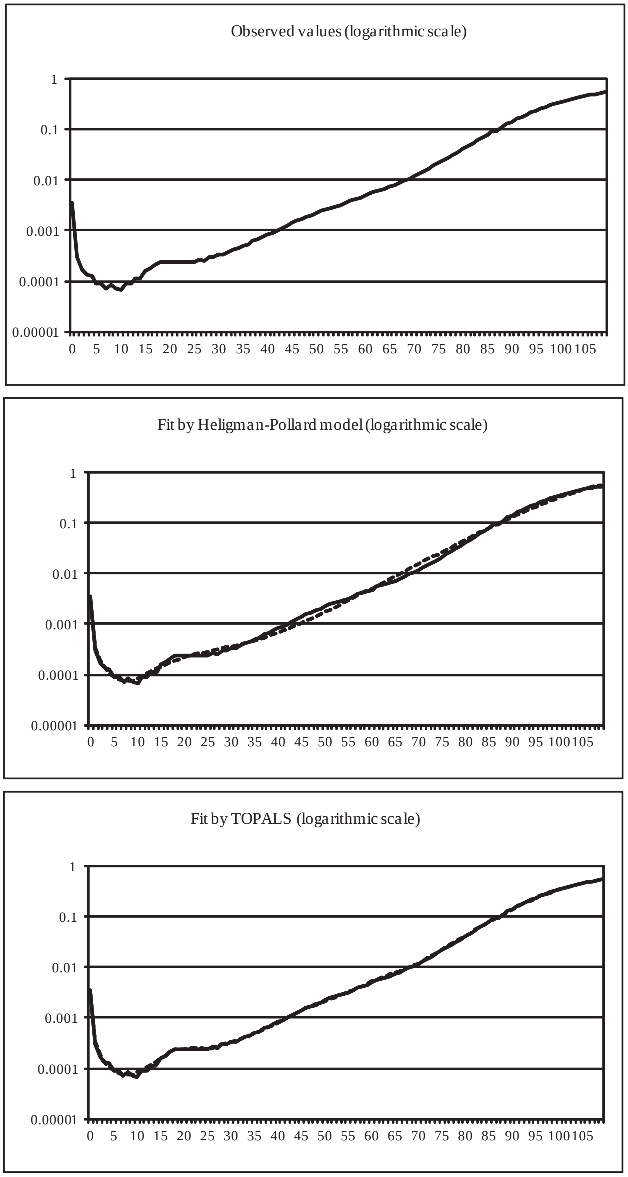
6.5. Smoothing age-specific probabilities of death

Age-specific death probabilities are obtained from the Human Mortality Database (2010). This database includes life tables for 29 European countries. These countries include 23 of the 27 EU countries: Austria, Belgium, Bulgaria, Czech Republic, Denmark, Estonia, Finland, France, Germany, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Netherlands, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden and United Kingdom. Six non-EU countries are included: Belarus, Iceland, Norway, Russia, Switzerland, Ukraine. For our analyses we used data for 26 countries. We did not include Iceland and Luxembourg because of their small population size, and we did not include Slovenia because the time series is shorter than for the other countries. The most recent year for which the database includes data for all 26 countries is 2006. For the sake of cross-country comparability we used this year as jump-off year for the projections for all countries. For our analyses we use the probabilities of death from these life tables. The probabilities of death included in the Human Mortality Database are smoothed at the highest ages using a logistic model (Wilmoth *et al.*, 2007). It is assumed that across all countries the death probability at age 110 equals 1. As a consequence the age patterns at the oldest ages look similar across countries. Even though there is discussion whether mortality rates at the oldest ages increase with age as described by a logistic or a Gompertz model (Boleslawski and Tabeau, 2001 and Booth, 2006), we decided to use the estimates included in the Human Mortality Database as these data are comparable across countries.

For applying TOPALS we need to specify a smooth standard age schedule. For this purpose we calculated the weighted average of the age-specific death probabilities for men and women of 15 Northern, Western and Southern European countries included in the Human Mortality Database: Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom. We weighted the death probabilities by total population size for men and women separately. We label this as the NWS European average. *Figure 6.4* shows the logarithms of the average age-specific death probabilities for females. The upper panel of figure 6.4 shows that the average probabilities are not smooth at ages below 20. At older age there are some irregular fluctuations as well. In order to obtain a smooth pattern for the whole age schedule we estimated the Heligman-Pollard model. The middle panel of figure 6.4 shows that this does not produce a perfect fit for all ages. Around age 50 the fitted values are too low and around age 70 too high. For that reason we applied TOPALS using the Heligman-Pollard curve as standard age schedule. For women the risk ratio at age 50 equals 1.3 and for age 70 it equals 0.8. For men the risk ratios are much closer to 1 as the fit of the Heligman-Pollard function is better. Multiplying the age-specific death probabilities according to the Heligman-Pollard model by the fitted linear spline (not shown here) produces the smooth curve shown in the lower panel of figure 6.4, which turns out to provide a very accurate fit. This curve is used as standard age schedule for smoothing the age-specific death probabilities for the 26 European countries in this study.

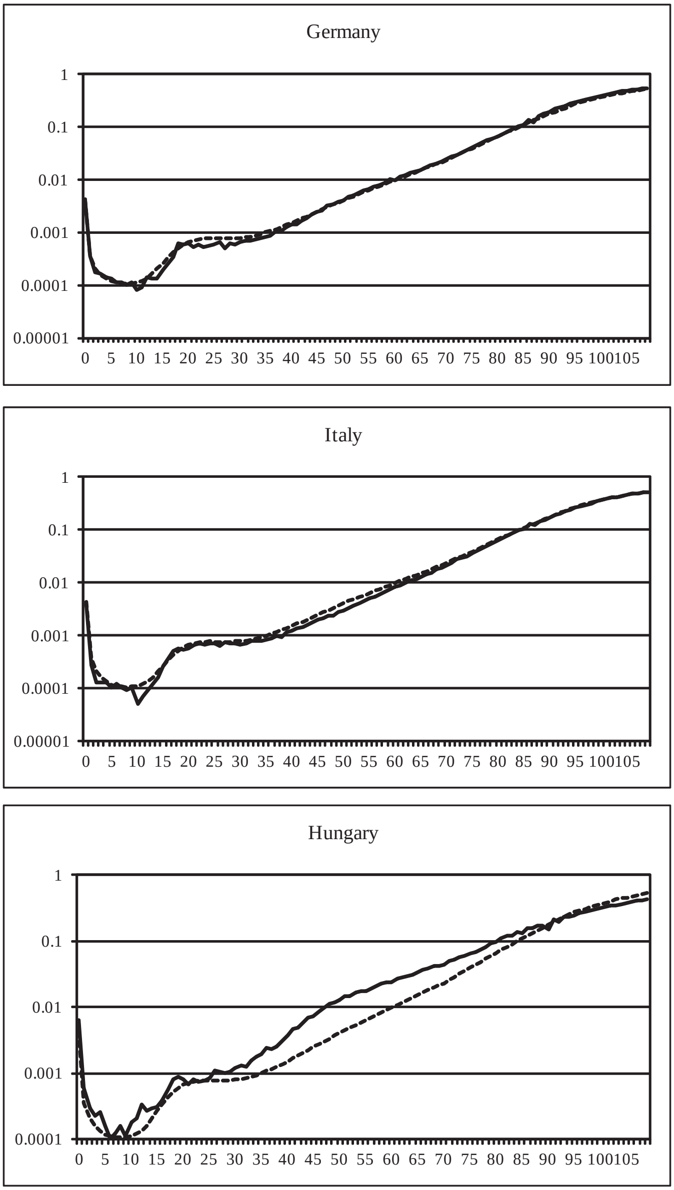
We illustrate the use of TOPALS by applying the method to three countries which are representative for the variation in mortality patterns in Europe. Germany has death probabilities that are close to the European average, Italy has lower death probabilities and Hungary has high probabilities. *Figures 6.5a and 6.5b* compare age-specific probabilities for men and women for these countries with the NWS European average. Life expectancy at birth for Germany equals 77.2 years for men and 82.3 for women. The NWS European average equals 77.4 years for men and 82.8 years for women. Italy has lower death probabilities for almost all ages. Life expectancy at birth for Italian men equals 78.6 years, thus 1.2 years above the average and for women the Italian life expectancy of 84.1 years is 1.3 years above the average. For Hungary life expectancy for men equals 69.2, thus 8.2 years below the NWS European average and for women 77.7 years, thus 5.1 years below the average.

Figure 6.4. Age-specific death probabilities, females, weighted average of 15 Northern, Western and Southern European countries, 2006



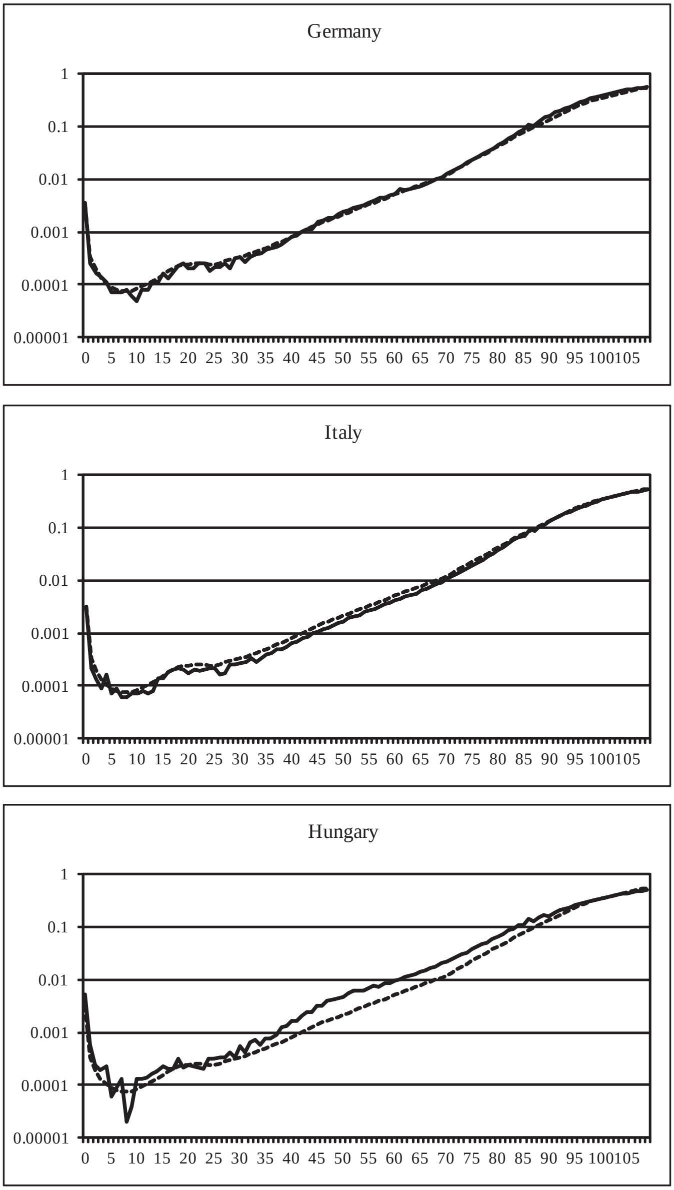
Solid line: observed values; dotted line: fitted values.

Figure 6.5a. Age-specific death probabilities of Germany, Italy and Hungary compared with average of Northern, Western and Southern Europe, 2006, males



Solid line: observed values; dotted line: average of 15 Northern, Western and Southern European countries.

Figure 6.5b. Age-specific death probabilities of Germany, Italy and Hungary compared with average of Northern, Western and Southern Europe, 2006, females

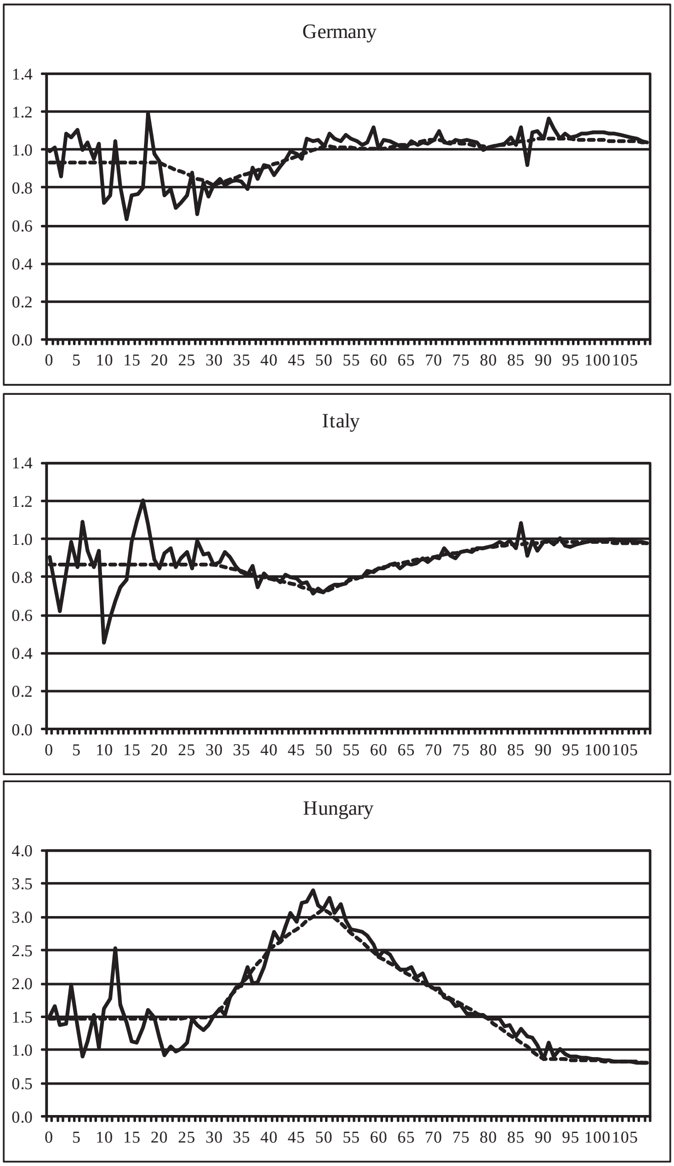


Solid line: observed values; dotted line: average of 15 Northern, Western and Southern European countries.

Figure 6.6a and *6.6b* show the risk ratios for the three countries compared with the NWS European average for men and women respectively. For each country we estimated linear splines. Since the death probabilities show large fluctuations at young ages, for estimating the spline we calculated the average value of the risk ratios for ages 0-20. For subsequent ages we use knots at intervals of ten years. *Table 6.1* shows the values of the risk ratios at the knots. The table shows that high or low life expectancies do not imply that the death probabilities across all ages are relatively low or high. The differences of the age-specific death probabilities with the NWS European average differ by age. Table 6.1 shows that until age 50 the age-specific death probabilities for Germany are slightly below the NWS European average and at higher ages slightly above the average. For Italian men the age-specific death probabilities are relatively low around age 50, but close to the average for ages 80 and older. For women the death probabilities are 20 percent lower than the average for most ages. *Tables B.1* and *B.2* in Annex B show the values of the risk ratios for all countries in this study for men and women respectively. These tables show that for other countries with low mortality the age pattern may be different than for Italy. For example, for French women life expectancy at birth is the same as for Italian women, but for French women the death probabilities for women in their 40s and 50s are higher than the NWS European average. French women have remarkably low mortality at higher ages. The low life expectancy of both Hungarian men and women is mainly caused by the very high mortality between ages 40 and 60. For men the death probabilities around age 50 are even over three times as high as the NWS European average. This pattern is typical for most Eastern European countries. Table B.1 shows that for Russia and Ukraine the death probabilities are very high at ages 30 and 40. At older ages the differences are smaller. For women the differences are considerably smaller than for men. Note that for most countries the risk ratios at the oldest ages are close to 1. This is caused by the fact that in the Human Mortality Database the age-specific death probabilities at old ages are smoothed using the same method across countries.

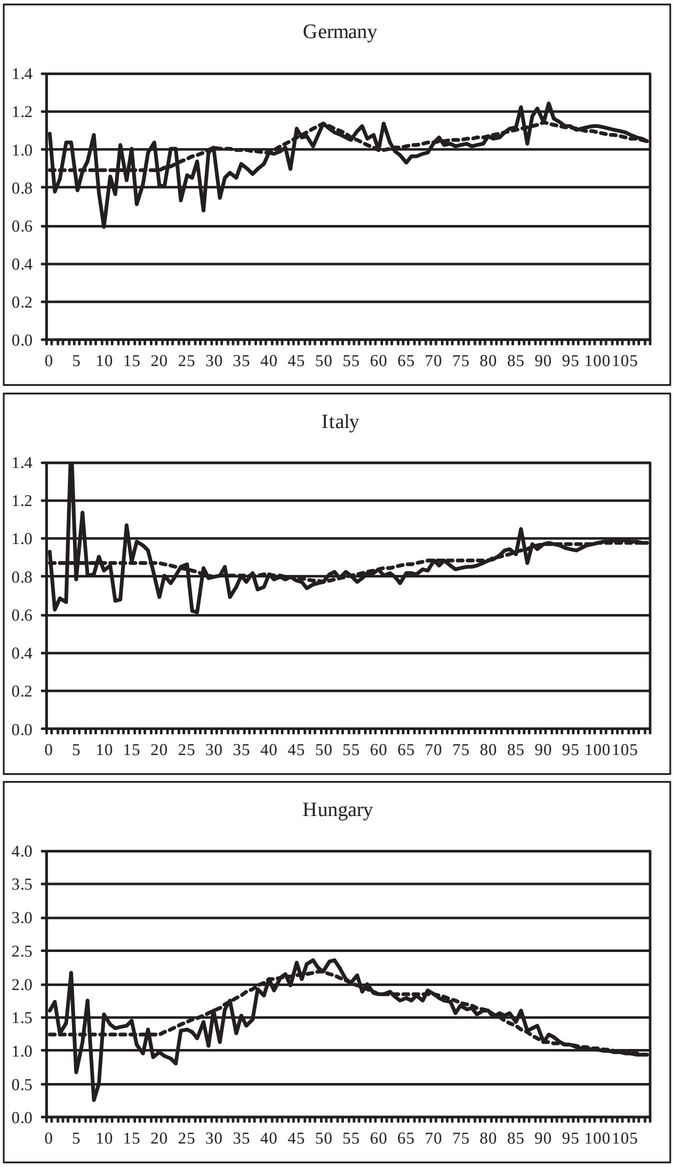
Figures 6.7a and *6.7b* show the fit of TOPALS for men and women respectivily. This is the product of the linear splines shown in figure 6.6 and the NWS European average shown in figure 6.5. Clearly the fit is accurate. *Table 6.2* compares the fit of TOPALS with those of the Heligman-Pollard and Brass models for all 26 European countries in this study. We fitted the Heligman-Pollard model to the logarithms of the death probabilities, the Brass model to the logits of the survival probabilities and TOPALS to the risk ratios of the death probabilities. For all three methods we calculated the

Figure 6.6a. Risk ratios of age-specific death probabilities of Germany, Italy and Hungary compared with average of Northern, Western and Southern European countries, 2006, males



Solid line: observed values; dotted line: linear spline.

Figure 6.6b. Risk ratios of age-specific death probabilities of Germany, Italy and Hungary compared with average of Northern, Western and Southern European countries, 2006, females



Solid line: observed values; dotted line: linear spline.

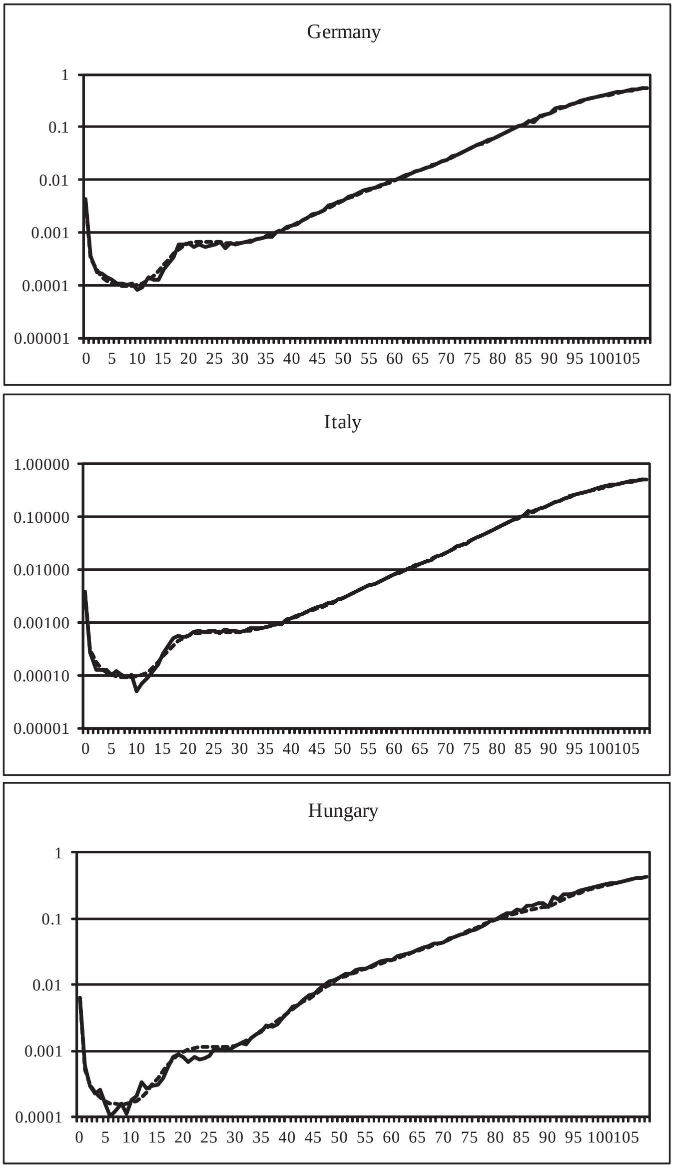
root mean square error (RMSE) for the logarithms of the death probabilities. For males the RMSE for TOPALS is smaller than for the Heligman-Pollard and Brass models for 16 countries out of the 26 countries, and for females for 15 countries. The Heligman-Pollard model performs best for ten countries for males and seven countries for females respectively. The Brass model outperforms the other two for four countries for females and for none for males. Thus on average TOPALS produces a better fit than the Heligman-Pollard and Brass models. However for many countries the differences are small. If we look at differences across the methods exceeding five percent of the RMSE only we find that TOPALS outperforms the other two methods in 12 countries for males and nine countries for females. The Heligman-Pollard model outperforms the other two methods for five countries for both males and females.

One benefit of using TOPALS rather than the Brass relational model is that TOPALS is less sensitive to the choice of the standard age schedule. For example, in the next section we use projected age-specific death probabilities of Japanese women as a standard age schedule for making scenarios of the death probabilities for European countries. The Japanese age pattern differs

Table 6.1. Values of the risk ratios of age-specific death probabilities of Germany, Italy and Hunganry compared with the average of 15 Northern, Western and Southern European countries, 2006

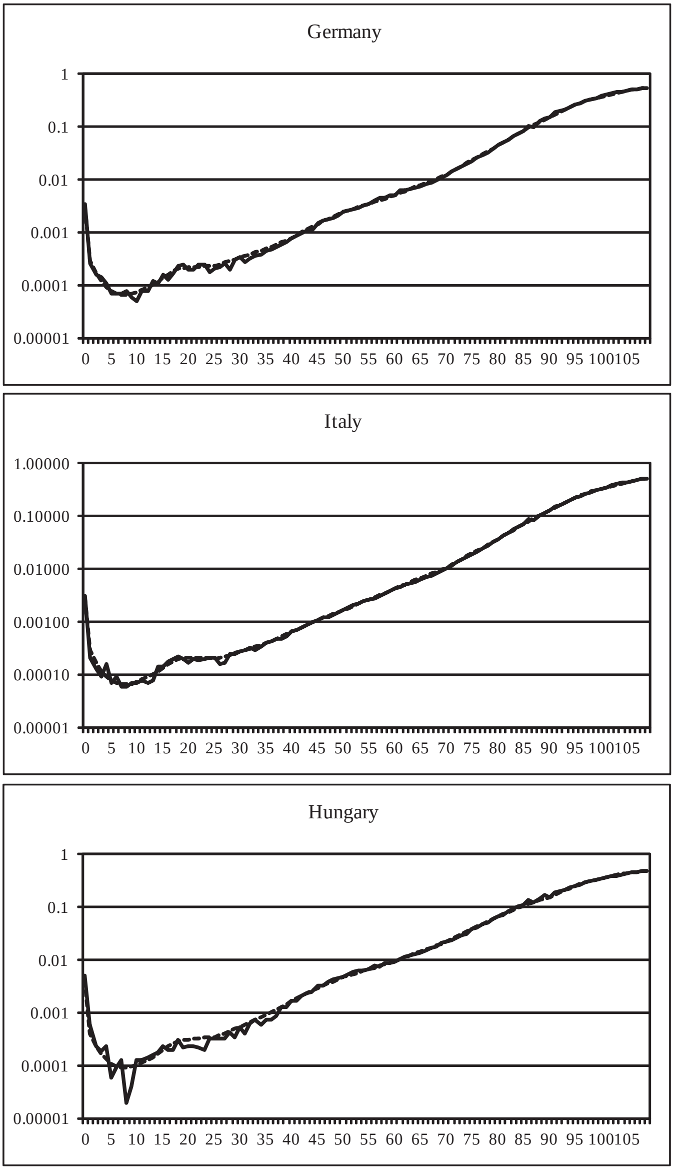
	Males			Females		
Ages	Germany	Italy	Hungary	Germany	Italy	Hungary
0-20	0.93	0.86	1.47	0.89	0.87	1.25
30	0.81	0.86	1.50	1.01	0.80	1.60
40	0.91	0.79	2.51	0.98	0.81	2.07
50	1.02	0.72	3.12	1.14	0.78	2.19
60	1.00	0.85	2.40	1.00	0.84	1.85
70	1.05	0.91	1.92	1.04	0.89	1.86
80	1.01	0.96	1.46	1.07	0.89	1.60
90	1.06	0.98	0.87	1.14	0.97	1.14
100	1.09	0.99	0.86	1.12	0.98	1.01
109	1.04	0.98	0.81	1.05	0.98	0.94
Life expectancy at birth	77.2	78.6	69.2	82.3	84.1	77.7

Figure 6.7a. Age-specific death probabilities of Germany, Italy and Hungary and fit by TOPALS, 2006, males



Solid line: observed values; dotted line: fit by TOPALS.

Figure 6.7b. Age-specific death probabilities of Germany, Italy and Hungary and fit by TOPALS, 2006, females



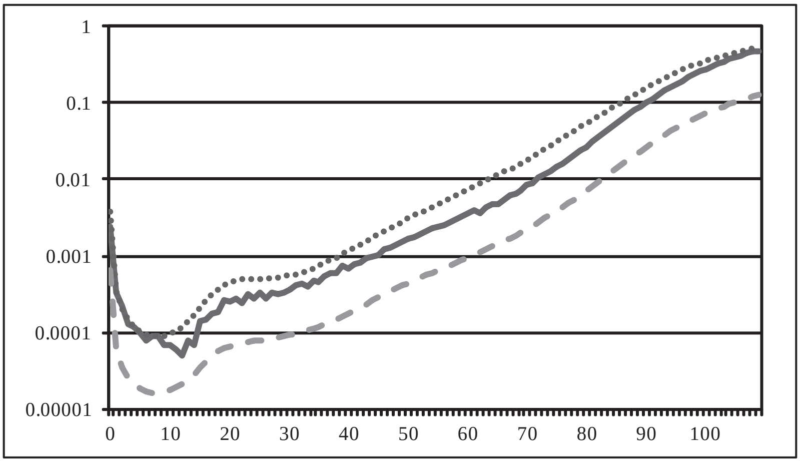
Solid line: observed values; dotted line: fit by TOPALS.

quite strongly from the European: the age-specific death probabilities at higher ages are considerably lower. Using this age schedule as standard for fitting TOPALS the RMSE increases only slightly compared with that shown in table 6.2. For males the average RMSE increases by 13 percent and for females by 10 per cent. However, using this age schedule as standard for fitting the Brass relation model the fit of the Brass model becomes rather poor: the RMSE becomes 2.9 times as high. Thus the Brass model is much more sensitive to the choice of the standard age schedule.

6.6. Scenarios of age-specific probabilities of death

The use of TOPALS for projecting death probabilities will be illustrated by making three types of scenarios. For each scenario we use the same 'target' age-specific probabilities of death. For this purpose we use age-specific probabilities of death of Japanese women. *Figure 6.8* shows the age-specific death probabilities of Japanese women in 2008 and compares these with the NWS European average. In line with Oeppen and Vaupel (2002) we assume that life expectancy at birth of Japanese women will increase linearly. In section 6.3 we showed that this implies that life expectancy at birth of Japanese women would increase to 99.6 years in 2060 (see figure 6.2). We

Figure 6.8. Age-specific death probabilities



Solid line: Japanese women in 2008.
Dotted line: average of 15 Northern, Western and Southern European countries.
Dashed line: target values (Japanese women in 2060).

Table 6.2. Goodness of fit (measured by Root mean square error) of the logarithms of age-specific probabilities of death in 26 European countries, 2006

	Males			Females		
	TOPALS	Heiligman-Pollard	Brass	TOPALS	Heiligman-Pollard	Brass
	RMSE (x 10^{-3})					
Austria	191	184	197	223	255	232
Belarus	144	255	284	214	192	249
Belgium	157	154	164	208	248	218
Bulgaria	164	145	206	255	184	353
Czech Republic	145	260	167	231	244	229
Denmark	201	190	207	232	234	271
Estonia	319	355	347	427	418	417
Finland	334	283	343	307	319	296
France	67	136	133	95	212	135
Germany	91	95	117	100	176	132
Hungary	143	433	324	232	327	302
Ireland	275	247	297	341	306	330
Italy	105	108	118	106	172	112
Latvia	364	343	408	411	369	419

Lithuania	188	201	308	222	222	265
Netherlands	148	110	222	173	192	184
Norway	268	268	291	360	362	350
Poland	67	171	162	116	170	159
Portugal	128	170	200	210	214	247
Russia	72	126	255	93	88	246
Slovakia	232	236	246	262	235	322
Spain	129	131	130	106	214	142
Sweden	188	292	221	228	298	229
Switzerland	196	217	214	274	299	282
Ukraine	126	102	251	122	92	253
United Kingdom	76	91	94	90	116	117
Average	174	204	227	217	237	250

calculated age-specific probabilities of death by reducing the age-specific probabilities of death of 2008 by 74 percent which corresponds with a life expectancy at birth of 99.6 years. Since this produces a rather irregular age pattern we used TOPALS to smooth the age pattern. For this purpose we used the NWS European average as standard age curve, *i.e.* the same age schedule that we used as standard in section 6.5. Figure 6.8 shows the smooth target pattern. Instead of assuming the same percentage decrease in death probabilities across all ages to calculate the target pattern we could have assumed a change in the age pattern, *e.g.* we could have assumed that the decline at older ages will be larger than at young ages. However, note that the difference between the death probabilities of Japanese women in 2008 and the NWS European average between ages 50 and 85 is larger than that at younger ages. This implies that an equal percentage reduction of death probabilities of Japanese women across all ages produces a target pattern which shows a stronger reduction of death probabilities between ages 50 and 85 compared with the current NWS European than for younger ages.

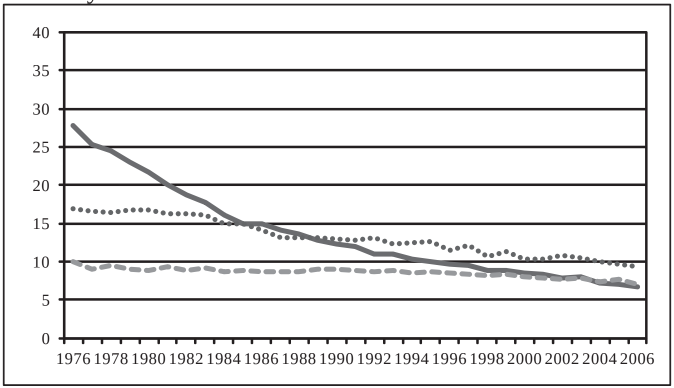
We calculated three scenarios based on this same target pattern. For the ages at the knots we make time series of risk ratios by dividing the death probabilities for each country by the target values. *Figures 6.9a and 6.9b* show the time series of risk ratios for three selected ages for Germany, Italy and Hungary, for men and women for the period 1976-2006. The scenarios are based on projections of these risk ratios into the future. The scenarios differ by the speed with which the target values will be reached, *i.e.* the speed with which the risk ratios move towards 1. The figures show that the rate of decline has differed across ages and across countries. For the youngest age group 0-20 years the decline has been strong. For age 90 there has been a moderate decline only. For Hungary, the development of mortality of middle-aged men showed an increase in the 1970s and 1980s. For the first scenario we estimate the partial adjustment model for each country separately. We call this the Baseline scenario. The second scenario assumes that the values of φ are equal for all countries. This scenario assumes that there will be a similar trend across European countries. We call this the Convergence scenario. The third scenario assumes that the future decrease in death probabilities will exceed that in the last three decades. We label this as the Acceleration scenario.

6.6.1. Baseline Scenario

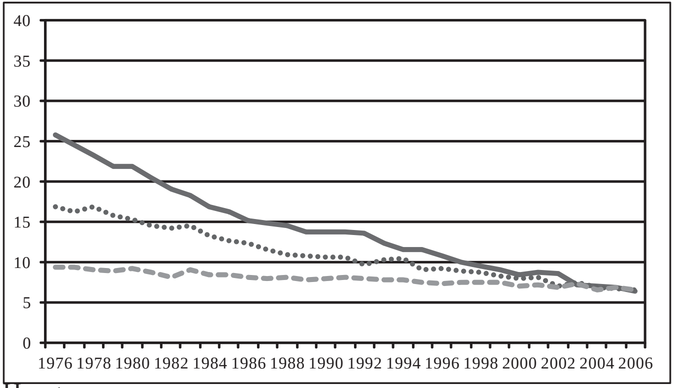
Oeppen and Vaupel (2002) suggest that life expectancy for individual countries can be projected by assuming that the gap with the best-practice level stays the same. This would imply a linear increase in life expectancy

Figure 6.9a. Risk ratios compared with target pattern, Germany, Italy and Hungary, ages 0-20, 50 and 90, 1976-2006, men

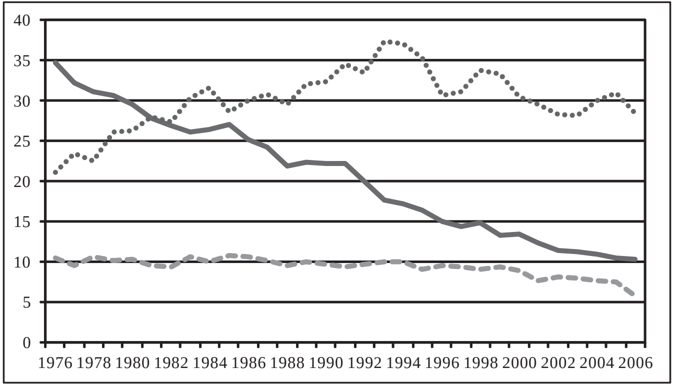
Germany



Italy



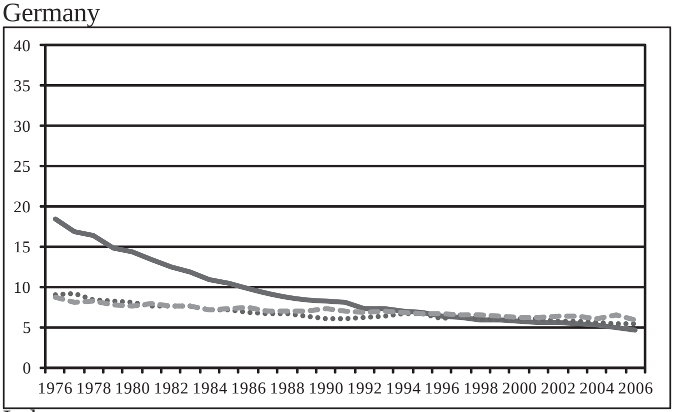
Hungary



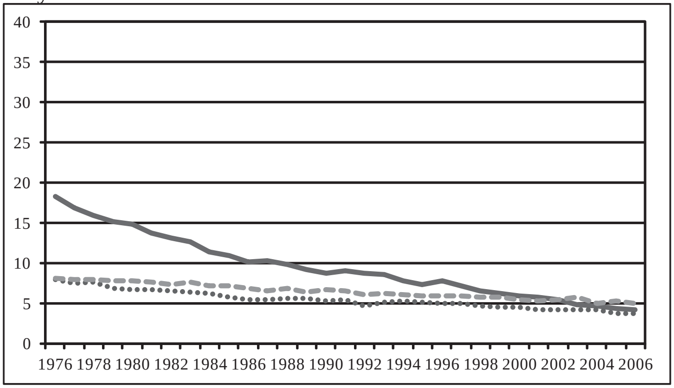
Solid line: Age 0-20; dotted line: Age 50; dashed line: Age 90.

Figure 6.9b. Risk ratios compared with target pattern, Germany, Italy and Hungary, ages 0-20, 50 and 90, 1976-2006, women

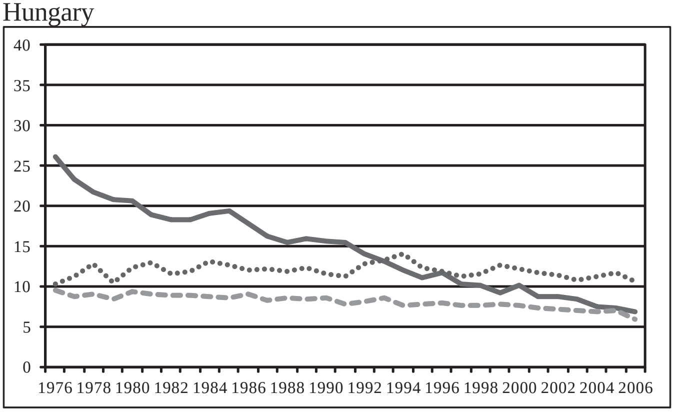
Germany



Italy



Hungary



Solid line: Age 0-20; dotted line: Age 50; dashed line: Age 90.

for all countries. However, Oeppen and Vaupel acknowledge that life expectancy has not increased with the same speed across all countries during the last century. Therefore we assume that death probabilities move towards the best practice levels, but that the speed may differ. Thus we specify the year that the best practice level will be reached not *a priori*: it depends on the value of φ. The Baseline scenario projects the risk ratios using equation (15). This scenario can be considered as an extrapolation of past trends. We estimate the parameter φ for each country at each knot separately for the period 1976-2006 by minimizing the sum of squared residuals of equation (14). This estimation period is similar to the period that Eurostat chooses as basis for their latest scenarios (Lanzieri, 2009). The values of φ indicate how strongly the observed probabilities of death move towards the low levels corresponding with a life expectancy of 99.6 years. *Table 6.3* shows the estimated values of φ for Germany, Italy and Hungary. If φ is close to 1, the projections will move very slowly to the target value, and thus death probabilities will decline slowly. If φ equals 1, the projected value equals the last observed value and does not move towards the target level. This is the case for Hungarian men at ages 50 and 60. For Italy for most ages the values of φ are lower than for the other two countries, thus the model will project a more rapid decline of death probabilities for Italy. *Tables B.3* and *B.4* in Annex B show the estimated values of φ for all countries in this study for men and women respectively. The tables show that the estimated values of φ for older ages tend to be closer to one than for younger ages. The explanation is that at the older ages there has been a slow decrease in death probabilities. This implies that the Baseline scenario projects only limited decrease at older ages in the future. Note that even though we assume the same target levels of probabilities of death across all countries and for both sexes, this does not imply that this is a convergence scenario. The projections differ across countries for two reasons: the differences between the current and target values of the death probabilities differ and the values of φ are different across countries.

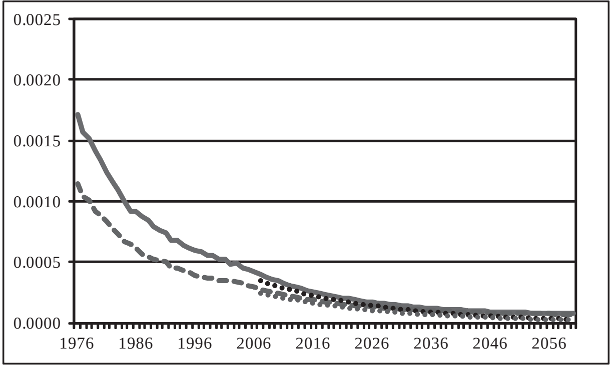
By multiplying the projected risk ratios by the target values of the probabilities of death we obtain projections of the death probabilities for each country. *Figures 6.10a, 6.10b* and *6.10c* show the projections for three selected ages for Germany, Italy and Hungary. The figures compare the Baseline scenario with the projections according to the Lee-Carter model. Generally the projections according to the Baseline scenario are rather close to the Lee-Carter projections. However, for German and Hungarian men aged 90 the jump-off point of the Lee-Carter projections differs from the last point in the observation period. The explanation is that the Lee-Carter

Table 6.3. Estimated values of coefficient φ of partial adjustment model, Baseline scenario for Germany, Italy and Hungary and Convergence and Acceleration scenarios

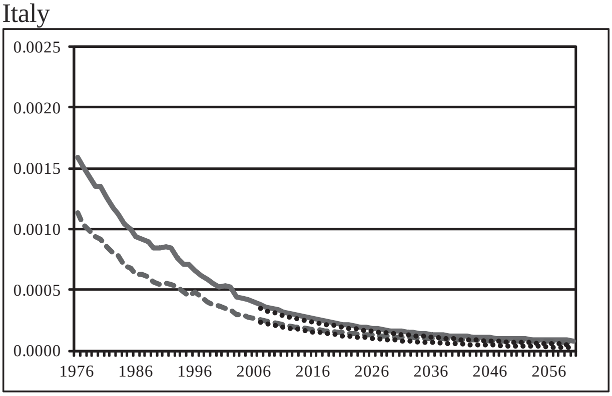
	0-20	30	40	50	60	70	80	90	100	109
	Males									
Baseline scenario										
Germany	0.9415	0.9702	0.9711	0.9794	0.9812	0.9737	0.9795	0.9865	0.9966	1.0000
Italy	0.9592	0.9829	0.9341	0.9563	0.9703	0.9755	0.9784	0.9826	0.9944	0.9979
Hungary	0.9574	0.9823	0.9966	1.0000	1.0000	0.9908	0.9889	0.9819	0.9859	0.9888
Convergence scenario	0.9546	0.9857	0.9794	0.9773	0.9794	0.9756	0.9811	0.9871	0.9949	0.9987
Acceleration scenario	0.9116	0.9715	0.9588	0.9548	0.9642	0.9517	0.9622	0.9747	0.9899	0.9974
	Females									
Baseline scenario										
Germany	0.9425	0.9748	0.9689	0.9777	0.9769	0.9702	0.9746	0.9834	0.9947	0.9992
Italy	0.9449	0.9730	0.9707	0.9671	0.9685	0.9715	0.9715	0.9804	0.9908	0.9971
Hungary	0.9526	0.9745	0.9815	0.9961	0.9905	0.9863	0.9853	0.9828	0.9906	0.9945
Convergence scenario	0.9537	0.9817	0.9790	0.9755	0.9781	0.9734	0.9749	0.9829	0.9932	0.9983
Acceleration scenario	0.9057	0.9642	0.9576	0.9517	0.9563	0.9481	0.9499	0.9659	0.9865	0.9966

Figure 6.10a. Projections of death probabilities for ages 0-20 years, Germany, Italy and Hungary, observations 1976-2006, Baseline scenario and Lee-Carter projections 2007-2060

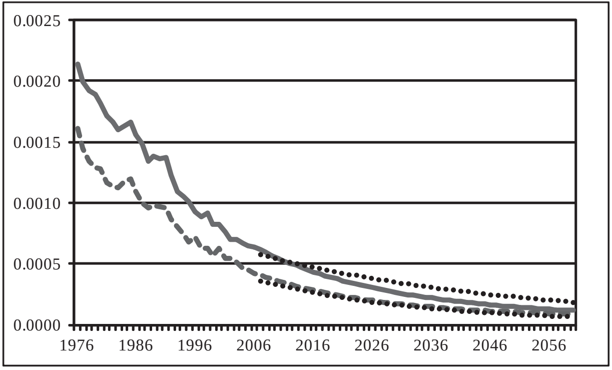
Germany



Italy

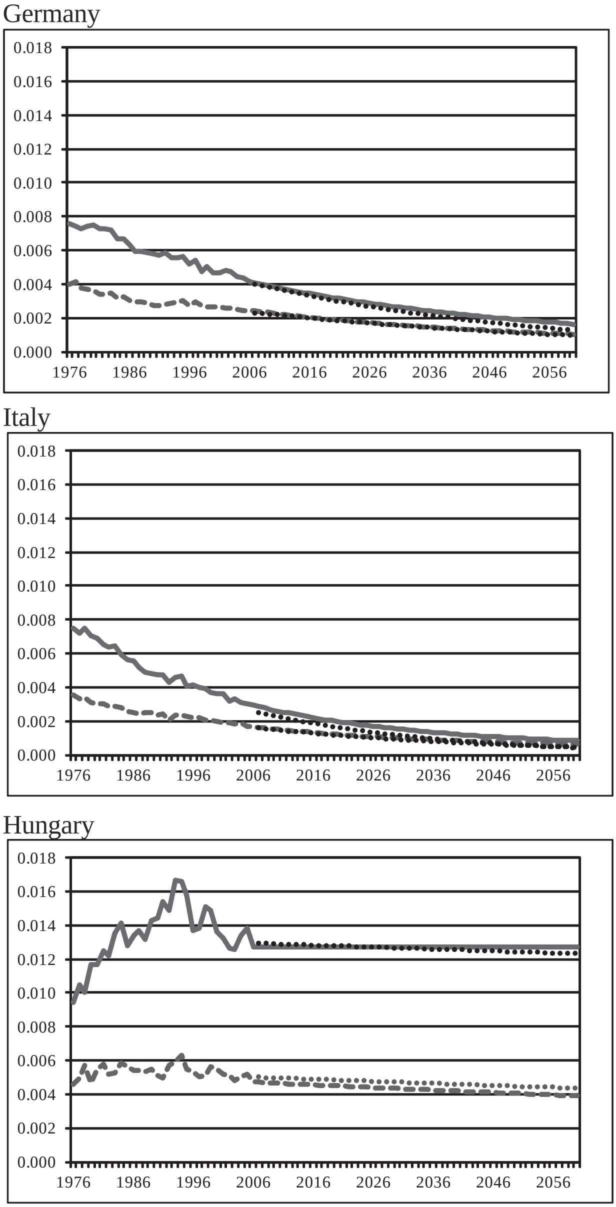


Hungary



Solid line: Men; dashed line: Women; dotted line: Lee-Carter model.

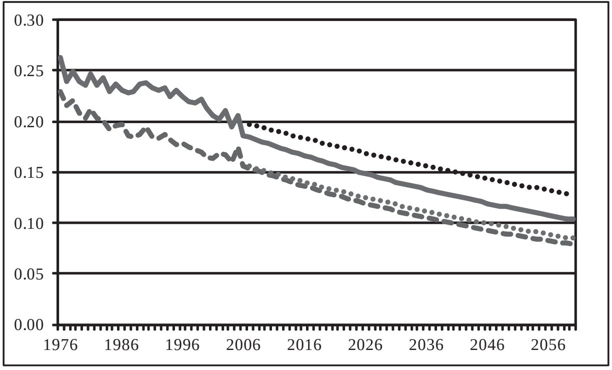
Figure 6.10b. Projections of death probabilities for age 50 years, Germany, Italy and Hungary, observations 1976-2006, Baseline scenario and Lee-Carter projections 2007-2060



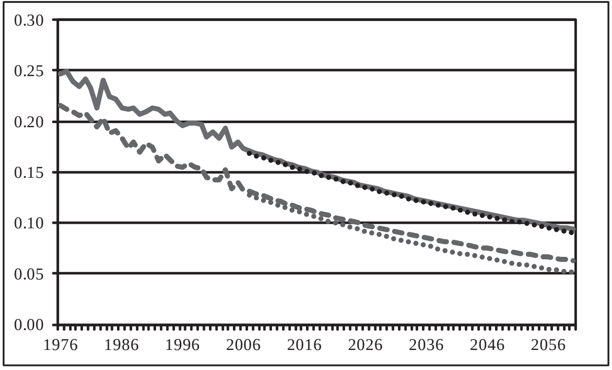
Solid line: Men; dashed line: Women; dotted line: Lee-Carter model.

Figure 6.10c. Projections of death probabilities for age 90 years, Germany, Italy and Hungary, observations 1976-2006, Baseline scenario and Lee-Carter projections 2007-2060

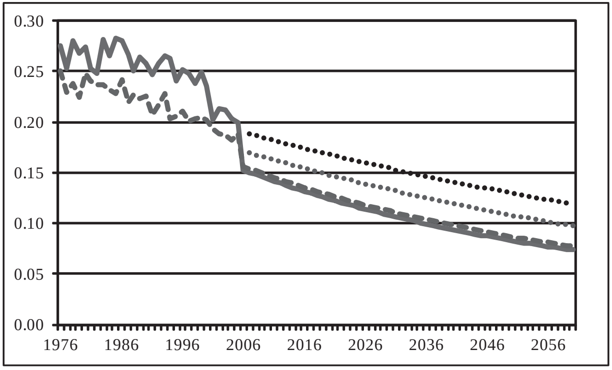
Germany



Italy



Hungary



Solid line: Men; dashed line: Women; dotted line: Lee-Carter model.

projections are based on a random walk with drift projection starting from the last estimated value of the death probability according to equation (4) rather than from the last observed value. We discussed this issue at the end of section 6.4 (see figure 6.3). For this reason Lee and Miller (2001) suggest to use the last observed value as jump-off value for the projections of the Lee-Carter model. This would make the Lee-Carter projections closer to the Baseline scenario.

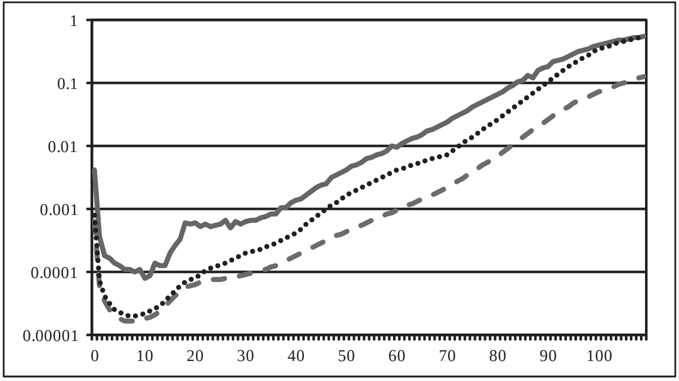
Figures 6.11a and *6.11b* compare the age-specific death probabilities projected by the Baseline scenario with the pattern in the last observation year and with the target pattern for men and women respectivily. The figures show that for young ages the projected death probabilities are rather close to the target pattern, whereas for the oldest ages the projections are close to the last observed values. This reflects the relatively strong decline in death probabilities at young ages and the slow decline at older ages during the observation period. For Hungarian men the death probabilities at middle ages hardly decline. This is the result of the fact during a large part of the estimation period death probabilities at middle age increased and that in more recent years there has been only a moderate decrease as a consequence of which death probabilities in 2006 were higher than in 1976.

Tables 6.4 and *6.5* show the values of life expectancy at birth in 2060 for men and women respectively which result from the projections of the age-specific death probabilities according to the Baseline scenario. In 2008 life expectancy at birth for Japanese women equalled 86 years. Table 6.5 shows that according to Baseline scenario in all Northern, Western and Southern European countries life expectancy of women is expected to reach that level well before 2060. In most Eastern European countries that level would not be reached before 2060. Table 6.4 shows that for men life expectancy at birth for most Northern, Western and Southern European countries would approach the current world record level around 2060. The target pattern assumes an increase in life expectancy to 99.6 years in 2060. None of the European countries is expected to approach that level in this century.

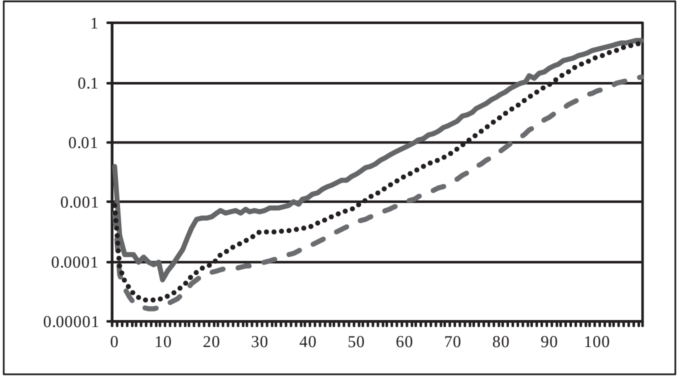
Tables 6.4 and 6.5 show the Lee-Carter projections as well. For most Northern, Western and Southern European countries the differences between the Baseline scenario and the Lee-Carter projection are moderate. On average the projected life expectancy according to the Baseline scenario in 2060 for the 15 Northern, Western and Southern European countries is 0.2 years lower than the Lee-Carter projection. For four Eastern European countries (Belarus, Bulgaria, Russia and Ukraine) the differences are large.

Figure 6.11a. Age-specific death probabilities Germany, Italy and Hungary in 2006 and 2060, men

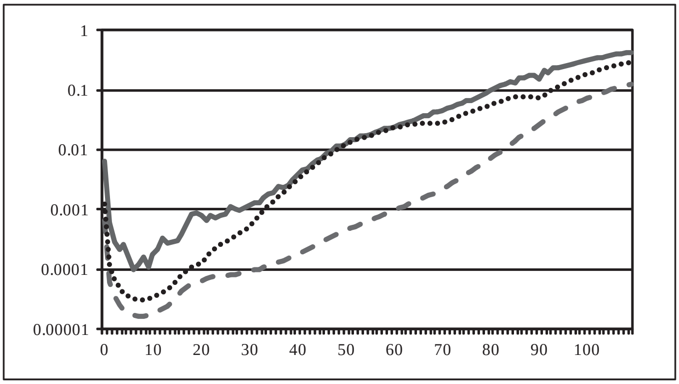
Germany



Italy



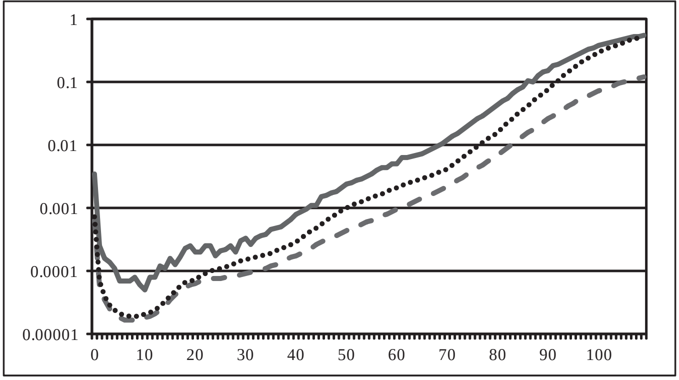
Hungary



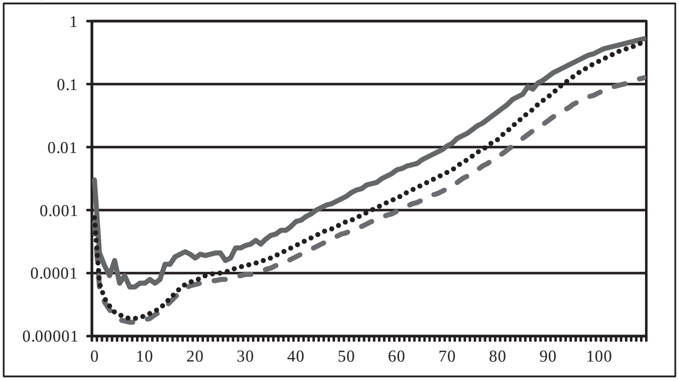
Solid lines: 2006; dotted line: baseline scenario for 2060; dashed line: Target pattern.

Figure 6.11b. Age-specific death probabilities Germany, Italy and Hungary in 2006 and 2060, women

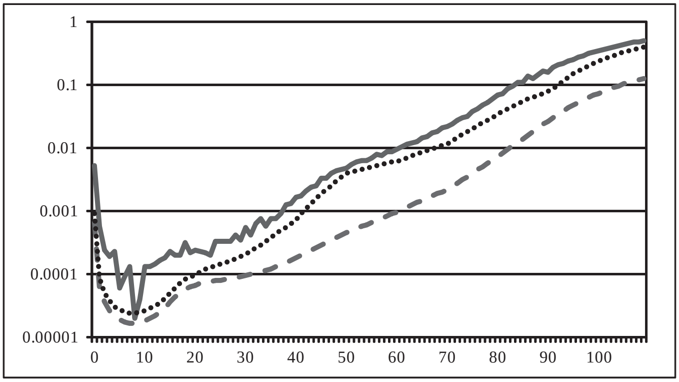
Germany



Italy



Hungary



Solid lines: 2006; dotted line: baseline scenario for 2060; dashed line: Target pattern.

The explanation is that the projections based on the partial adjustment projections are restricted since it is assumed that $\varphi \leq 1$. Thus if at certain ages death probabilities have increased in the observation period, the model projects a constant future level, whereas the projections of the Lee-Carter model project an increase in death probabilities at those ages. Figure 6.11 shows that the age pattern of death probabilities projected by TOPALS is smooth. In contrast, the age pattern projected by the Lee-Carter model (not shown here) is rather irregular. Lee and Carter (1992) suggest to use five-years age groups. However, since mortality rates increase strongly by age, this is a rather crude approximation. As an alternative Renshaw and Haberman (2003) and Currie *et al*. (2004) suggest to smooth the age-specific death probabilities using splines.

6.6.2. Sensitivity analysis

The projections of the Baseline scenario depend on different choices: (1) the choice of the estimation period for estimating the parameter of the partial adjustment model, (2) the choice of the partial adjustment model for making projections and (3) the choice of the target pattern for calculating the risk ratios. It is useful to examine how sensitive the projections are to these choices. *Table 6.6* compares the Baseline scenarios for Germany, Italy and Hungary with projections based on alternative assumptions. The table shows that choosing a shorter, more recent estimation period for estimating the value of φ would result in considerably higher projections for Hungary, especially for men. The reason is that the development of mortality in Hungary in recent years has been more favourable than in the 1970s and 1980s as was shown in figure 6.9. For Germany and Italy the effect of choosing a different estimation period is clearly smaller. If the Lee-Carter model is estimated for a shorter period the effects on the projections are similar.

If instead of using the partial adjustment model (15) we use the random walk model with drift (18) for making projections for Germany and Italy the projections become higher. The reason is that the projections of the random walk model with drift are unconstrained. The projections of the random walk model are closer to those of the Lee-Carter model as would be expected since the parameter k_t of the Lee-Carter model is projected by a random walk model as well. However, the projections are not equal to those of the Lee-Carter model. The reason was explained at the end of section 6.4. For Hungarian men the random walk model projection is very low. The explanation is that the random walk model projects an increase in death probabilities of men in their 50s and 60s, since the death probabilities in 2006 exceeded those in 1976. In using the partial adjustment model we assume that $\varphi \leq 1$. As noted

Table 6.4. Projections of life expectancy at birth in 2060, males

	Observed in 2006	Baseline scenario	Convergence scenario	Acceleration scenario	Lee Carter model	EUROPOP 2008	Linear projection life expectancy
Austria	77.1	86.6	86.2	90.7	87.5	84.9	93.1
Belarus	63.6	67.9	78.3	86.9	59.9		57.6
Belgium	76.5	86.6	85.6	90.4	85.8	84.4	90.3
Bulgaria	69.2	72.6	81.4	88.3	69.6	81.6	69.6
Czech Republic	73.5	82.0	83.8	89.5	82.3	83.2	85.0
Denmark	75.9	82.5	85.1	90.2	82.2	84.3	84.9
Estonia	67.4	72.7	80.1	87.8	72.2	80.8	72.5
Finland	75.8	86.5	85.1	90.2	86.1	84.3	90.7
France	77.2	86.7	86.4	91.1	87.4	85.1	91.6
Germany	77.2	86.4	86.0	90.7	86.9	84.9	92.8
Hungary	69.2	73.9	82.3	89.5	72.9	81.9	73.6
Ireland	77.3	87.1	85.8	90.6	85.8	85.2	91.5
Italy	78.6	87.5	86.8	91.1	88.9	85.5	94.4
Latvia	65.6	70.2	78.8	87.3	70.8	80.5	68.6

Lithuania	65.3	69.2	79.6	88.8	67.0	80.4	63.5
Netherlands	77.6	83.8	85.9	90.5	84.0	84.9	88.6
Norway	78.1	85.9	86.4	90.7	84.8	85.2	89.2
Poland	70.9	76.3	82.8	89.4	76.3	82.5	78.2
Portugal	75.5	86.1	85.2	90.3	85.3	84.1	93.5
Russia	60.3	62.7	76.1	85.9	58.3		57.0
Slovakia	70.4	74.8	82.1	88.9	74.5	82.0	76.5
Spain	77.6	85.9	86.4	91.0	85.8	84.9	89.9
Sweden	78.7	85.8	86.5	90.8	85.9	85.4	90.4
Switzerland	79.1	87.8	87.2	91.3	87.6	85.8	92.5
Ukraine	62.3	64.6	77.3	86.5	58.3		56.2
United Kingdom	77.2	86.2	86.2	90.9	86.6	85.0	91.0

Table 6.5. Projections of life expectancy at birth in 2060, females

	Observed in 2006	Baseline scenario	Convergence scenario	Acceleration scenario	Lee Carter model	EUROPOP 2008	Linear projection life expectancy
Austria	82.7	90.1	89.8	93.2	91.1	89.2	96.3
Belarus	75.5	79.7	86.5	92.1	73.7		74.1
Belgium	82.2	90.7	89.7	93.3	90.4	88.9	94.3
Bulgaria	76.3	81.2	86.6	91.9	78.4	86.5	80.6
Czech Republic	79.9	86.5	88.4	92.6	87.7	87.8	90.1
Denmark	80.5	86.7	89.2	93.2	87.7	88.4	87.4
Estonia	78.6	86.0	88.0	92.6	82.7	87.5	86.1
Finland	82.8	89.7	89.7	93.3	90.4	89.3	94.5
France	84.1	91.5	91.0	94.0	92.6	90.1	96.6
Germany	82.3	89.6	89.6	93.2	90.5	89.1	95.2
Hungary	77.7	84.4	87.8	92.8	84.3	87.3	86.7
Ireland	81.9	90.5	89.8	93.4	90.0	89.2	95.2
Italy	84.1	91.4	90.6	93.7	93.6	90.0	98.2
Latvia	76.5	82.0	86.7	92.0	79.9	86.8	80.4

Lithuania	77.1	82.0	87.1	92.2	78.1	86.9	79.2
Netherlands	81.9	87.4	89.4	93.1	86.6	88.9	89.1
Norway	82.7	88.6	89.8	93.2	89.2	89.2	90.8
Poland	79.6	85.5	88.5	92.9	85.5	88.0	88.5
Portugal	82.2	90.5	89.4	93.1	90.8	88.8	99.3
Russia	72.4	75.9	84.7	91.2	69.2		73.6
Slovakia	78.4	85.3	87.9	92.6	84.4	87.4	86.2
Spain	84.1	91.0	90.5	93.6	91.9	89.6	97.4
Sweden	82.9	89.2	90.0	93.4	88.8	89.3	91.8
Switzerland	84.0	91.6	90.7	93.7	90.9	89.9	94.4
Ukraine	73.8	78.0	85.5	91.5	71.9		73.0
United Kingdom	81.5	89.0	89.6	93.4	88.7	88.9	92.1

above for Hungarian men the estimated value of φ at knots 50 and 60 equals 1 (see table 6.3). This implies that the projection equal the last value in the observation period.

The projections of the partial adjustment model are based on assuming target values of the death probabilities that would result in a life expectancy at birth of 99.6 years. If higher target values of the death probabilities would be assumed, the projected life expectancy would be lower. However table 6.6 shows that the change in the projected value is considerably smaller than the difference between the target values. If it would be assumed that the target level of life expectancy equals 95 years instead of 99.6 years the projected life expectancy for men would hardly be affected. For women the projections would be 0.3 to 1.1 years lower. If the target levels of death probabilities are chosen so that they result in a life expectancy at birth of 110 years rather than 99.6 years, the projected life expectancy for Italian men and women would become about 1 year higher. For the other two countries the differences would be considerably smaller. The explanation is that the estimated values of φ change if another target level is chosen. If the target value is lower the estimated value of φ becomes higher, which implies that the model projects that it will take much more time before that lower target level will be reached. Instead of assuming the same rate of decline across all ages one could specify target values assuming a different age pattern. For example, one might assume that the decrease in death probabilities at older ages is larger than at younger ages. However, that would not result in strongly different projections, since the estimated values of φ at older ages are close to one. This would lead to different projections only if one would assume that in the future different values of φ would apply than in the observation period. The conclusion of the sensitivity analysis is that even though different choices would result in different projections, the differences are moderate only.

6.6.3. Convergence scenario

There is ample empirical evidence that there has been a converging tendency in mortality declines during the last decades (Wilson, 2001; White, 2002; Janssen *et al.*, 2004; Bongaarts, 2006 and Lanzieri, 2009). Life expectancy has increased more strongly in countries that had relatively low life expectancies. The latest Eurostat projections, EUROPOP2008, are based on the assumption that there is a converging trend in the long run (Lanzieri, 2009). The main underlying assumption is that the socioeconomic differences between Member States of the European Union will fade out in the long run (Lanzieri, 2009). The scenario assumes that advanced medical

Table 6.6. Sensitivity analysis of projections of life expectancy at birth in 2060, Germany, Italy and Hungary

	Germany		Italy		Hungary	
	men	women	men	women	men	women
Baseline scenario	86.4	89.6	87.5	91.4	73.9	84.4
Estimation period 1986-2006	86.5	89.2	88.2	91.7	78.5	86.7
Random walk model	87.4	90.4	88.7	92.8	70.5	84.2
Target value life expectancy = 95 years	86.0	88.5	87.6	90.7	73.8	84.1
Target value life expectancy = 110 years	87.1	89.9	88.8	92.6	74.1	84.7

Note: Baseline scenario: estimation period 1976-2006; target value of life expectancy = 99.6 years.

techniques will be accessible in each country and healthy life styles will be homogeneously spread in Europe. Gender differences in life style are assumed to diminish. Differences in smoking between men and women have decreased. Moreover, improvement of standards of living will have a stronger positive effect on male life expectancy as they are more sensitive to economic conditions, which will narrow the gender gap in life expectancy (Brunner, 1997). The convergence scenario of EUROPOP2008 assumes that full convergence will be reached in 2150. In specifying our Convergence scenario we follow a different approach. We follow the recommendation by Janssen and Kunst (2007) that rather than assuming that mortality rates of different countries will reach the same target level by the end of the projection period, the average mortality change among similar countries should be used as the basis for the long-run projection of the mortality levels for the individual countries. One reason is, that as Bongaarts (2006) argues, the average pace of mortality decline across a number of countries reflects the effects of improvements in medical technology and behaviour whereas country-specific deviations are unpredictable. Tuljapurkar *et al.* (2000) found that the time-dependent parameter of the Lee-Carter model follows a common pattern for the G7 countries. Li and Lee (2005) argue that long-run forecasts for individual countries can be improved by estimating the time-dependent parameter in the Lee-Carter model for a group of countries. Thus there may be two reasons for specifying a Convergence scenario. One obvious reason is that one assumes that there is a converging tendency among European countries. But another important reason is that estimating a common long-run trend for a group of countries may provide a more reliable basis for long-run projections as it excludes the effect of temporary deviations in individual countries. Therefore we specified a Convergence scenario by estimating the values of φ for time series of the average probabilities of death of 15 Northern, Western and Southern European countries. We calculated weighted averages using population size as weight. We did not include the Central and Eastern European countries in the estimation of the common parameter as these have clearly followed a different development in the sample period. The estimated values of φ are given in table 6.3. For the Convergence scenario we use these estimated values of φ for making projections for all European countries including the Central and Eastern European countries.

EUROPOP2008 projects the levels of the age-specific death probabilities in 2100 by applying the Lee-Carter model to the average across 12 Northern, Western and Southern European countries. Compared with the 15 countries we use, Eurostat did not include Ireland, Norway and Switzerland. The Lee-Carter model is fitted to the period 1977-2005. The levels for 2060 are obtained by exponential interpolation. This results in a rather strong converging trend. Whereas in 2006 the difference between the lowest and highest values of life expectancy shown in tables 6.4. and 6.5. were 13.8 years for men and 7.9 years for women respectively, Eurostat assumes that this will be decreased to 5.4 years for men and 3.6 years for women in 2060 (we exclude Belarus, Russia and Ukraine from these comparisons as Eurostat did not produce scenarios for these three countries). According to the Baseline scenario the differences between the countries with the highest and lowest levels of life expectancy in 2060 would be larger than in 2006. This is mainly due to the relatively small increases projected for the Eastern European countries. Our Convergence scenario projects that the differences between the lowest and highest life expectancy in 2060 would become 8.4 years for men and 4.3 years for women. This is smaller than the differences in 2006 but larger than the Eurostat scenario. Again, this can mainly be explained by the Eastern European countries. If we look at the Northern, Western and Southern European countries, the Convergence scenario projects that the differences between the lowest and highest values of life expectancy would become 2.1 years for men and 1.8 years for women compared with 1.7 years for both men and women according to Eurostat.

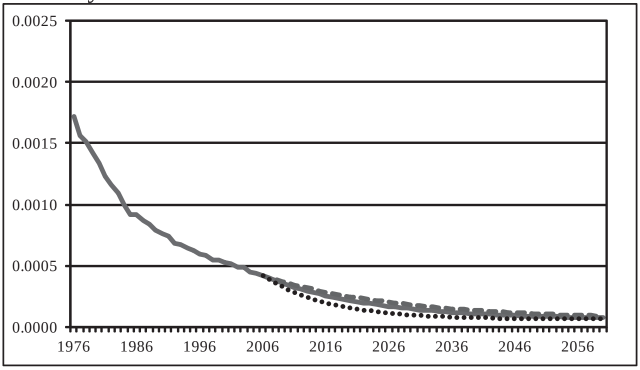
Figures 6.12a, 6.12b and *6.12c* compare the projected death probabilities for men according to the Convergence scenario with those according to the Baseline scenario for three ages. For Germany the projections according to both scenarios are similar, which can be explained by the fact that Germany is a rather average country. For Italy there is slightly less decrease in the death probabilities according to the Convergence scenario. The reason is that the Italian decrease according to the Baseline scenario is above average. Figure 6.12b shows that for Hungarian middle aged men the decrease in death probabilities according to the Convergence scenario is considerably stronger than according to the Baseline scenario. Figure 6.12c shows that at the oldest ages the opposite is true. Tables 6.4 and 6.5 show that for all Central and Eastern European countries life expectancy according to the Convergence scenario is considerably higher than according to the Baseline scenario. On average the Convergence scenario is 3.6 years higher for men and 2.0 years for women than the Baseline scenario. For most Northern, Western and Southern European countries the differences between both scenarios are under one year. There are two clear exceptions: for Denmark and the Netherlands the Baseline scenario projects only a moderate increase since both countries have shown a below average increase in the observation period. For both countries life expectancy according to the Convergence scenario is two years higher than according to the Baseline scenario.

6.6.4. Acceleration scenario

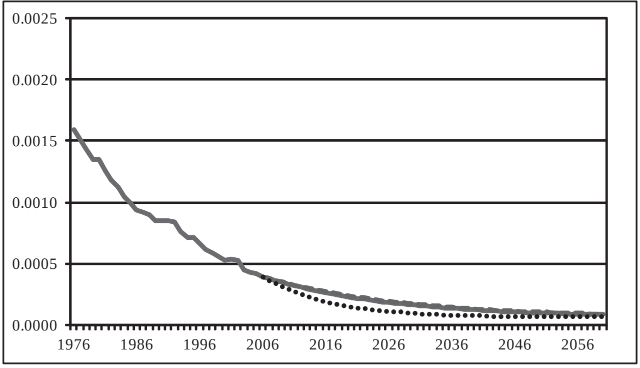
The future may differ from the past. Even though mortality has declined steadily for a long period, the causes of this decline have changed over time. In the past the main cause of increase in life expectancy at birth was a decline in infant mortality. This was mainly caused by advances in hygiene, medicine and improvement of living conditions. In the first half of the 20th century the main cause of death were infectious diseases. In the second half of the century death by infectious diseases has declined strongly across all ages. The main causes of death have become cardiovascular diseases and cancer. During the last 50 years mortality by cancer has increased. One main cause has been smoking. In recent decades in many countries mortality by cardiovascular diseases has decreased as a consequence of advances in prevention and treatment. In recent years mortality from lung cancer has

Figure 6.12a. Projections of death probabilities for ages 0-20 years Germany, Italy and Hungary, observations 1976-2006, Baseline, Convergence and Acceleration scenarios 2007-2060, men

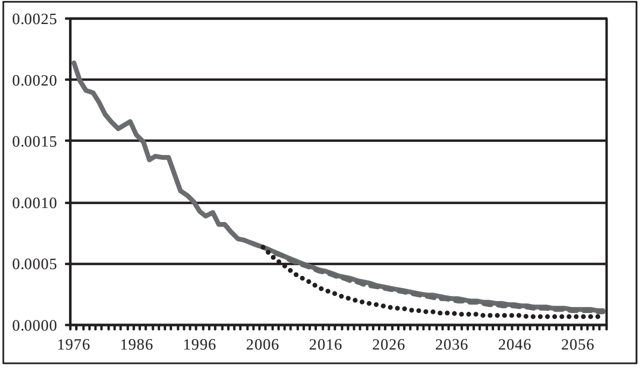
Germany



Italy



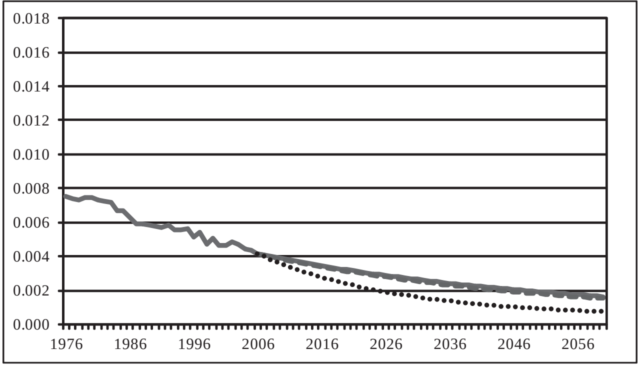
Hungary



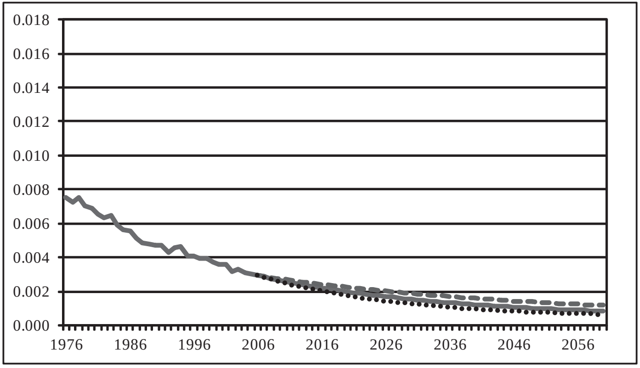
Solid line: Baseline scenario; dashed line: Convergence scenario; dotted line: Acceleration scenario.

Figure 6.12b. Projections of death probabilities for age 50 years Germany, Italy and Hungary, observations 1976-2006, Baseline, Convergence and Acceleration scenarios 2007-2060, men

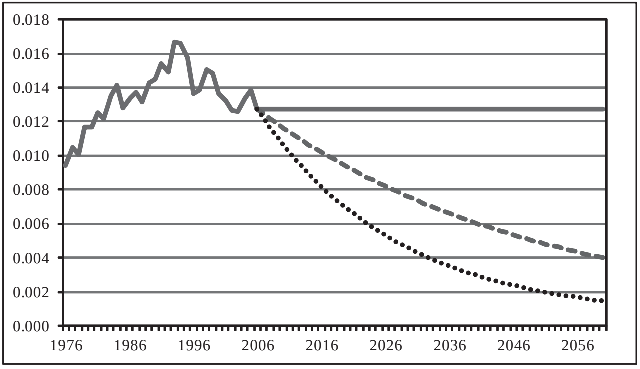
Germany



Italy



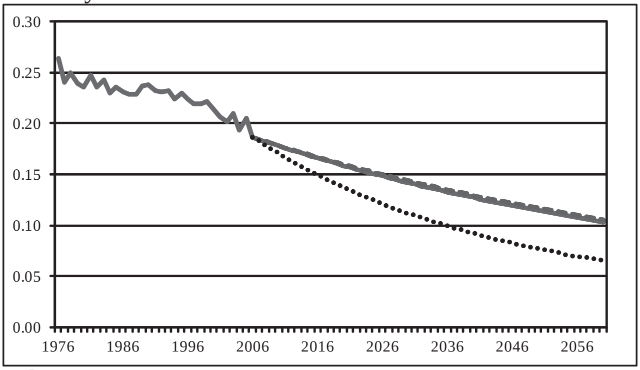
Hungary



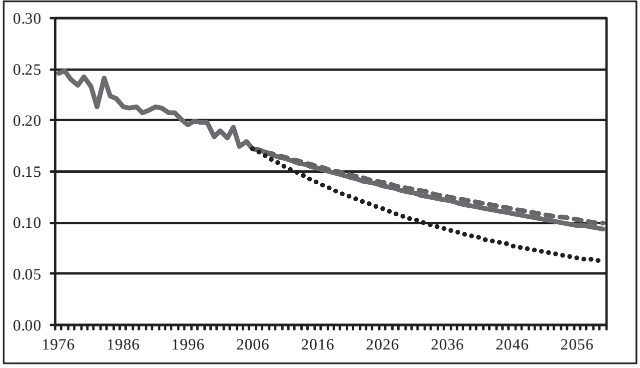
Solid line: Baseline scenario; dashed line: Convergence scenario; dotted line: Acceleration scenario.

Figure 6.12c. Projections of death probabilities for age 90 years Germany, Italy and Hungary, observations 1976-2006, Baseline, Convergence and Acceleration scenarios 2007-2060, men

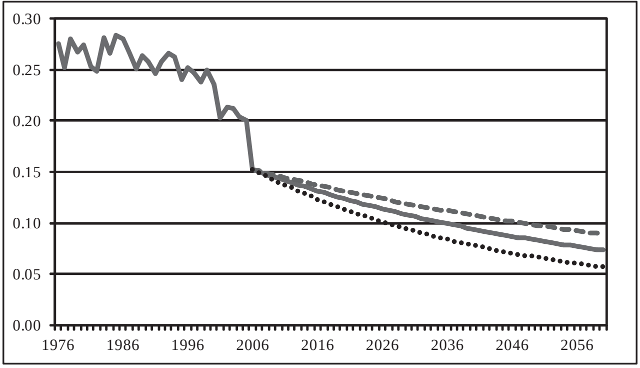
Germany



Italy



Hungary

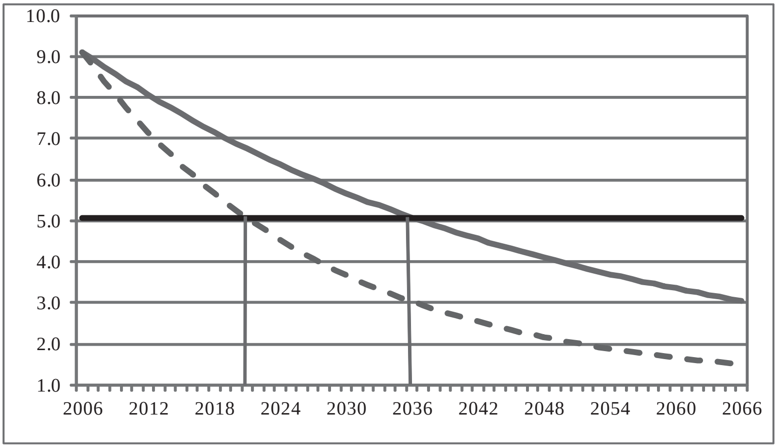


Solid line: Baseline scenario; dashed line: Convergence scenario; dotted line: Acceleration scenario been falling as well due to a decline in smoking. As the causes of changes in death probabilities have changed over time there is no a priori reason why the decline of mortality in the future should be the same as in the past. Olshansky *et al*. (2009) assume that in the next 50 years the risk of death may be influenced by accelerated advances in biomedical technology, by changes in behavioural risk factors and by aggressive management of symptoms. Therefore we developed a third scenario assuming that the future rate of decline in mortality will be stronger than during the observation period. In the Acceleration scenario we assume that the time needed for reaching a 50 percent reduction in the difference between the current age-specific probabilities of death of each country and the target values will be half of that according to the Convergence scenario. We calculate our Acceleration scenario by reducing the values of φ for each age. This is illustrated in *figure 6.13* which shows the projection of the risk ratios for men aged 50 according to the Convergence scenario. The estimated value of φ equals .977. Starting from a risk ratio of 9.12 in 2006, this value of φ implies that it will take 30 years (in the year 2036) to reach a 50 percent reduction in the value of the risk ratio compared with the target value of 1. In order to reach this value within 15 years (in the year 2021) the value of φ has to be reduced to .955. The latter value is used for the calculation of the Acceleration scenario. The values of φ for the Acceleration scenario are shown in table 6.3.

Figures 6.12a, 6.12b and 6.12c show that according to the Acceleration scenario the death probabilities decline at a higher rate than during the observation period. From tables 6.4 and 6.5 it can be calculated that average life expectancy according to the Acceleration scenario would be six years higher than according to the Convergence scenario for men and four years for women respectively. The tables show that for two thirds of the Northern, Western and Southern European countries the linear projection of life expectancy leads to a higher projection than the Acceleration scenario. This clearly illustrates that a linear increase in life expectancy can only be achieved by an acceleration in the decrease of age-specific death probabilities. For most Eastern European countries the linear projections of life expectancy are lower than the three scenarios. The explanation is that death probabilities at middle ages have shown an increase during the observation period, but that the three scenarios do not project an increase in death probabilities because it was assumed that $\varphi \leq 1$.

Figures 6.14a and *6.14b* compare the age pattern of the death probabilities of the Acceleration scenario for Germany, Italy and Hungary with those of the other two scenarios and with the target pattern. The figures show that for old

Figure 6.13. Values of risk ratio for Convergence and Accelaration scenarios, men aged 50, average of Northern, Western and Southern European countries



Solid line: Convergence scenario. Dashed line: Acceleration scenario.

ages the Acceleration scenario differs considerably from the target pattern. The reason is that the values of φ for ages 80 and over (shown in table 6.3) are closer to 1 than the values for middle ages. Olshansky *et al*. (2009) specify one scenario in which they assume that the slope of the mortality age schedule will be reduced. Using TOPALS and the partial adjustment model we can specify such a scenario by assuming lower values of φ for older ages. Note that the target pattern implies a decreasing slope for ages 80 and over. By reducing the value of φ the projections move more strongly in that direction. For example one could assume that the values of φ for ages 80 and over are equal to those for ages 50 to 70. Such a scenario would lead to an additional increase in life expectancy of three to four years compared with the Acceleration scenario.

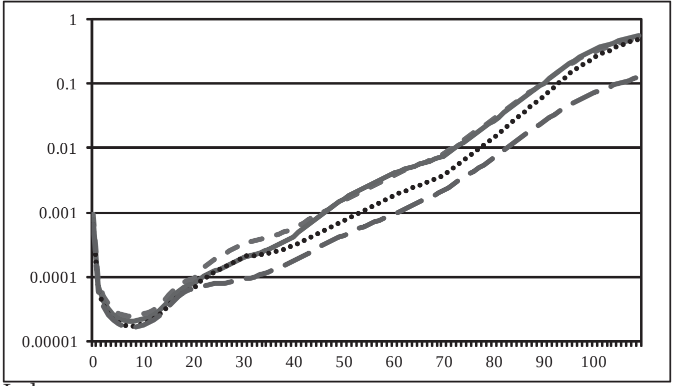
6.7. Conclusion and discussion

TOPALS is a relational model that can be used to smooth and project age-specific probabilities of death. The benefits of TOPALS are that the method is easy and flexible, while its performance is comparable with that of more complex methods. TOPALS uses a linear spline to model the ratios between the age-specific probabilities of death of a given country and a smooth standard age schedule. This implies that the relationship of the age-specific death probabilities of that country and the standard age schedule can be described by risk ratios at selected ages only, the so-called knots. The use of a spline makes TOPALS flexible: it can describe different types of age schedules. This chapter uses TOPALS to smooth age-specific probabilities of death for 26 European countries. If the standard age schedule is the average over a number of countries the risk ratios simply indicate to what extent the death probabilities of a country at different ages are higher or lower than the average. Using the average of 15 Northern, Western and Southern European countries as standard schedule, TOPALS turns out to produce smooth age curves for the 26 European countries. On average the goodness of fit of TOPALS is better than that of the Heligman-Pollard model and the Brass relational model.

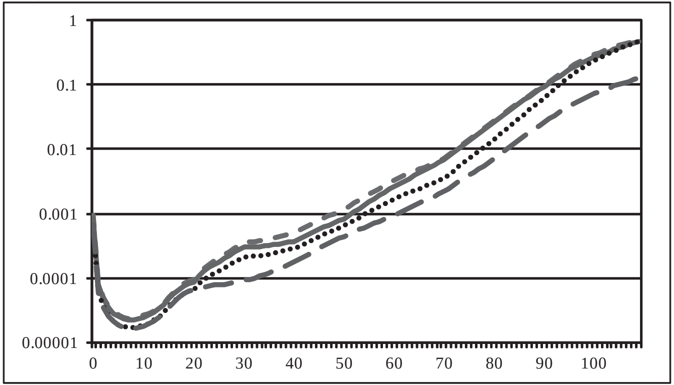
If the standard age schedule describes the best practice level of mortality, the risk ratios show how much higher death probabilities at different ages are than the best practice level. A partial adjustment model can be used to project how rapidly the death probabilities at the knots will move towards the best practice level. Oeppen and Vaupel (2002) argue that best practice life expectancy at birth has followed a linear increase for a century and a half and that a reasonable scenario is that this linear trend will continue for decades to come. Since the early 1980s life expectancy at birth of Japanese women has been the highest in the world. Thus a linear projection of life expectancy of Japanese women five decades ahead may be assumed to indicate the minimum future levels of age-specific death probabilities in the next 50 years. Using these levels as standard age schedule, TOPALS and a partial adjustment model can be used to project to what extent death probabilities in European countries will move in the direction of those levels. Instead of a priori assuming that the record level will be reached by other countries before a given forecast horizon, we estimate the parameter φ of the partial adjustment model which determines how rapidly death probabilities move towards this pattern. The value of φ can be estimated for each country separately. This produces a Baseline scenario. Table 6.4 shows that the Baseline scenario projects that life expectancy at birth in 2060 for men in Northern, Western and Southern European countries will range from 83 to 88 years and table 6.5 shows that the range for women will be from 87 to 92 years. For Central and Eastern European countries the range is wider: from 63 to 82 years for men and from 76 to 87 years for women. This is due to the fact that the development of death probabilities in countries such as Belarus, Ukraine and Russia has been much worse than in countries such

Figure 6.14a. Age-specific death probabilities in 2060, Germany, Italy and Hungary: Baseline, Convergence and Acceleration scenarios, men

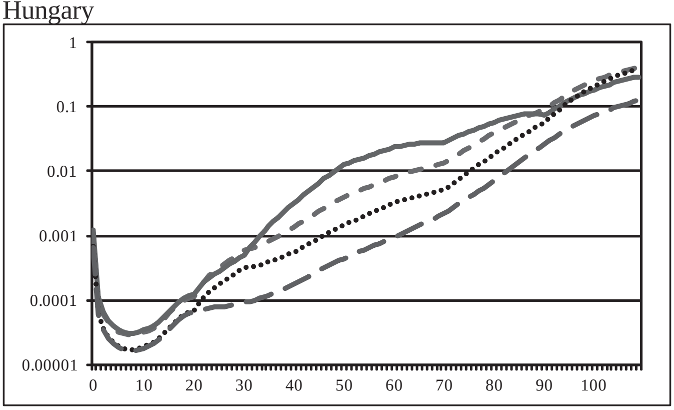
Germany



Italy



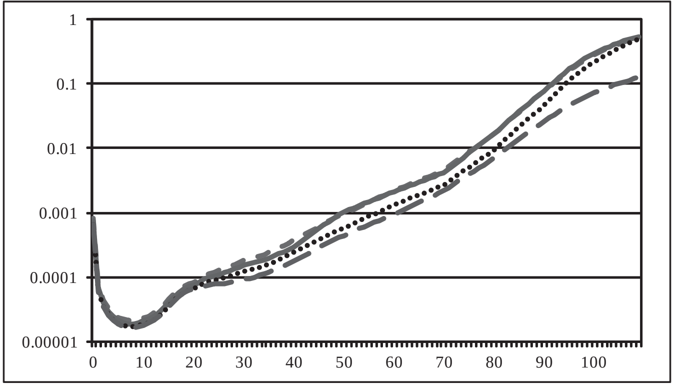
Hungary



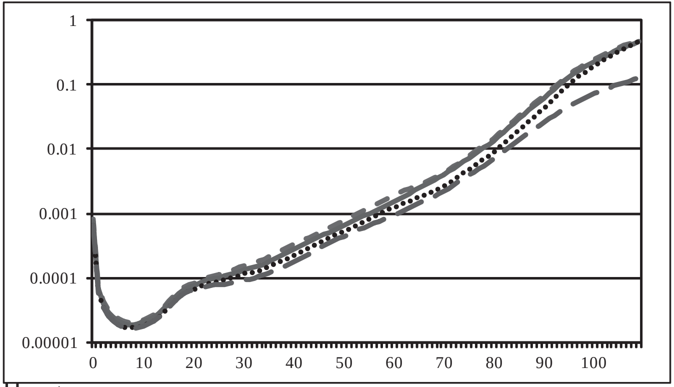
Solid lines: Baseline scenario; dashed line: Convergence scenario; dotted line: Acceleration scenario; long-dashed line: target pattern.

Figure 6.14b. Age-specific death probabilities in 2060, Germany, Italy and Hungary: Baseline, Convergence and Acceleration scenarios, women

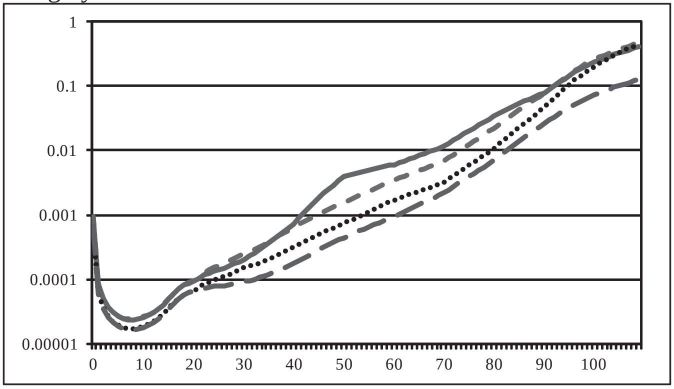
Germany



Italy



Hungary



Solid lines: Baseline scenario; dashed line: Convergence scenario; dotted line: Acceleration scenario; long-dashed line: target pattern.

as Czech Republic, Slovakia and Poland. On average the Baseline scenario is slightly higher than the Eurostat scenario for Northern, Western and Southern European countries. For Central and Eastern European countries the Eurostat scenarios are much higher as they assume strong convergence towards the low levels in Northern, Western and Southern Europe.

TOPALS can be used to calculate alternative scenarios as well. One scenario is to assume that different European countries will follow similar trends. There are two reasons for making such a scenario. One reason is that there is empirical evidence that mortality trends in developed countries have followed a converging trend. Another reason is that estimation of a common trend in mortality decline across a number of countries may produce a more stable trend than separate estimates of the trend for individual countries which are more sensitive to temporary deviations from the long-run trend. In this chapter TOPALS is used to calculate a Convergence scenario which uses estimates of the values of φ for average death probabilities across 15 Northern, Western and Southern European countries. The Convergence scenario projects a narrow range for Northern, Western and Southern European countries in 2060: from 85 to 87 years for men and from 89 to 91 years for women. For Central and Eastern European countries the Convergence scenario projects a rather narrow range as well: from 76 to 84 years for men and from 85 to 89 years for women.

During the last decades the decline in death probabilities at older ages has been moderate in many countries. This implies that even if very low target values are assumed, the projections will move only very slowly to these low values, and thus within the foreseeable future not very low levels will be reached. An alternative assumption is to assume that in the future the death probabilities will move more quickly to the target values than they have done during the last decades. The Acceleration scenario assumes that the time needed to reduce the difference between the current level of age-specific death probabilities and the target level by 50 percent will be half that according to the Convergence scenario. According to the Acceleration scenario life expectancy of men in Northern, Western and Southern European countries would range from 90 to 91 years and for women from 93 to 94 years in 2060. For Central and Eastern Europe life expectancy would range from 86 to 89 years for men and from 91 to 93 years for women. The gender gap would become about three years. The Acceleration scenario is closer to a linear projection of life expectancy than the Baseline scenario. Thus assuming a linear increase in life expectancy at birth is a rather optimistic scenario as it assumes an acceleration in the decrease of age-specific death probabilities.

When making projections of age-specific death probabilities one important decision to be made is the choice of the base period (Janssen and Kunst, 2007; Alders and De Beer, 2006). Forecasters tend to follow the general rule that for making long-run forecasts, one should use a long base period, *i.e.* a period that is at least as long as the period for which projections are made (Janssen and Kunst, 2007). However, this simple rule of thumb does not always lead to satisfactory projections. Lee and Miller (2001) suggest to fit the Lee-Carter model to the period since 1950 in order to avoid departures of the time series of the time-dependent parameter from linearity. In many Western European countries developments in mortality of men were not very favourable in the 1950 and 1960s. As a consequence projections based on time series of the last 50 or 60 years seem to be rather pessimistic. In most European countries the decline in mortality of men in the last ten years has been stronger than in previous decades. Thus if the projections would be based on the last ten years of the observation period projections of life expectancy of men would have been higher. In contrast, in many Northern, Western and Southern European countries, the increase in life expectancy of women in the last ten years have been smaller than before. Thus using a short base period would result in lower projections of life expectancy of women. However, one may question whether a base period of ten years is a sound basis for making projections for several decades into the future since developments over such a short period may be caused by temporal deviations from the long-term trend (Janssen and Kunst, 2007). Booth, Maindonald and Smith (2002) proposed a method for determining the optimal fitting period of the Lee-Carter model. Their criterion is whether the recent trend is linear. This seems to produce reasonably accurate forecasts for the relatively short run. Booth, Tickle and Smith (2005) examine this procedure for 15-years ahead forecasts for different countries. They find that their procedure improves average forecast accuracy in a number of cases, but not in all cases. Moreover, accuracy of short-term projections does not necessarily imply that long-term projections will be accurate. One way of examining the effect of the choice of the base period on forecast accuracy is to calculate the size of ex ante forecast errors, *i.e.* to examine the accuracy of projections of observations outside the base period. However, this procedure is not very helpful for examining the accuracy of long-run projections, as this would imply that one would need to examine whether very old data would help in projecting recent observations. It is questionable to what extent this would provide useful information for new projections.

One benefit of the Lee-Carter model is the assessment of forecast intervals. The use of the random walk model with drift to project the time-dependent parameter of the Lee-Carter model allows to calculate forecast intervals (Lee and Carter, 1992). However, one should note that the estimation of the forecast intervals depends on the choices to be made when estimating the model. For example, the choice of the estimation period does not only affect the point projections of the Lee-Carter model but the estimate of the forecast intervals as well. The uncertainty of the projections based on TOPALS together with the partial adjustment model could be assessed by Monte Carlo simulation assuming some distribution of the target values and of the values of φ. Expert opinion can be used to formulate assumptions about the probability distribution of the target values (Lutz *et al.*, 1998; Alders and De Beer, 2006).

The method described in this chapter projects period and age effects of changes in death probabilities and do not take into account cohort effects. Booth (2006) and Janssen and Kunst (2007) note that only few forecasts of mortality are based on cohort models. Cohort effects can lead to non-linear developments (Renshaw and Haberman, 2006). For example, changes in smoking behaviour have caused non-linear effects. It caused an increase in death by lung cancer between 1950 and 1990 among cohorts who started to smoke in the first half of the 20th century (Peto *et al.*, 2005). After the prevalence of smoking declined, death by lung cancer has started to decline. Bongaarts (2006) and Janssen and Kunst (2007) suggest that forecasts of mortality can be improved by estimating which part of mortality changes can be explained by changes in smoking behaviour. Because of the long time lag between smoking and death by lung cancer, recent statistics on smoking behaviour can be used to project smoking-related mortality for the next decades. The part of mortality that is not affected by smoking can be projected using a linear projection model. TOPALS could be used for this purpose by estimating the partial adjustment model for time series of risk ratios that are ‘corrected’ for the effect of smoking.

This chapter describes a method for making scenarios of future mortality on the basis of an analysis of past time series of death probabilities and life expectancy. One alternative is to look for determinants of changes in mortality. For example, the increase in life expectancy can be explained by changes in life style behaviour (diet, smoking, physical exercise), the availability of medical and long-term care, the improvement of medical technology, prevention, and living conditions. However, it is very difficult to assess the individual effects of these determinants as they have changed simultaneously. Moreover, it is difficult to make projections of these underlying causes (Booth, 2006). This would imply that one would need to make forecasts of medical technical progress and its effects on mortality, forecasts of the availability of medical care, which is dependent on both economic developments and political choices, and forecasts of future changes in behaviour. Nevertheless, one may develop scenarios based on alternative assumptions about the future developments in these determinants and their effect on mortality. This can result in alternative assumptions about the target pattern of age-specific death probabilities and TOPALS together with the partial adjustment model can be used to make projections of age-specific death probabilities towards these target levels. The value of the coefficient φ of the partial adjustment model can be estimated on the basis of the times series of risk ratios. However, that would not lead to widely different scenarios as low target levels will lead to relatively high estimated values of φ. Alternatively one can determine the value of φ on the basis of an assumption about the number of years it will take until the difference with the target pattern will be reduced by 50 percent. For example, if φ equals 0.933 it will take ten years to halve the distance to the target value. If one would assume that it would take 20 years to reduce the difference with the target value by 50 percent one should assumes that the value of φ equals 0.966.

7. Conclusions and discussion

Population forecasts project the future population by age and sex. Usually the cohort component model is used for making the calculations. Starting from the current population numbers by age and sex the cohort component model projects how the population will change as a consequence of changes in the levels of fertility, mortality, and migration. Thus for making population projections assumptions need to be made about future changes in fertility, mortality, and migration. These assumptions can be based on quantitative models, e.g. time series models or explanatory models. On the basis of an assessment of trends observed in the past, time series models can be used to make projections showing what will happen if past trends will continue. Even though such model based projections may seem objective, the application of models for making projections is based on a number of choices that are not always made explicit by the forecaster. Alternatively forecasts can be argument based, *i.e.* they can be based on expert opinions about likely future developments in the main drivers of changes in fertility, mortality and migration. Whether forecasts are based on models or arguments, they are based on choices and assumptions.

This book describes different models that can be used for making assumptions about future changes in fertility, mortality, and migration. The emphasis is on quantitative methods. We argue that the choices and assumptions that need to be made when using these methods should be made explicit. The methods should be regarded as tools for forecasters rather than as statistical models that ‘automatically’ produce forecasts. If in using these methods the forecaster would have made different choices the outcomes of the forecasts or scenarios would have been different. In order to make it possible for users to judge the quality of forecasts and scenarios the forecasting process should be transparent. The outline of this chapter is as follows. First, we summarize the main findings of Chapters 2 and 3 about migration, Chapters 4 and 5 about fertility and Chapter 6 about mortality. Subsequently we discuss the use of methods for making population projections, scenarios and forecasts in a transparent way.

7.1. Migration

Many countries do not have reliable or detailed statistics of international migration. Due to differences in definitions and measurement methods cross-country comparisons of international migration patterns are difficult. Moreover, under-registration, under-coverage and accuracy of the collection system affect the measurement of migration. The main source of differences in definition of migration are the place of residence and the duration of stay. The legal place of residence is not necessarily the same as the actual place of residence. For example, emigrants may be registered in their country of citizenship even after several years of living abroad. The duration of stay criterion in most European countries ranges from three months to one year. But some countries measure permanent change of residence only and some other countries do not take duration of stay into account at all. One main cause of under-registration of migration is that migrants may not report a change in place of residence. In general migrants have more incentive to report their arrival than their departure. Therefore immigration statistics are generally considered more reliable than emigration statistics. Under-coverage occurs because particular migrant groups may not be included in statistics, *e.g.* nationals, asylum seekers or students. In some countries migration statistics are based on sample surveys. These may be unreliable due to sampling errors if the sample size is small. Because of the differences in definition, registration, coverage and accuracy the numbers of emigrants reported by sending countries differ from the numbers of immigrants reported by receiving countries.

Chapter 2 shows how comparing migration statistics from sending and receiving countries can help in making internationally consistent estimates of migration flows. The idea behind the method is simple. If we compare emigration statistics of country A to countries B, C and D with the immigration statistics of countries B, C and D, and we find that the emigration statistics are x percent lower than the corresponding immigration statistics, we assume that emigration statistics of country A should be multiplied by $100/(100\text{-}x)$ in order to obtain estimates of the ‘true’ size of emigration. In fact, the calculations are a bit more complicated than this, because this calculation implies that it is assumed that the immigration statistics are correct. But immigration statistics are affected by differences in definition and measurement errors as well. Thus the immigration numbers need to be adjusted too. For that reason two tables of migration flows between countries are compared. One table is based on immigration statistics reported by receiving countries and shows immigration numbers by country of immigration and by country of origin. The other table is based on emigration statistics by country of destination reported by sending countries. One can calculate adjustment factors for immigration and emigration for each country in such a way that for each country the total number of immigrants calculated from the adjusted immigration table equals the total from the adjusted emigration table. In order to find a unique set of adjustment factors one needs to impose one restriction. For example, one may assume that immigration statistics of one particular country are reliable. This implies that one can assume that the adjustment factor for immigration for that country equals one. The calculations can easily been done in a spreadsheet program. One basic assumption underlying these calculations is that the distributions of reported immigration by country of origin and reported emigration by country of destination correspond with the distribution of actual migration flows. This assumption may not hold in all cases. For example, reported emigration from country *A* to country *B* may be *x* percent too low, but reported emigration from country *A* to country *C* may be *y* per cent too low. Estimating one adjustment factor for emigration from country *A* would result in overestimating one flow and underestimating the other. For that reason parameters are added to the model in order to take this effect into account. However, the number of additional parameters should be limited. Otherwise, the estimate of the adjustment factor for a particular country is based on a very limited number of data only. Thus if there is reason to believe that for one country the reported distribution of immigration or emigration by country of origin or destination differs strongly from the actual distribution it should be concluded that comparing the immigration and emigration tables does not provide sufficient information. Additional information will be needed. This may be obtained from expert opinions. The IMEM project is aimed to develop a Bayesian method which includes expert opinion (Raymer and Smith, 2010). The method described in chapter 2 can only be applied for countries reporting immigration and emigration numbers by country of origin and destination respectively. If data are missing, a two-step procedure can be followed (Raymer *et al*., 2011). The first step estimates missing data using covariate information. The second step harmonises the estimates using the procedure described in chapter 2.

International migration flows tend to show more short-run fluctuations than developments in fertility and mortality. One reason is that changes in migration are heavily dependent on economic and political changes, whereas fertility and mortality are dependent on gradual long-term trends, such as cultural changes and changes in living conditions and health care. Another reason is that immigration totals include different categories of migrants, such as labour migrants, family migrants, asylum seekers and returning nationals and these categories show different changes across time. The same applies to emigration. Chapter 3 shows how time-series models can be applied to extrapolate immigration, emigration, and net migration. Because of the strong fluctuations in migration, different extrapolation models may produce a wide range of projected outcomes. Deterministic time series models assume that there is a fixed trend that is not affected by random fluctuations. These models project long-term trends. If recent changes deviate from this trend, they are not projected into the future because they are assumed to be temporary only. As a result projections made in successive years react only slowly to recent changes in the time series. In contrast stochastic time series models, such as ARIMA models, are based on the assumption that the trend is subject to random changes. Recent fluctuations affect the level of the trend. The results of extrapolations depend on various choices made by the forecaster. In addition to the choice between a deterministic and a stochastic time series model the choice of the base period makes a difference. For example, a long base period may suggest that migration shows random fluctuations around a constant level, whereas a short base period suggests that there is an increasing trend, as is illustrated in chapter 3. Since there is no single extrapolation method that outperforms all other models under all circumstances and each model has its pros and cons, the logical way to improve projections is to examine the explanations behind the changes in migration.

There is not a single explanation of changes in migration. The types and mechanisms of migration have changed. In the 1960s shortages in the Western European labour market created opportunities for labour migrants from Southern countries. After the rise of the unemployment level during the economic recession in the early 1970s, most Western European countries imposed immigration restrictions. Many labour migrants returned home, but those who stayed brought their families over, which lead to an increase of family reunification. The collapse of communism in Eastern Europe resulted in an increase in immigration from Eastern to Western Europe in the 1990s. At the end of the 20th and the start of the 21st century wars and unrest in former Yugoslavia and the Middle East lead to an increase in asylum seekers.

Labour migration is primarily affected by the situation in the labour market, marriage migration is affected by the partner choice of the resident migrant population, the migration of asylum seekers is affected by political turmoil in sending countries and asylum policies in receiving countries, and return migration of nationals is affected by the size of emigration of nationals in previous years. Argument based forecasts should take these driving forces into account. For projecting the future number of labour migrants the main question is whether the decline in the working age population will lead to shortages in the labour market. An increase of labour force participation rates may lead to an increase in labour supply, whereas an increase of labour productivity, of imports and of investments in other countries may reduce labour demand. On the other hand, ageing of the working age population may lead to a decrease in labour supply, whereas population ageing may lead to an increase in the demand of health care and of long term care, and this may cause additional labour demand, as these sectors tend to be labour intensive.

A forecast of marriage migration can be based on assumptions about the choice of partners by resident migrants. This may differ strongly between origins of migrants. Some migrant groups tend to marry a partner from the country of origin, whereas others choose a partner in the country of residence.

Thus a projection should take these differences into account. For making assumptions about the future number of asylum seekers one can make separate assumptions about the total inflow to the European Union and the distribution between EU countries. The former depends primarily on the situation in sending countries, whereas the latter depends on differences in the strictness of policies across receiving countries. The analysis in chapter 3 shows that a larger part of changes in the number of asylum seekers in individual European countries were due to changes in the distribution of asylum seekers over European countries than to changes in the total inflow to the EU. Thus projections of the future number of asylum seekers should be based on assumptions about future co-ordination of migration policies between EU countries.

Projections of the future size of emigration depend on assumptions about return migration of immigrants and about the propensity to emigrate of nationals. Return migration varies strongly by type of immigrant. The return migration rate of labour migrants tends to be considerably higher than that of family migrants. Nationals may emigrate for different reasons. Students and labour migrants may be expected to leave the country of origin temporarily. People who emigrate because they are not satisfied with the situation in their home country and retired people who emigrate to Southern Europe because of the warmer climate may be expected to stay in the destination country for a longer period. In short, many factors affect different types of immigration and emigration and thus argument based forecasts of future migration depend on many underlying assumptions. Moreover, the interdependency of forecasts of immigration and emigration should be taken into account. Emigration of foreigners depends on the size of immigration flows in previous years, whereas immigration of nationals depends on emigration in previous years.

Thus if immigration increases, one may expect an increase in emigration of migrants some years later, whereas if emigration of nationals increases, one may expect an increase in immigration of nationals some years later.

7.2. Fertility

Forecasts of fertility can be based on expectations, explanations, or extrapolations. In various surveys young women are asked how many children they expect to have during their lifetime. As expectations are not always realised, the results of these surveys cannot be used at face value for making forecasts (De Beer, 1991, 2000). To some extent the deviations between expectations and actual behaviour are systematic. For example, one reason for not having the intended number of children is the break-up of a relationship. Another reason is infecundity. As a result expectations of future fertility of young cohorts tend to be higher than actual fertility. To the extent that the differences between intentions and realisations are systematic, a model may be used to adjust the expectations (De Beer, 1991). However, the realisation of expectations does not only depend on individual factors but also on changes in the social, economic and political environment. If respondents are not better capable of projecting these changes than population forecasters, the use of expectations data will not improve forecast accuracy (De Beer, 2000).

Differences in levels of fertility can be used for forecasting by means of distinguishing population categories with different levels of fertility. For example, one may distinguish fertility by the level of educational attainment. One can use this difference for making forecasts of future fertility in either of two ways. One may either assume that the proportion of people with a high level of educational attainment will increase or one may assume that the differences in the level of fertility by level of educational attainment will diminish, *e.g.* because the fertility level of people with a lower level of education will move towards that of people with a high education. Another example of using fertility differences for making forecasts is to examine regional or international differences. Chapter 4 examines regional differences in the level of fertility and chapter 5 international differences. If one assumes that the fertility level of one region or country will move into the direction of the current level of fertility of another region or country, this can be used for making forecasts for the former region or country.

Chapter 4 examines regional differences in the level of fertility in the Netherlands. The focus is on differences between small and large municipalities. The level of fertility in small cities exceeds that in large cities. An explanatory model is specified in order to explain these differences. Four categories of variables are used: demographic, socioeconomic, cultural and regional variables. The demographic variables include the household structure and the ethnic structure of the population. The socioeconomic variables include the proportion of newly built houses as a percentage of the stock of houses, the percentage of the population with low income and the percentage of the population receiving social benefits. The cultural variables include religion and the degree of urbanization. The regional variables are included because not all systematic regional patterns can be accounted for by the other explanatory variables. The explanatory model can be used for argument based forecasting. First, one should make assumptions about whether or not the differences in the explanatory variables will persist or will diminish in the future. Second, the model can be used to assess the consequences of these assumptions for future differences in fertility levels. For example, the demographic variables show two opposite effects. In large municipalities both the percentage of young Moroccan and Turkish women and the percentage of young women living alone are relatively high. The former has an upward on the level of fertility and the latter has a downward effect. If the differences between small and large cities in the ethnic and household structure would become smaller, this would not have a strong effect on differences in the average level of fertility, because both variables have opposite effects. However, one may argue that demographic differences between small and large cities will not become smaller. Selective migration may cause differences in the population structure to be persistent. Moreover, if the level of fertility of ethnic groups will decline into the direction of the level of nationals, this will have a downward effect on the level of fertility in big cities, and as result the difference in the level of fertility between large and small cities may increase rather than decrease. Chapter 4 shows how assumptions can be made about the future effects of the other explanatory variables as well. In this way the model can be used as an instrument for argument based forecasting. Statistics Netherlands and the Netherlands Environmental Assessment Agency use this model for making assumptions about the future level of fertility for the official regional population forecasts for the Netherlands (De Jong *et al*., 2005).

Chapter 5 shows how international comparisons of fertility can be used for making projections and scenarios of future fertility. The most widely used indicator of fertility is the Total Fertility Rate (TFR). The level of the TFR is determined not only by changes in the average number of children per woman across successive cohorts, but by changes in the timing of fertility as well. Since the effects of changes in the timing of fertility are temporary, we cannot simply extrapolate recent changes of the TFR into the future. For that reason it is useful to make assumptions about the future values of the age-specific fertility rates rather than about the level of the TFR. Separate projections of individual age-specific fertility rates will lead to irregular patterns. Chapter 5 shows how the relational model TOPALS (Tool for projecting age-specific rates using linear splines) can be used to project a smooth age schedule. TOPALS describes the ratios of the age-specific fertility rates to be projected and those of a smooth standard age schedule by a linear spline. One benefit is that one does not need a complex model to describe the age pattern. One only needs to describe the differences compared with that standard age schedule. If one uses the average fertility rates over a number of countries as the standard age schedule, the rate ratios indicate to what extent age-specific fertility rates of a country deviate from the average. The benefit of using a linear spline is that one needs to specify the values of the rate ratios at selected ages, the so-called knots, only.

TOPALS makes it possible to create scenarios in which the shape of the age schedule changes. This allows the forecaster to make a distinction between a rise in the mean age at childbearing due to a decrease in fertility rates at very young ages and a rise at older ages caused by the catching up of postponed births. If one assumes that the differences of the fertility rates in a country with the European average will become smaller, TOPALS can be used for making a convergence scenario. Alternatively the current fertility pattern of a ‘forerunner’ country (after smoothing) or an assumption about the future fertility age schedule of a young cohort can be used as standard age schedule. Chapter 5 shows how a partial adjustment model can be estimated to determine how quickly the age-specific fertility rates will move in the direction of the current fertility pattern of Sweden. Sweden is generally considered to be a forerunner country. Since both the Eurostat and the Swedish national population projections assume that age-specific fertility rates in Sweden will hardly change in the future, the current age pattern of fertility reflects the future fertility pattern of young cohorts and can be regarded as a ‘target’ for other European countries. One benefit of using a partial adjustment model is that the forecaster does not need to specify a priori in which year other countries will reach the current Swedish pattern. In contrast the most recent Eurostat scenario assumes that convergence will be reached in the year 2150. The projections calculated by TOPALS exceed those of Eurostat. In addition, chapter 5 shows how TOPALS can be used for making a scenario of future fertility on the basis of assumptions about changes in the age pattern of fertility in Nordic, Western, Central, Southern and Eastern European countries compared with the current average European pattern. This scenario can produce different age patterns across countries.

7.3. Mortality

Changes in mortality rates can be explained by improvements in public health, advances in medical treatment, changes in bio-medical technology, availability and quality of long term care, improvements in the standard of living, introduction of safety measures, effectiveness of preventive screening, changes in socioeconomic inequality and changes in health-related behaviour such as smoking, alcohol use, diet and physical exercise. If these explanations are to be used for making forecasts, the magnitude of the effects of these developments on the level of mortality needs to be assessed and assumptions be made about future developments in the main driving forces of mortality changes. Even though many factors have an influence on changes in mortality, the development of life expectancy at birth has shown a gradual development over time. Oeppen and Vaupel (2002) show that 'best practice' life expectancy at birth has shown a linear trend for more than a century and a half. Thus rather than assessing the separate effects of all underlying forces, one may choose to forecast future mortality by extrapolating the linear trend of life expectancy into the future. Oeppen and Vaupel calculate that life expectancy has increased by 2.5 years per decade. They argue that there is no reason why this linear trend will not continue in the coming decades. However, the underlying age-specific death probabilities have changed in different directions. If changes in age-specific death probabilities are projected into the future life expectancy will increase slower than a linear projection. This raises the question whether projections of future mortality should be based on an extrapolation of life expectancy or of age-specific death probabilities.

Chapter 6 shows how both approaches can be combined by using TOPALS. Since 1981 life expectancy of Japanese women is the highest in the world. A linear projection of life expectancy of Japanese women results in a level of life expectancy of almost 100 years in 2060. This corresponds with a 74 percent reduction of age-specific death probabilities of Japanese women compared with the 2008 levels. These values can be regarded as the target level of mortality for other countries. Chapter 6 describes how a partial adjustment model can be estimated in order to assess with what speed age-specific death probabilities of 26 European countries will move towards the target values. Note that the use of this model does not imply that it is assumed that the target level will be reached within the forecast period, but rather that the death probabilities will move in that direction. TOPALS uses a linear spline to describe the ratio between the death probabilities of each country and the target level. The use of a linear spline implies that the partial adjustment model needs to be estimated for selected ages (the knots) only. The partial adjustment model can be estimated separately for each country. For Northern, Western and Southern European countries this results in a projected life expectancy at birth in 2060 ranging from 83 to 88 years for men and ranging from 87 to 92 years for women. Thus for all countries life expectancy of women in 2060 would be higher than the current Japanese level of 86 years.

One alternative approach is to estimate the partial adjustment model for the average death probabilities across a number of countries. In chapter 6 the average death probabilities over 15 Northern, Western and Southern European countries are calculated. One benefit of using average death probabilities rather than the separate probabilities for each individual country is that the average trends may be more stable in the long run. These estimates produce converging projections. According to this Convergence scenario life expectancy in 2060 would range from 85 to 87 years for men and from 89 to 91 years for women.

Another scenario can be based on the assumption that in the future the decrease in mortality may be stronger than in the past, *e.g.* due to medical progress. An Acceleration scenario is calculated under the assumption that the number of years needed to reach a reduction of the difference between the current and target pattern by 50 percent is halved. This would result in a projected life expectancy of men from 90 to 91 years and for women from 93 to 94 years.

In all scenarios the projected life expectancy for Central and Eastern European countries will be lower than in the Northern, Western and Southern European countries, but the differences in the Convergence and Acceleration scenarios will be much smaller than in the Baseline scenario which is based on the projection of a continuation of past trends in each country. Obviously other scenarios can be specified as well. For example, one may assume that the future decline in death probabilities at older ages will surpass that at younger ages. Another scenario could be to assume that the differences in mortality between East and West Europe would become smaller in the long run. The aim of describing these scenarios is to illustrate how TOPALS can easily be used to make alternative scenarios rather than to present the most likely scenario.

7.4. Transparency of population projections, scenarios and forecasts

Calculations of the future size and age structure of the population are based on assumptions about future changes in the levels of fertility, mortality, and migration. Depending on the type of assumptions the outcomes of these calculations can be considered as projections, scenarios or forecasts. Projections are aimed to describe what will happen in the future if current trends will continue. Time-series models seem the most appropriate instrument to calculate projections. They identify past trends and show the effects of a continuation of these trends in the future. Scenarios describe alternative futures that may occur assuming different future developments in the driving forces of fertility, mortality and migration. Explanatory models can be used to assess to what extent future fertility, mortality and migration may vary depending on alternative assumptions about future social, economic, cultural, political or technological developments.

Forecasts are aimed to describe the most likely future. The difference with projections and scenarios is not the method that is used but the interpretation of the underlying assumptions. If the forecaster assumes that a continuation of trends represents the most likely future, then the projection of these trends can be interpreted as a forecast. If the forecaster shows how extrapolations based on different assumptions lead to different outcomes, these projections can be interpreted as alternative scenarios. If the forecaster makes different scenarios, *e.g.* based on alternative assumptions about future developments in driving forces, and considers one of these scenarios as most likely, the latter scenario can be considered as a forecast, whereas the other scenarios show possible alternative developments. Thus a forecast does not follow automatically from the application of a method. The distinction between projections, scenarios and forecasts cannot be made solely on the basis of the methods that are applied. A projection or a scenario can be regarded as a forecast if the forecaster assumes that this will be a likely future. However, as Keyfitz (1972) notes, statistical agencies usually label the outcomes of their calculations as projections, whereas users interpret them as forecasts. Keilman (2008) argues that, unless the agency presents its assumptions as unrealistic, the projections published by statistical agencies can be regarded as forecasts indicating a likely development, given the current knowledge of the forecaster. Eurostat uses the term scenarios for their projections. In 2008 Eurostat published a ‘Convergence scenario’ and a ‘No migration scenario’. Since the ‘No migration’ scenario is not considered as a realistic scenario, the convergence scenario is used as a forecast by other European agencies. Eurostat argues that a converging tendency is in line with past trends (Lanzieri, 2009). This suggests that this scenario should be considered as a forecast of a likely development rather than one scenario of a possible future.

Consequently the use of the labels projections and scenarios is not sufficient to distinguish them from forecasts. Both projections and scenarios are based on choices and assumptions. Even the assumption that past trends will continue in the future does not automatically lead to one projection. The choice of the time series model, the choice of the base period and the choice of the indicator to be projected can make a lot of difference. A deterministic model, *e.g.* a linear time trend, assumes that there is a fixed trend that is not affected by random fluctuations, whereas a stochastic model, such as an ARIMA model, assumes that random fluctuations affect the level of the trend. Deterministic models emphasise long-run developments. Projections based on this model tend to react slowly to recent changes in the time series. In contrast, projections based on a stochastic model tend to react very quickly. A long base period may result in quite different projections than a short period. Chapter 3 shows that a long base period may suggest that there is no increasing trend in migration, whereas a short base period does. Chapter 6 shows that in projecting mortality the choice of the indicator to be projected makes a difference: a projection of life expectancy at birth results in different values than a projection of age-specific death probabilities.

If an explanatory model is used, the forecaster needs to make assumptions about the future values of the explanatory variables. Lutz (2009) developed a questionnaire including the main driving forces of future changes in fertility, mortality and migration. For example for changes in life expectancy the main forces are biomedical technology, effectiveness of health care, behavioural changes, possible new infectious diseases, environmental change, and changes in population composition. For each of these forces a set of arguments is defined that would have an influence on the future effects of these forces. For example, for the effects of health care systems on changes in life expectancy the arguments are: The costs of new treatments will be prohibitive for a large part of the population, there will be very effective new technologies, waiting times for treatment will increase, society will afford expensive new treatments, progress in preventive medicine will lead to lower death rates, and dissemination of health information will increase longevity. For each of these arguments experts are asked to weigh the validity of the argument (ranging from ‘very likely to be right’ to ‘very likely to be wrong’) and to indicate the impact of the argument (ranging from ‘a large upward influence on life expectancy’ to ‘a large downward influence on life expectancy’). Both the answers to the validity question and to the impact question are given a weight. In addition the experts are asked the relative importance of the six forces. These are used to weigh the scores in order to produce one number for each expert for life expectancy. These numbers are not directly used to project the future level of life expectancy. Lutz asks each expert what will be the likely future value of life expectancy in a given year. The scores of the experts are used to assess the relative importance of the forces. For example, Lutz (2009) describes the results of a survey among international experts in which experts expect that life expectancy will increase by two years per decade on average. The results of the survey show that experts attribute about a half of the increase in life expectancy to bio-medical progress. Lutz concludes that technological progress will lead to an increase in life expectancy by one year per decade. However, there are two problems in following this approach. First, the forces are not independent and thus their effects on life expectancy cannot simply be added up. Secondly, the assumption that life expectancy will increase by two years per decade is the average of the increase in life expectancy expected by experts and does not follow directly from the arguments. Even though this exercise is useful in assessing forces underlying future changes in mortality, fertility, and migration, the resulting forecasts are not purely argument based since the projected changes in fertility, life expectancy and net migration do not follow directly from the arguments but rather are averages of expert opinions. Thus the forecast is expert based rather than argument based.

Rather than emphasising the distinction between the terms projections, scenarios and forecasts, it is important to make the underlying choices and assumptions as well as the reasons for making the choices and assumptions explicit. The forecaster should make the methods and assumptions transparent in order to make it possible for the user to determine how to interpret the outcomes of the calculations. Armstrong (2001) describes 139 principles for forecasting. They cover the collection and preparation of data, the selection and application of methods, and the evaluation and presentation of forecasts. Armstrong (2001) argues: “When managers receive forecasts, they often cannot judge their quality. Instead of focusing on the forecasts, however, they can decide whether the forecasting process was reasonable for the situation.” This requires that it is necessary for users to know which decisions are made by the forecaster. Two principles mentioned by Armstrong are “Provide complete, simple and clear explanations of methods” and “Describe your assumptions.” In other words: the forecasting process should be transparent. One main reason given by Armstrong is that by examining forecasting processes and improving them, accuracy can be increased.

In order to achieve transparency, it is not sufficient to make choices and assumptions explicit. In addition, it is important that the forecaster gives arguments for the choices and provides information about the consequences of these choices. For example, when using an extrapolation model the forecaster should indicate which difference it would have made if another base period or another model would have been chosen. When using an explanatory model the forecaster should indicate to what extent alternative assumptions about the future developments of the explanatory variables would have resulted in different scenarios. Transparency is a necessary condition for users to be able to assess whether a projection can be regarded as a forecast of a likely future or a scenario of only one possible future and whether a scenario can be regarded as a projection that extrapolates past trends or as a forecast of likely developments.

One obvious criterion for regarding a projection or scenario as forecast is accuracy. If past projections or scenarios produced by the same method or by the same forecaster have turned out to be accurate, the user may regard the projections or scenarios as forecasts. If short-term forecasts have been published regularly, such as daily weather forecasts or quarterly economic forecasts, there is sufficient empirical evidence to assess the forecast accuracy. If projections have repeatedly been proven to be reasonably accurate, the user can regard new projections made by the same method or by the same forecaster as reliable forecasts. However, for long run forecasts there are only few forecasts of which the accuracy can be examined. Moreover the methods may have changed or trends may have changed, which make it much more difficult to assess whether forecast accuracy in the past will be relevant for the future. In those cases the user cannot simply conclude that a method that produced accurate forecasts in the past or a forecaster with a good track record in the past is likely to produce accurate forecasts in the future. The user needs information about the reasons for the choice of a particular method and the underlying assumptions in order to be able to assess the validity of new forecasts for the long run. Thus transparency of forecast and scenarios is a necessary condition for a user to be able to decide whether a projection or scenario can be considered as forecast.

The uncertainty of the validity of the choices and assumptions underlying projections, scenarios and forecasts implies that population forecasts are uncertain. De Beer (2000) gives an overview of issues related with uncertainty of population projections. The traditional way to deal with the uncertainty of population forecasts is to present deterministic variants or scenarios. This implies that alternative sets of assumptions about the future levels of fertility, mortality and migration have to be made. These assumptions can be combined into a limited set of scenarios, *e.g.* a low variant combining low values of the total fertility rate, life expectancy and net migration and a high variant based on high values of these components. The reason for combining low values of the components of change in one variant and high values in another is not that it is assumed that these values are assumed to be interdependent, but because these variants result in low and high projections of population growth. If it is assumed that there is no perfect correlation between the levels of fertility, life expectancy and net migration the range between these variants overestimates the uncertainty of future population growth.

For that reason several researchers have proposed to make stochastic or probabilistic population projections (*e.g.* Alho and Spencer, 2005 and Lutz and Goldstein, 2004). In 1998 Statistics Netherlands was the first national statistical institute that published stochastic population forecasts (Alders and de Beer, 1998 and Keilman, 2008). These projections are based on assumptions about the future probability distribution of fertility, mortality, and migration. This requires that assumptions are made about the form of the distribution and about parameters of that distribution. For example if a normal distribution is used the forecaster has to make an assumption about the future values of the variance of the total fertility rate, life expectancy at birth and net migration. These assumptions can be based on an analysis of forecast errors in the past, the variance can be estimated on the basis of a time series model, or the assumption about the future value of the variance can be based on expert judgment. If past forecast errors are analysed one problem is that the results depend on the particular period for which the errors are examined. For example, if the level of an indicator has not changed much during the last ten years, a random walk forecast made ten years ago projecting that the indicator would remain constant, would have produced accurate forecasts, and thus one could conclude that the variance is relatively small. However, in another period in which the indicator showed an increasing or decreasing trend, this projection method would have lead to poor results and thus the variance would be large. One alternative approach is to estimate the variance of the forecast errors from a stochastic time series model. This calculation is based on the assumption that the correct time series model is specified, *i.e*. it is assumed that future developments will be like the past. However, one source of uncertainty of forecasts is that future developments may not be a continuation of past trends. This would imply that past changes may not inform us on possible future changes. For that reason forecast variance can be determined on the basis of expert opinions about the probability of future events that have not yet occurred, *e.g.* medical breakthroughs leading to a strong increase in longevity. This implies that assumptions about the probability distribution of forecasts can be based on arguments just like forecasts themselves. The assessment of the probability of forecasts is a forecast itself. This does not imply that probabilistic forecasts are not useful. In contrast, rational decision making requires that a proper assessment of the probability of forecasts should be taken into account, even though the assessment of the probability is to some extent subjective (Raiffa, 1997).

The aim of this book is to show how methods can be used to make projections and scenarios in a transparent way. Chapters 2 to 5 illustrate the usefulness of using quantitative methods for making assumptions about future changes in fertility, mortality, and migration forecasts. In order for the forecasts to be transparent, the methods should be as simple as possible. Both for the forecaster and the user it should be clear what choices are made and what the consequences of these choices are. If methods are complicated, forecasts come from a black box. Forecasts are projections or scenarios that result from applying a method but if the forecaster cannot explain why the method produces that forecast, it will be difficult for the user to judge the validity of the forecast. The aim of chapters 2 to 5 is not to present one model that will outperform all other models. Neither is the aim to find one model that will produce objective forecasts, *i.e.* forecasts that do not depend on choices to be made by the forecaster. It is inevitable that the forecaster has to make choices and it is important that these choices are made on the basis of arguments and do not remain implicit. The user should know which choices are made, what the reasons for those choices are and what the impact of those choices is on the outcomes.

The first step in making forecasts is to assess the quality of data. As Keilman (2008) notes “poor data quality tends to go together with poor forecast performance”. In most countries the quality of data on international migration is considerably poorer than data on fertility and mortality. Particularly the size of emigration tends to be underestimated in most countries because of under-registration. Chapter 2 shows how migration data can be improved by using a simple model that compares data from different countries. Apart from systematic errors in statistics due to under-registration, data may be affected by random fluctuations. In order to avoid that random fluctuations are projected into the future, it is useful to smooth data before making projections. Chapters 5 and 6 show how TOPALS can be used for smoothing age-specific fertility rates and death probabilities respectively. TOPALS is a relational model that can be used to smooth all types of age-specific rates or probabilities using a smooth standard age schedule. Both chapters show that for many countries the fit of TOPALS is better than that of complicated models.

Once reliable and smooth estimates are available, they can be used as a basis for projections. Chapter 3 shows how time series models can be used to make projections. Since different time series models may lead to different projections, Chapter 3 argues that it is useful to examine explanations behind the changes in migration. Since different types of migration are affected by different driving forces, an argument-based forecast of migration should be based on a distinction of types of immigration and emigration. Chapter 4 illustrates how an explanatory model can be used for making assumptions about future changes in fertility. The model is used to assess the effects of different types of explanatory variables on regional differences in the level of fertility. The chapter shows how assumptions about future developments in the explanatory variables and their effect on the level of fertility can be used as arguments for forecasting whether or not regional differences in the level of fertility will disappear.

Chapters 5 and 6 show how TOPALS can produce time series projections as well as alternative scenarios for fertility and mortality respectively. In both cases several choices have to be made, particularly about the choice of the standard age schedule which can be used as target pattern and about the way the values of the rate or risk ratios are determined. These values can be estimated on the basis of a time series for one country or for a group of countries or assumptions about the future values of the rate or risk ratios can be made on the basis of qualitative arguments. Chapter 5 shows how TOPALS can be used to make projections of future age-specific fertility rates assuming that the fertility rates of countries in different European regions will move towards the Swedish pattern. The extent to which this target will be reached within the projection period depends on the estimated coefficient of a time series model that is fitted to past time series. Alternatively, chapter 5 shows how TOPALS can be used to make scenarios assuming that the shape of the age pattern of fertility will change. Chapter 6 shows how TOPALS can be used to make projections of age-specific death probabilities assuming that they will move in the direction of the world record level. The extent to which this level will be approached differs across countries. To project future changes a time series model can be estimated for each country separately or for a combination of countries. Alternatively TOPALS can be used to make a scenario assuming that future declines in mortality will be stronger than in the past. TOPALS is a tool rather than a statistical model.

TOPALS is a useful instrument for making transparent projections and scenarios because it is both conceptually and computationally simple. Because TOPALS is a relational model it does not include a complicated mathematical formula to describe age schedules. Rather it uses a standard age schedule. Because different standard age schedules can be used, TOPALS is flexible. As a result, TOPALS can be used to describe different age patterns of fertility and mortality. This makes it possible to use TOPALS both for cross-country comparisons and for analyses and projections of changes over time. Because TOPALS uses linear splines it does not need a complex model to describe the relationship between the age-specific rates to be projected and the standard age schedule. Rather it specifies ratios between the age-specific rates to be projected and the standard age schedule for selected ages and interpolates the value for ages in between. The use of TOPALS is transparent because it describes differences in age patterns across countries and changes over time in a rather intuitive way. It does not use parameters that may be difficult to interpret. If the standard age schedule is the average over a number of countries and the age-specific rates or probabilities for specific ages for a particular country are higher than the average, the rate or risk ratios are larger than one. If one assumes that convergence will occur, the rate or risk ratios will move towards a value of one.

References

Abel, G.J. (2009), International migration flow table estimation. PhD thesis, University of Southampton, School of Social Sciences.

Alders M. (2005), Prognose van gezinsvormende migratie van Turken en Marokkanen (Forecast of marriage migration of Turks and Moroccans). *Bevolkingstrends*, 53(2), 46-49.

Alders, M. and J. de Beer (1998), Kansverdeling van de bevolkingsprognose. (Probability distribution of population projections) *Maandstatistiek van de Bevolking*, 46, 8-11.

Alders, M. and J. de Beer (2006), An expert knowledge approach to stochastic mortality forecasting in the Netherlands. In: N. Keilman (Ed.), *Perspectives on mortality forecasting II. Probabilistic models*. Swedish Social Insurance Agency, Stockholm, 39-64.

Alho, J. and B. Spencer (2005), *Statistical demography and forecasting*. New York: Springer.

Armstrong, J. S. (2001). Standards and practices for forecasting. In: J.S. Armstrong (Ed.), *Principles of Forecasting: A Handbook for Researchers and Practitioners.* Norwell, MA: Kluwer Academic Publishers, 679-732.

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

Becker, G. (1991), *A treatise on the family*. Cambridge, Mass.: Harvard University Press.

Billari, F. and C. Wilson (2001), Convergence towards diversity? Cohort dynamics in the transition to adulthood in contemporary Western Europe. Working Paper 2001-039, Rostock: Max Planck Institute for Demographic Research.

Bilsborrow, R.E., G. Hugo, A.S. Oberai and H. Zlotnik (1997), *International migration statistics: Guidelines for improving data collection systems*. Geneva: International Labour Office.

Boleslawski, L. and E. Tabeau (2001), Comparing theoretical age patterns of mortality beyond the age of 80. In: E. Tabeau, A. van den Berg Jeths and C. Heathcote (Eds.), F*orecasting mortality in developed countries: Insights from a statistical, demographic and epidemiological perspective*. Dordrecht: Kluwer Academic Publishers, 127-155.

Bongaarts, J. (2002), The end of the fertility transition in the developed world. *Population and Development Review*, 28, 419-443.

Bongaarts, J. (2006), How long will we live? *Population and Development Review*, 32, 605-628.

Booth, H. (2006), Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting*, 22, 547-581.

Booth, H., J. Maindoland and L. Smith (2002), Applying Lee-Carter under conditions of variable mortality decline. *Population Studies*, 56, 325-336.

Booth, H., L. Tickle and L. Smith (2005), Evaluation of the variants of the Lee-Carter method of forecasting mortality: A multi-country comparison. New Zealand *Population Review*, 31,13-37.

Box G.E.P and G.M. Jenkins (1970), *Time series analysis: Forecasting and control*. San Francisco: Holden-Day.

Boyle, P. (2003), Population geography: Does geography matter in fertility research? *Progress in Human Geography*, 27(5), 615-626.

Brass, W. (1960), The graduation of fertility distribution by polynomial functions. *Population Studies*, 14, 148-162.

Brass, W. (1971), On the scale of mortality. In: W. Brass (ed.), *Biological aspects of demography.* London: Taylor & Francis, 69-110.

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.

Brass, W. (1975), Methods for estimating fertility and mortality from limited and defective data. Chapel Hill, NC: Carolina Population Center, University of North Carolina.

Brunborg, H. and A. Cappelen (2010), Forecasting migration flows to and from Norway using an economic model. Paper presented at Joint Eurostat/UNECE Work Session on Demographic Projections, 28-30 April 2010, Lisbon, Portugal

Brunetta, G. and G. Rotondi (1989), Différenciation régionale de la fécondité italienne depuis 1950. *Espace, Populations, Sociétés*, 2, 271-291.

Brunner, E. (1997), Socioeconomic determinants of health: Stress and the biology of inequality. *British Medical Journal*, 314, 1472.

Carriere, J.F. (1992), Parametric models for life tables. *Transactions of the Society of Actuaries*, 44, 77-99.

Champion, A.G. (1994), International migration and demographic change in the developed world. *Urban Studies*, 31(4/5), 653-677.

Chandola, T., D.A. Coleman and R.W. Hiorns (1999), Recent European fertility patterns: Fitting curves to 'distorted' distributions. *Population Studies*, 53(3), 317-329.

Coale, A.J. and T.J. Trussell (1974), Model fertility schedules: Variations in the age structure of childbearing in human populations. *Population Index*, 40, 185-258.

Coleman, D. (2004), Why we don't have to believe without doubting in the 'Second Demographic Transition' - some agnostic comments. *Vienna Yearbook of Population Research*, 11-24.

Courgeau, C. and B. Baccaini (1998), Multilevel analysis in the social sciences. *Population* (English selection), 10, 39-71.

Cuadrado-Roura, J.R. (2001), Regional convergence in the European Union: From hypothesis to the actual trends. *The Annals of Regional Science*, 35(3), 333-356.

Currie, I.D., M. Durbin and P.H.C. Eilers (2004), Smoothing and forecasting mortality rates. *Statistical Modeling*, 4, 279-298.

De Beer, J. (1991), From birth expectations to birth forecasts: A partial-adjustment approach. *Mathematical Population Studies*, 3, 127-144.

De Beer J. (1993), Forecast intervals of net immigration: The case of the Netherlands. *Journal of Forecasting*, 12, 585-599.

De Beer J. (1997), The effect of uncertainty of migration on national population forecasts: The case of the Netherlands. *Journal of Official Statistics*, 13 (3), 227-243.

De Beer, J. (2000), Dealing with uncertainty in population forecasting. Voorburg: Statistics Netherlands. Available at www.cbs.nl.

De Beer, J. (2006a), An assessment of the tempo effect for future fertility in the European Union. Research Note, European Commission, Directorate-General Employment, Social Affairs and Equal Opportunities.

De Beer, J. (2006b), *Future trends in life expectancies in the European Union*. Research Note, European Commission. http://www.nidi.nl/Content/NIDI/output/2006/sso-2006-02-nidi-debeer.pdf

De Beer, J., R. van der Erf and J. Raymer (2009), Estimates of OD matrix by broad group of citizenship, sex and age, 2002-2007. Report for MIMOSA-project. Available at http://mimosa.gedap.be/Documents/Mimosa_2009b.pdf.

De Beer, J., N. van der Gaag and F. Willekens (2007), A tool for projecting age patterns based on a standard age schedule and assumptions about relative risks using linear splines: TOPALS. In: Eurostat, Work session on demographic projections; Bucharest, 10-12 October 2007. Methodologies and working papers; 2007 edition, 211-235.

De Jong, A. and H. Nicolaas (2005), Prognose van emigratie op basis van een retourmigratiemodel (Forecast of emigration based on a return migration model). *Bevolkingstrends*, 53(1), 24-31.

De Jong, A., M. Alders, P. Feijten, P. Visser, I. Deerenberg, M. van Huis and D. Leering (2005), *Achtergronden en veronderstellingen bij het model PEARL*. (Backgrounds and assumption of the PEARL model) Rotterdam: NAI Uitgevers.

Del Boca, D. (2002), The effect of child care and part time opportunities on participation and fertility decisions in Italy. *Journal of Population Economics*, 15, 549-573.

Del Bono, E. (2002), Total fertility rates and female labour force participation in Great Britain and Italy: Estimation of a reduced form model using regional panel data. Paper presented at the ESPE 2002 conference, 12-15 June 2002, Bilbao.

Diniz-Filho, J., L. Bini and B. Hawkins (2003), Spatial autocorrelation and red herrings in geographical ecology. *Global Ecology and Biogeography*, 12, 53-64.

Diprete, T.A., S.P. Morgan, H. Engelhardt and H. Pacalova (2003), Do cross-national differences in the costs of children generate cross-national differences in fertility rates? *Population Research and Policy Review*, 22, 439-477.

Duchêne, J., A. Gabadinho, M. Willems and P. Wanner (2004), Study of low fertility in the regions of the European Union: Places, periods and causes. *Eurostat: Population and social conditions* 3/2004/F/no.4.

Engelhardt, H., T. Kögel and A. Prskawetz (2004), Fertility and female employment reconsidered: A macro-level time series analysis for developed countries, 1960-2000. *Population Studies*, 58(1), 109-120.

Eurostat (2010), Statistics [electronic resource]. Luxembourg: Eurostat. http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database.

Eurostat (2011), Statistics [electronic resource]. Luxembourg: Eurostat. http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database

Fahey, T. and Z. Spéder (2004), *Fertility and family issues in an enlarged Europe*. European Foundation for the Improvement of Living and Working Conditions (www.eurofound.eu.int).

Fassmann, H. (2009), European migration: Historical overview and statistical problems. In: H. Fassmann, U. Reeger and W. Sievers (Eds.), *Statistics and Reality. Concepts and Measurements of Migration in Europe* (pp. 21-44). Amsterdam: Amsterdam University Press.

Fingleton, B. (1999), Estimates of time to economic convergence: An analysis of regions of the European Union. *International Regional Science Review*, 22(1), 5-34.

Fokkema, T., H. de Valk, J. de Beer and C. van Duin (2008), The Netherlands: Childbearing within the context of a ‘Poldermodel’ society. *Demographic Research,* 19, 743-794.

Fox, J. (2000), *Nonparametric simple regression: smoothing scatterplots*. Thousand Oaks CA: Sage.

Frejka, T. (2008). Determinants of family formation and childbearing during the societal transition in Central and Eastern Europe. In: Frejka, T. et al. (eds.), *Childbearing trends and policies in Europe. Demographic research, Special collection* 7 19: 139-170.

Frejka, T. and T. Sobotka (2008), Fertility in Europe: Diverse, delayed and below replacement. In: T. Frejka *et al*. (Eds.), *Childbearing trends and policies in Europe. Demographic research, Special collection 7*, 19, 15-46.

Frejka, T., T. Sobotka, J.M. Hoem and L. Toulemon (2008), Summary and general conclusions: Childbearing trends and policies in Europe. In: T. Frejka *et al*. (Eds.), *Childbearing trends and policies in Europe. Demographic research, Special collection 7*, 19, 5-14.

Garssen, J. (2006), Will life expectancy continue to increase or level off? Weighing the arguments of optimists and pessimists. Voorburg: Statistics Netherlands.

Gauthier, A.H. and J. Hatzius (1997), Family benefits and fertility: An econometric analysis. *Population Studies 51*; 295-306.

Gayawan, E., S.B. Adebayo, R.A. Ipinyomi and B.A. Oyejola (2010), Modeling fertility curves in Africa. *Demographic Research*, 22, 211-236.

George, M.V., S.K. Smith, D.A. Swanson and J. Tayman (2004), Population projections. In: J.S. Siegel and D.A. Swanson (Eds.) *The methods and materials of demography*. Second edition. San Diego: Elsevier Academic Press, 561-601.

Giannokouris, K. (2008), Ageing characterises the demographic perspectives of the European societies. Eurostat: Statistics in focus.

Gilje, E. (1972), Fitting curves to age-specific fertility rates: Some examples. *Statistical Review of the National Central Bureau of Statistics of Sweden*, 7, 118-134.

Gilks, W.R. (1986), The relationship between birth history and current fertility in developing countries. *Population Studies*, 40, 437-455.

Goldstein, J.R. (2006), How late can first births be postponed? Some illustrative population-level calculations. *Vienna Yearbook for Population Research*: 153-165.

Goldstein, J., W. Lutz and M.R. Testa (2003), The emergence of sub-replacement family size ideals in Europe. *Population Research and Policy Review*, 22, 479-496.

Goldstein, J., T. Sobotka and A. Jasilioniene (2009), The end of 'lowest-low' fertility? *Population and Development Review*, 35, 663-699.

Hank, K. (2001), Regional fertility differences in Western Germany: An overview of the literature and recent descriptive findings. *International Journal of Population Geography*, 7, 243-257.

Hank, K. (2002), Regional social contexts and individual fertility decisions: A multilevel analysis of first and second births in Western Germany. *European Journal of Population*, 18, 281-299.

Heligman, L. and J.H. Pollard (1980), The age pattern of mortality. *Journal of the Institute of Actuaries*, 107, 49-80

Herm, A. (2006a), Recommendations on international migration statistics and development of data collection at an international level. In: M. Poulain (2006), *THESIM: Towards Harmonised European Statistics on International Migration* (pp. 77-106). Louvain-la-Neuve: Presses universitaires de Louvain.

Herm, A. (2006b), Country report Sweden, In: M. Poulain (2006), *THESIM: Towards harmonised European Statistics on international migration* (pp. 633-643). Louvain-la-Neuve: Presses universitaires de Louvain.

Hoem, B. (2000), Entry into motherhood in Sweden. The influence of economic factors on the rise and fall in fertility, 1986-1997. *Demographic Research*, 2(4).

Hoem, J.M., D. Madsen, J.L. Nielsen, E. Ohlsen, H.O. Hansen and B. Rennermalm (1981), Experiments in modelling recent Danish fertility curves. *Demography,* 18, 231-244.

Hofstede, G. (1981), *Culture's consequences: International differences in work-related values*. Beverly Hills: Sage Publications.

Howe, N. and R. Jackson (2005), Projecting immigration. A survey of the current state of practice and theory. A report of the CSIS global aging initiative. Washington, D.C.: Center for Strategic and International Studies.

Human Fertility Database (2010), The Human Fertility Database [electronic resource]. http://www.humanfertility.org/cgi-bin/main.php.

Human Mortality Database (2010), data were downloaded in August 2010, http://www.mortality.org/.

Janssen, F. and A. Kunst (2007), The choice among past trends as a basis for the prediction of future trends in old-age mortality. *Population Studies*, 61, 315-326.

Janssen, F., J.P. Mackenbach and A.E. Kunst (2004), Trends in old-age mortality in seven European countries, 1950-1999. *Journal of Clinical Epidemiology*, 57, 203-216.

Jennissen, R. (2004), *Macro-economic determinants of international migration in Europe*. Amsterdam: Dutch University Press.

Johansson, M. (2000), Population and regional development − the case of Sweden. Paper presented at the 40th ERSA Congress, Barcelona, 29 August-1 September, 2000.

Johansson, M. and D. Rauhut (Eds.) (2005), *The spatial effects of demographic trends and Migration*. Available at www.espon.lu.

Keilman, N. (2008), European demographic forecasts have not become more accurate over the past 25 years. *Population and Development Review*, 34, 137-153.

Keilman, N. and D.Q. Pham (2000), Predictive intervals for age-specific fertility. *European Journal of Population*, 16, 41-65.

Kelly, J. (1987), Improving the comparability of international migration statistics: Contribution by the conference of European Statisticians from 1971 to date. *International Migration Review*, 21(4), 1017-1037.

Keyfitz, N. (1972), On future population. J*ournal of the American Statistical Association*, 67, 347-363.

Kohler, H.P., F.C. Billari and J.A. Ortega (2002), The emergence of lowest-low fertility in Europe during the 1990s. *Population and Development Review*, 38, 641-680.

Kostaki, A. (1992), A nine-parameter version of the Heligman-Pollard formula. *Mathematical Population Studies*, 3(4), 277-288.

Kostaki, A., J.M. Moguerza, A. Olivares and S. Psarakis (2009), Graduating the age-specific fertility pattern using Support Vector Machines. *Demographic Research*, 20, 599-622.

Kraier, A., M. Jandl and M. Hofmann, (2006), The Evolution of EU migration policy and implications for data collection. In: M. Poulain (2006), *THESIM: Towards Harmonised European Statistics on International Migration* (pp. 35-75). Louvain-la-Neuve: Presses universitaires de Louvain.

Kraly, E.P. and K.S. Gnanasekaran (1987), Efforts to improve international migration statistics: A historical perspective. *International Migration Review (Special issue: Measuring international migration: Theory and practice)*, 21(4), 967-995.

Kravdal, O. (2002), The impact of individual and aggregate unemployment on fertility in Norway. *Demographic Research*, 6(10), 261-293.

Kupiszewska, D. and B. Nowok (2008), Comparability of statistics on international migration flows in the European Union. In: J. Raymer and F. Willekens (Eds.), *International Migration in Europe: Data, Models and Estimates* (pp. 41-71). Chichester: John Wiley and Sons.

Kupiszewska, D. and A. Wisniowski (2009), Availability of statistical data on migration and migrant population and potential supplementary sources for data estimation. Report for MIMOSA project. Available at http://mimosa.gedap.be/Documents/Mimosa_2009.pdf.

Lanzieri, G. (2009), EUROPOP2008*:* A set of population projections for the European Union. Paper presented at the IUSSP International Population Conference, Marrakech, 27 September-2 October 2009.

Lanzieri, G. (2010), Is there a fertility convergence across the Member States of the European Union? Paper presented at the Joint Eurostat/UNECE Work Session on Demographic Projections, Lisbon, 28-30 April 2010.

Lee, R. (2006), Mortality forecasts and linear life expectancy trends. In Bengtsson, T. (Ed.), *Prospectives on Mortality Forecasting. III*. Stockholm: National Social Insurance Board, pp. 19-40.

Lee, R.D. and L. Carter (1992), Modeling and forecasting the time series of U.S. mortality. *Journal of the American Statistical Association*, 87, 659-671.

Lee, R.D. and T. Miller (2001), Evaluating the performance of the Lee-Carter method for forecasting mortality. *Demography*, 38, 537-549.

Lesthaeghe, R. and D. van de Kaa (1986), Twee demografische transities?. In: R. Lesthaeghe and D. van de Kaa (Eds.), *Groei of krimp*? (Two demographic transitions?) Van Loghum-Slaterus, Deventer, 9-24.

Lesthaeghe, R. and K. Neels (2002), From the first to the second demographic transition: An interpretation of the spatial continuity of demographic innovation in France, Belgium and Switzerland. *European Journal of Population*, 18, 325-360.

Lesthaeghe, R. and J. Surkyn (2002), New forms of household formation in Central and Eastern Europe: Are they related to newly emerging value orientations. *Economic Survey of Europe*, 1, 197-216.

Li, N. and R. Lee, (2005), Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method. *Demography*, 42, 575-594.

Lutz, W. (2009), Toward a systematic, argument based approach to defining assumptions for population projections. Interim Report IR-09-037, Laxenburg, Austria: International Institute for Applied Systems Analysis.

Lutz, W. and J. Goldstein (2004), Introduction: How to deal with uncertainty in population forecasting. *International Statistical Review*, 72, 1-4.

Lutz, W., W.C. Sanderson and S. Scherbov (1998), Expert-based probabilistic population projections. *Population and Development Review*, 24, 139-155.

Lutz, W., V. Skirbekk and M.R. Testa (2006), The low-fertility trap hypothesis: Forces that may lead to further postponement and fewer births in Europe. *Vienna Yearbook for Population Research*, 167-192.

Massey, D.S., J. Arango, G. Hugo, A. Kouaouci, A. Pellegrino and J.E. Taylor (1993), Theories of international migration: A review and appraisal. *Population and Development Review*, 20, 699-751.

McNeil, D.R., T.J. Trussell and J.C. Turner (1977), Spline interpolation of demographic data. *Demography*, 14, 245-252.

McNown, R., A. Rogers and J. Little (1995), Simplicity and complexity in extrapolative population forecasting models. *Mathematical Population Studies*, 5, 235-257.

Mitra, S. and A. Romaniuk (1973), Pearsonian Type I curve and its fertility projection potentials. *Demography*, 10, 351-365.

Naz, G. (2000), Determinants of fertility in Norway. Working Papers in Economics, Bergen University, Norway.

Nicolaas, H. (2004), Helft Nederlandse emigranten keert terug (Half of Dutch emigrants returns). *Bevolkingstrends*, 53(2), 39-45.

Nowok, B., D. Kupiszewska and M. Poulain (2006), Statistics on international migration flows. In: M. Poulain (2006), *THESIM: Towards Harmonised European Statistics on International Migration* (pp. 203-232). Louvain-la-Neuve: Presses universitaires de Louvain.

Oeppen, J. and J.W. Vaupel (2002), Broken limits to life expectancy. *Science*, 296, 1029-1031.

Official Journal of the European Union (2007), Regulation (EC) No 862/2007 of the European Parliament and of the Council of 11 July 2007 on Community statistics on migration and international protection and repealing Council Regulation (EEC) No 311/76 on the compilation of statistics on foreign workers. Available at http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:199:0023:0029:EN:PDF.

Olshansky, S.J. and B.A. Carnes (1994), Demographic perspectives on human senescence. *Population and Development Review*, 20, 57-80.

Olshansky, S.J., D.P. Goldman, Y. Zheng and R.W. Rowe (2009), Aging in America in the twenty-first century: Demographic forecasts from the MacArethur Foundation Research Network on an aging society. *The Milbank Quarterly*, 87, 842-862.

Olshansky, S.J., D.J. Passaro, R.C. Hershow, J. Layden and B.A. Carnes *et al.* (2005), A potential decline in life expectancy in the United States in the 21st century. New England *Journal of Medicine*, 352, 1138-1145.

Peristera, P. and A. Kostaki (2007), Modeling fertility in modern populations. *Demographic Research*, 16, 141-194.

Persson, L. (2010), Trend reversal in childlessness in Sweden. Paper presented at the Joint Eurostat/UNECE Work Session on Demographic Projections, Lisbon, 28-30 April 2010.

Peto, R., A.D. Lopez, J. Boreham and M. Thun (2005), *Mortality from smoking in developed countries 1950-2000*. Oxford: Oxford University Press.

Poulain, M. (1993), Confrontation des Statistiques de migrations intra-européennes: Vers plus d'harmonisation? *European Journal of Population*, 9(4), 353-381.

Poulain, M. (1995), Towards a harmonisation of migration statistics within the European Community. In: S. Voets, J. Schoorl and B. de Bruijn (Eds.), *Demographic consequences of international migration* (pp. 11-25). The Hague: Netherlands Interdisciplinary Demographic Institute.

Poulain, M. (1999), International migration within Europe: Towards more complete and reliable data? Working Paper 12, joint ECE-Eurostat Work Session on Migration Statistics, Geneva, Switzerland.

Poulain, M. and L. Dal (2008), Estimation of flows within the intra-EU migration matrix. Report for the MIMOSA project. Available at http://mimosa.gedap.be/Documents/Poulain_2008.pdf

Raiffa, H. (1997). Decision analysis. Introductory readings on choices under uncertainty. Columbus, Ohio: McGraw-Hill.

Raymer, J. (2008), Obtaining an overall picture of population movement in the European Union. In: J. Raymer and F. Willekens (Eds.), *International Migration in Europe: Data, Models and Estimates* (pp. 209-234). Chichester: John Wiley and Sons.

Raymer, J. and G. Abel (2008), The MIMOSA model for estimating international migration flows in the European Union. Joint UNECE/Eurostat Work Session on Migration Statistics, Working paper 8. Geneva: UNECE / Eurostat. Available at: http://www.unece.org/stats/documents/ece/ces/ge.10/2008/wp.8.e.pdf.

Raymer, J. and P.W.F. Smith (2010), Modelling migration flows. Journal of the Royal Statistical Society: *Series A (Statistics in Society)*, 173, 703-705.

Raymer, J. and F. Willekens (Eds.) (2008), *International migration in Europe: Data, models and estimates*. Chichester: John Wiley and Sons.

Raymer, J., J. de Beer and R. van der Erf (2011), Putting the pieces of the puzzle together: Age and sex-specific estimates of migration between EU / EFTA countries, 2002-2007. *European Journal of Population*, 27, 185–215.

Reher, D.S. (1998), Family ties in Western Europe: Persistent contrasts. *Population and Development Review*, 24, 203-234.

Renshaw, A.E. and S. Haberman (2003), On the forecasting of mortality reduction factors. Insurance: *Mathematics and Economics*, 32, 379-401.

Renshaw, A.E. and S. Haberman (2006), A cohort-based extension to the Lee-Carter model for mortality reduction factors. Insurance: *Mathematics and Economics*, 38, 556-570.

Rogers, A. (1986), Parameterized multistate population dynamics and projections. *Journal of the American Statistical Association*, 81, 48-61.

Romaniuk, A. (1973), A three parameter model for birth projections. *Population Studies,* 27, 467-478.

Sandberg, K. and T. Westerberg (2005), Spatial dependence and the determinants of child births in Swedish municipalities 1974-2002. Paper presented at the workshop op Spatial Econometrics, Kiel, Germany, April 8-9, 2005.

Schmertmann, C.P. (2003), A system of model fertility schedules with graphically intuitive parameters. *Demographic Research*, 9, 82-110.

Schmertmann, C. (2005), Quadratic spline fits by nonlinear least squares. *Demographic Research*, 12, 105-106

Schoen, R. (2004), Timing effects and the interpretation of period fertility. *Demography*, 41, 801-819.

Sobotka, T. (2004), Is lowest-low fertility in Europe explained by the postponement of childbearing? *Population and Development Review*, 30, 195-220.

Sobotka, T. and F. Adigüzel (2002), Religiosity and spatial demographic differences in the Netherlands, SOM Research Report 02F65, University of Groningen.

Statistics Sweden (2009), *The future population of Sweden 2009-2060*. Statistics Sweden, Stockholm.

Stewart, S.T., D.M. Cutler and A.B. Rosen (2009), Forecasting the effects of obesity and smoking on U.S. life expectancy. *New England Journal of Medicine*, 36, 2252-2260.

Surkyn, J. and R. Lesthaeghe (2004), Value orientations and the second demographic transition (SDT) in Northern, Western and Southern Europe: An update. *Demographic Research, Special Collection*, 3, 43-86.

Tabeau, E., A. van den Berg Jeths and R. Heathcote (Eds.) (2001), *Forecasting mortality in developed countries*. Dordrecht: Kluwer Academic Publishers.

Ter Bekke, S., H. van Dalen and K. Henkens (2005), Emigratie van Nederlanders geprikkeld door bevolkingsdruk (Emigration of Dutch nationals triggered by population pressure). *Demos*, 21, 25-28.

Thierry, X., A. Herm, D. Kupiszewska, B. Nowok and M. Poulain (2005), How the UN recommendations and the forthcoming EU regulation on international migration statistics are fulfilled in the 25 EU countries? Paper presented at the XXV International Population Conference, Tours, 18-23 July 2005. Available at http://iussp2005.princeton.edu/download.aspx?submissionId=51414.

Tuljapurkar, S., N. Li and C. Boe (2000), A universal pattern of mortality decline in the G7 countries. *Nature*, 405, 789-792.

UNECE (2009), Improving migration statistics by exchange of data between countries. 95th DGINS conference “Migration – Statistical mainstreaming”, 1 October 2009, Malta. Available at http://epp.eurostat.ec.europa.eu/portal/page/portal/conferences/documents/95th_dgins_conference/UNECE.pdf.

United Nations (1998), Recommendations on statistics of international migration. Statistical Papers Series M, No. 58, Rev.1, Department of Economic and Social Affairs, Statistics Division, United Nations, New York. Available at http://unstats.un.org/unsd/publication/SeriesM/SeriesM_58rev1E.pdf.

United Nations (2000), Replacement migration: Is it a solution to declining and ageing populations? New York: United Nations, Department of Economic and Social Affairs.

United Nations (2002), *Measuring international migration: Many questions, few answers. Population Division*. Department of Economic and Social Affairs, United Nations, New York.

Van der Erf, R. and N. van der Gaag (2007), An iterative procedure to revise available data in the double entry migration matrix for 2002, 2003 and 2004. Discussion Paper, Netherlands Interdisciplinary Demographic Institute, The Hague. Available at http://mimosa.gedap.be/Documents/Erf_2007.pdf.

Van Imhoff, E. (2001), On the impossibility of inferring cohort fertility measures from period fertility measures. *Demographic Research*, 5, 23-64.

Van de Kaa, D. (1987), Europe's second demographic transition. *Population Bulletin*, 42(1).

Van de Kaa, D. (2001), Postmodern fertility preferences: From changing value orientation to new behavior. *Population and Development Review*, 27, 290-331.

Van Wissen, L. and R. Jennissen (2008), A simple method for inferring substitution and generation from gross flows: Asylum seekers in Europe. In: J. Raymer and F. Willekens (Eds.), *International Migration in Europe: Data, Models and Estimates*. Chichester: John Wiley, 235-251.

White, K.M. (2002), Longevity advances in high-income countries, 1955-96. *Population and Development Review*, 28, 59-67.

Willekens, F. (1994), Monitoring international migration flows in Europe: Towards a statistical data base combining data from different sources. *European Journal of Population*, 10, 1-42.

Willekens, F. (1999), Modeling approaches to the indirect estimation of migration flows: From entropy to EM. *Mathematical Population Studies*, 7(3), 239-276.

Wilmoth, J.R., K. Andreev, D. Jdanov and D.A. Glei (2007), Methods protocol for the Human Mortality Database. http://www.mortality.org/

Wilson, C. (2001), On the scale of global demographic convergence 1950-2000. *Population and Development Review*, 27, 155-172

Yntema, L. (1969), On Hadwiger's fertility function. *Statistical review of the Swedish National Central Bureau of Statistics*, 7, 113-117.

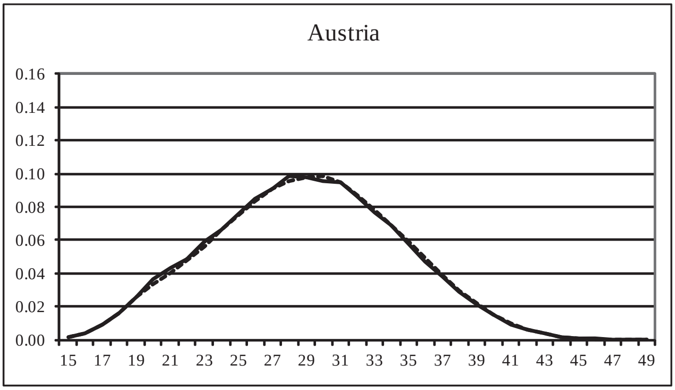
Zeng Yi, Wang Zhenglian, Ma Zhongdong and Chen Chunjun (2000), A simple method for projecting or estimating α and β, An extension of the Brass Relational Gompertz Fertility Model. *Population Research and Policy Review*, 19, 525-549.

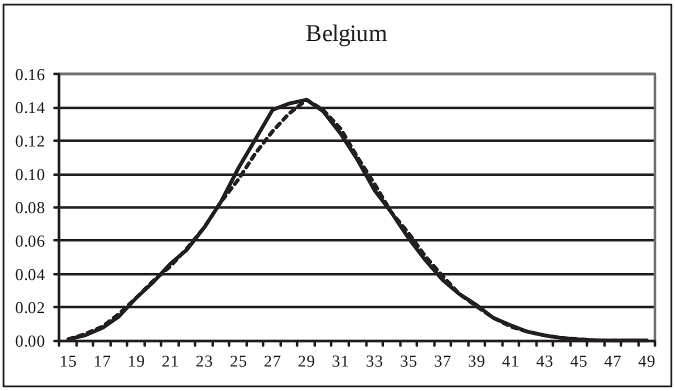
Zlotnik, H. (1987), The concept of international migration as reflected in data collection systems. *International Migration Review*, 21(4), 925-946.

Annex A

Annex to chapter 5

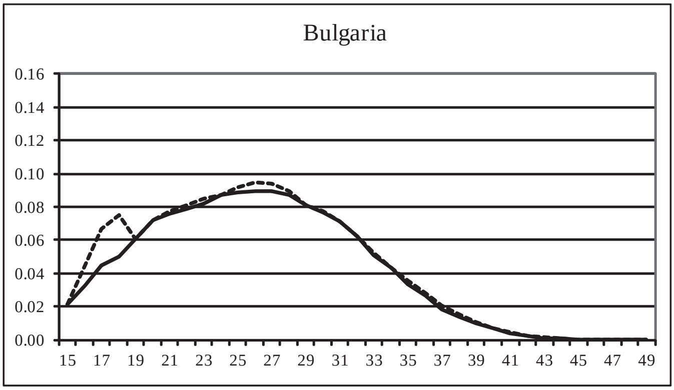
Figure A.1. Age-specific fertility rates of Austria and Belgium and fit by TOPALS, 2008

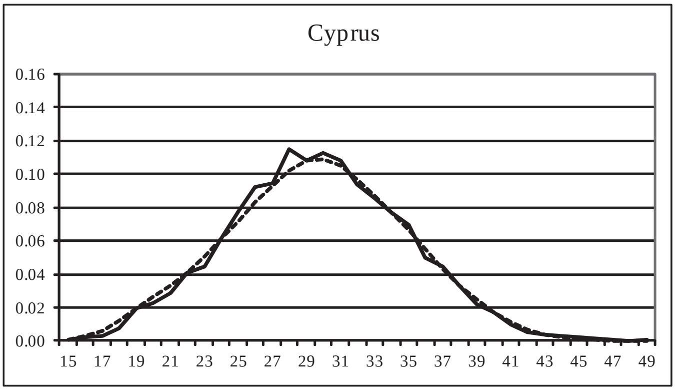




Solid lines: observations; doted lines: TOPALS.

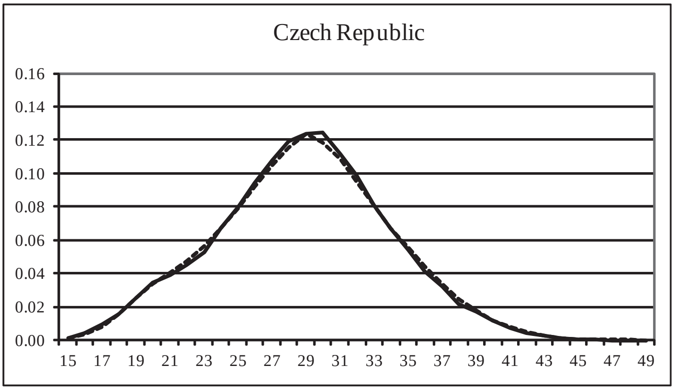
Figure A.2. Age-specific fertility rates of Bulgaria and Cyprus and fit by TOPALS, 2008

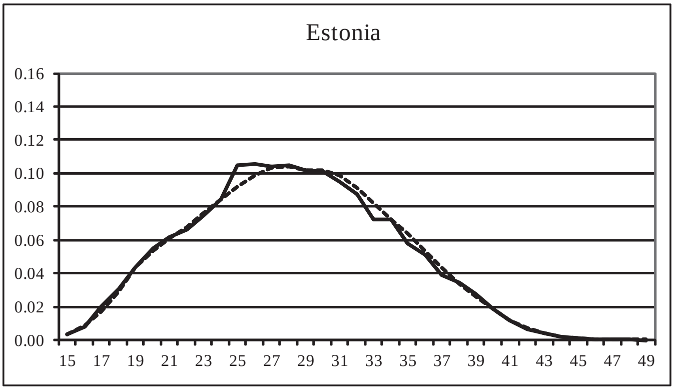




Solid lines: observations; doted lines: TOPALS.

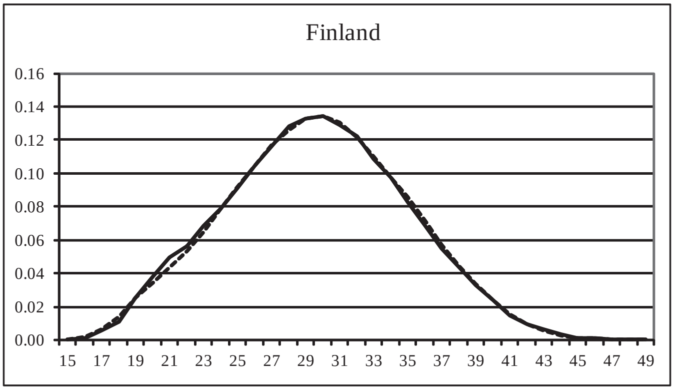
Figure A.3. Age-specific fertility rates of Czech Republic and Estonia and fit by TOPALS, 2008

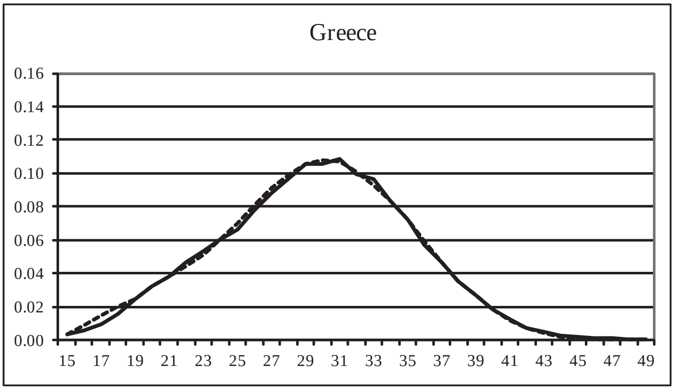




Solid lines: observations; doted lines: TOPALS.

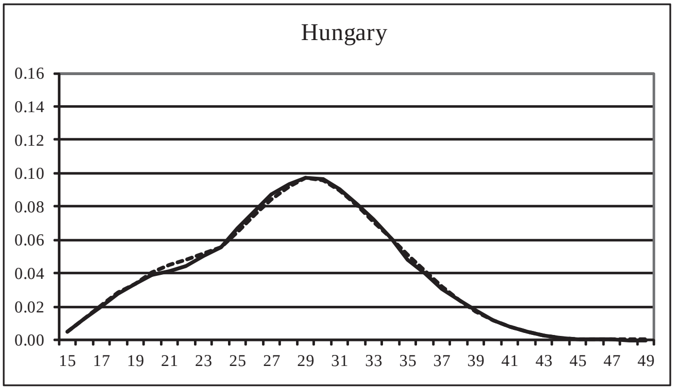
Figure A.4. Age-specific fertility rates of Finland and Greece and fit by TOPALS, 2008

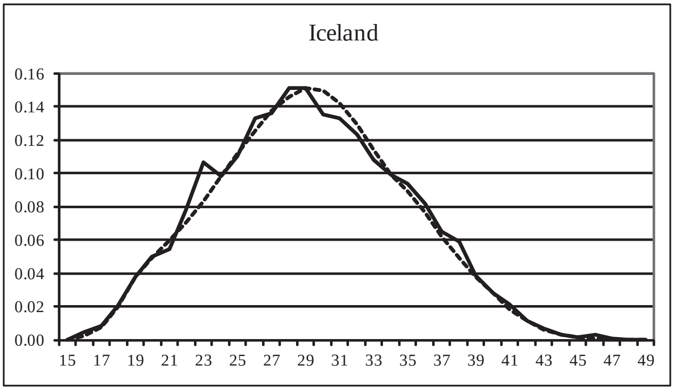




Solid lines: observations; doted lines: TOPALS.

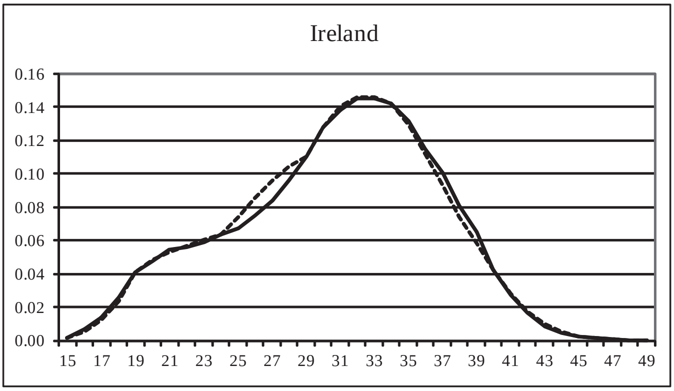
Figure A.5. Age-specific fertility rates of Hungary and Iceland and fit by TOPALS, 2008

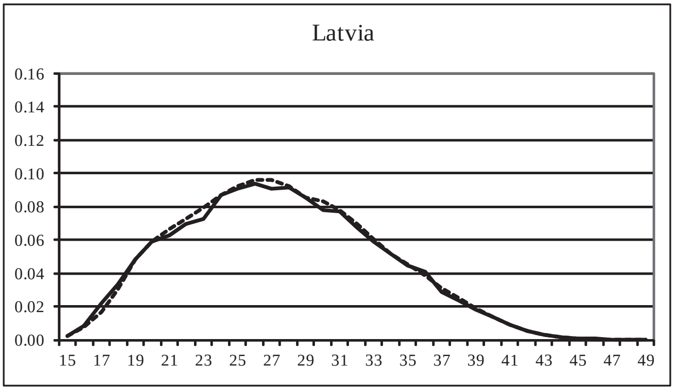




Solid lines: observations; doted lines: TOPALS.

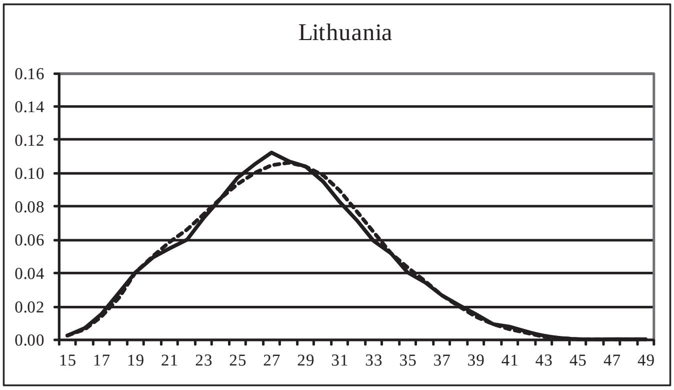
Figure A.6. Age-specific fertility rates of Ireland and Latvia and fit by TOPALS, 2008

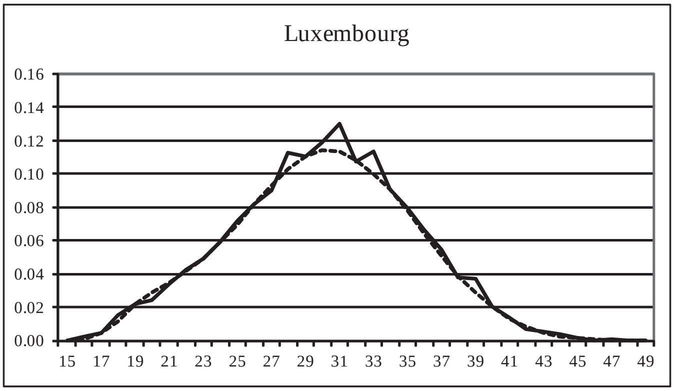




Solid lines: observations; doted lines: TOPALS.

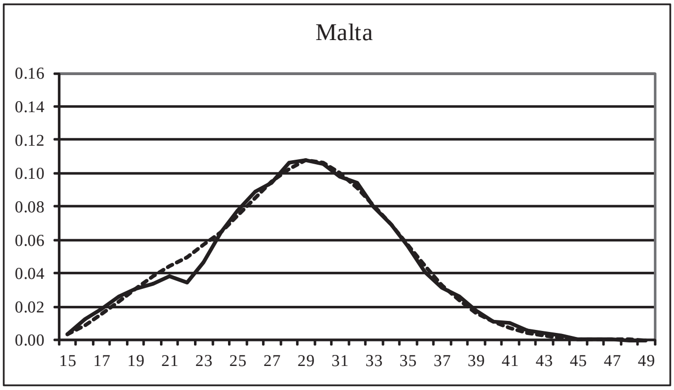
Figure A.7. Age-specific fertility rates of Lithuania and Luxembourg and fit by TOPALS, 2008

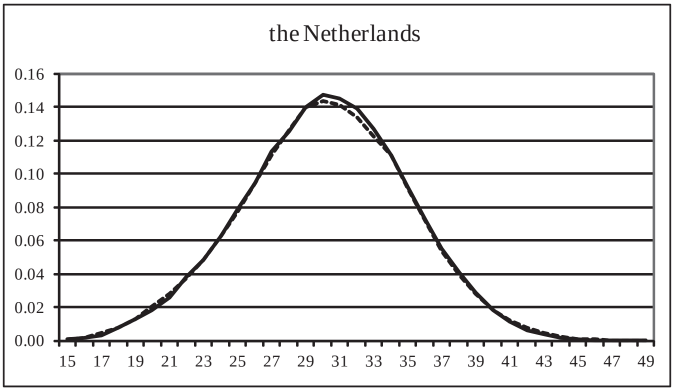




Solid lines: observations; doted lines: TOPALS.

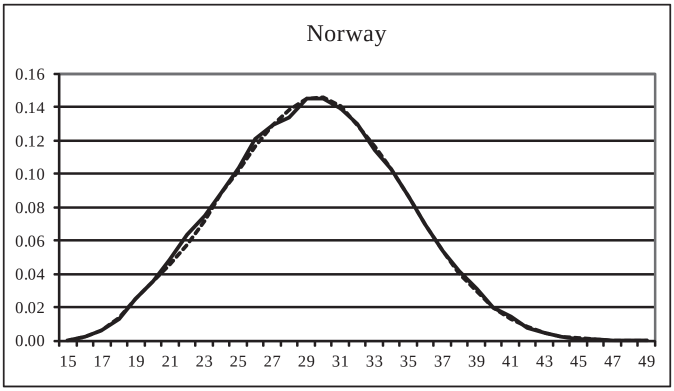
Figure A.8. Age-specific fertility rates of Malta and the Netherlands and fit by TOPALS, 2008

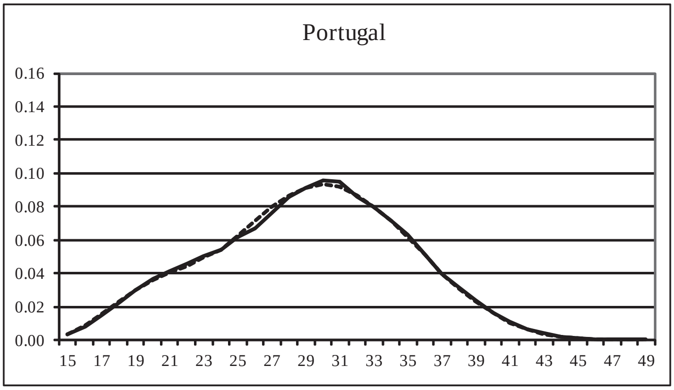




Solid lines: observations; doted lines: TOPALS.

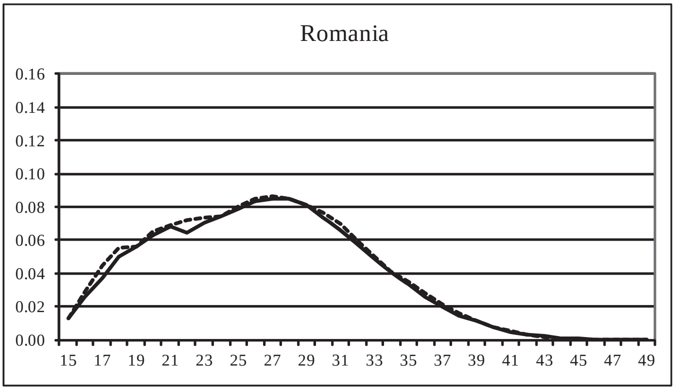
Figure A.9. Age-specific fertility rates of Norway and Portugal and fit by TOPALS, 2008

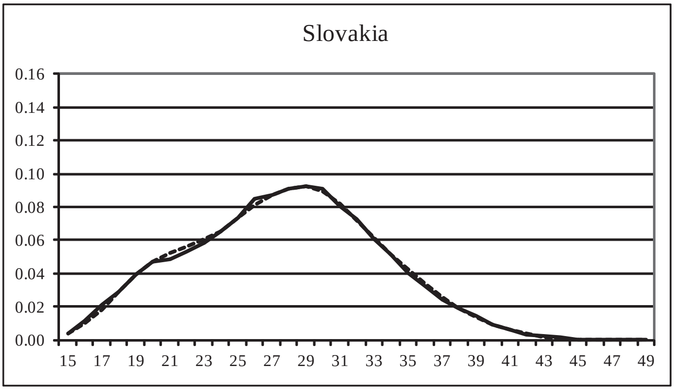




Solid lines: observations; doted lines: TOPALS.

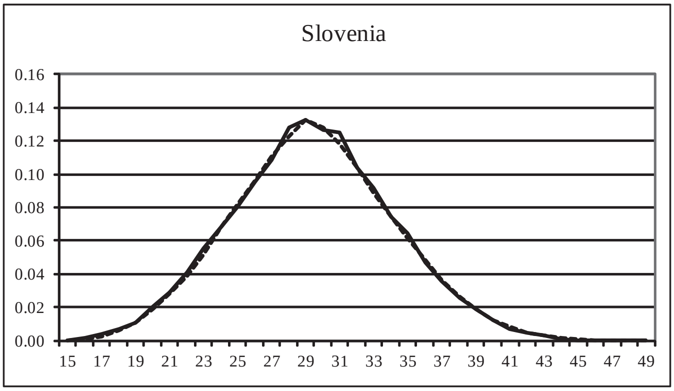
Figure A.10. Age-specific fertility rates of Romania and Slovakia and fit by TOPALS, 2008

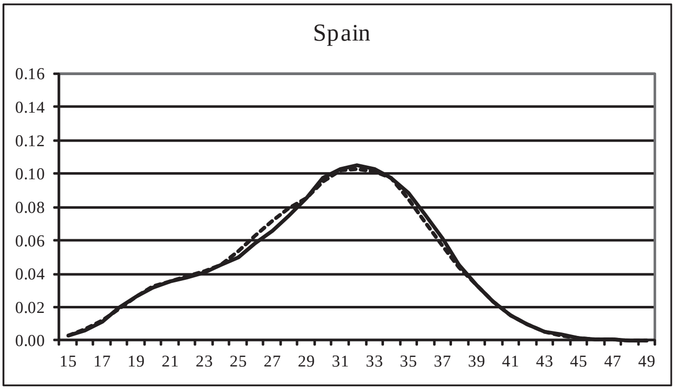




Solid lines: observations; doted lines: TOPALS.

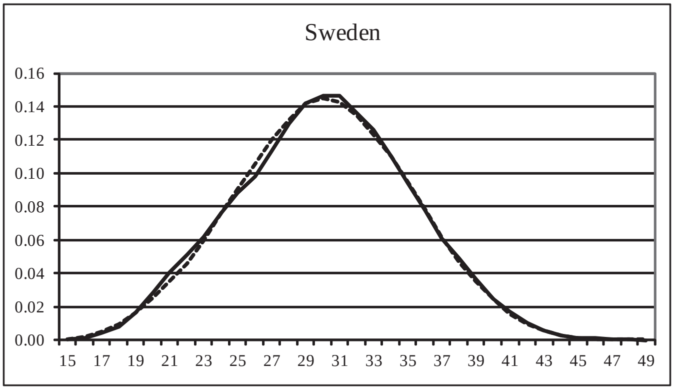
Figure A.11. Age-specific fertility rates of Slovenia and Spain and fit by TOPALS, 2008

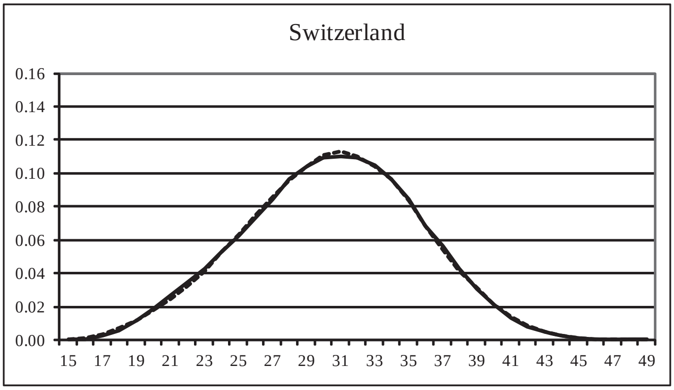




Solid lines: observations; doted lines: TOPALS.

Figure A.12. Age-specific fertility rates of Sweden and Switzerland and fit by TOPALS, 2008





Solid lines: observations; doted lines: TOPALS.

Annex B

Annex to chapter 6

Table B.1. Values of the risk ratios of age-specific death probabilities of European countries and the average of Northern, Western and Southern European countries at the knots, 2006, males

	0-20	30	40	50	60	70	80	90	100	109	e_o
Austria	1.16	1.03	0.78	0.87	1.00	1.00	0.99	1.04	1.10	1.04	77.1
Belarus	2.27	5.43	5.27	3.95	3.45	2.66	1.69	1.40	1.01	0.91	63.6
Belgium	1.06	1.07	0.99	1.06	1.16	1.02	1.10	1.11	1.09	1.03	76.5
Bulgaria	2.38	1.55	2.20	2.43	2.48	2.01	1.67	1.32	1.09	0.99	69.2
Czech Republic	1.33	0.98	1.32	1.44	1.68	1.49	1.31	1.25	1.17	1.06	73.5
Denmark	0.99	0.90	1.09	1.28	1.15	1.23	1.18	1.12	1.08	1.02	75.9
Estonia	2.09	3.97	3.27	3.15	3.08	2.18	1.57	1.26	1.09	1.00	67.4
Finland	1.10	1.44	1.22	1.37	1.39	1.12	1.07	1.12	1.06	1.01	75.8
France	1.01	1.11	1.31	1.35	1.09	0.94	0.91	0.89	1.01	1.00	77.2
Germany	0.93	0.81	0.91	1.02	1.00	1.05	1.01	1.06	1.09	1.04	77.2
Hungary	1.47	1.50	2.51	3.12	2.40	1.92	1.46	0.87	0.86	0.81	69.2
Ireland	1.17	1.27	0.64	0.72	0.95	1.07	1.17	1.09	1.04	0.99	77.3
Italy	0.86	0.86	0.79	0.72	0.85	0.91	0.96	0.98	0.99	0.98	78.6
Latvia	2.42	3.79	4.91	4.53	3.39	2.39	1.71	1.24	1.02	0.94	65.6

	0-20	30	40	50	60	70	80	90	100	109	e_o
Lithuania	2.66	4.86	5.44	4.29	3.09	2.05	1.51	1.11	0.85	0.79	65.3
Netherlands	0.92	0.76	0.69	0.81	0.85	1.11	1.14	1.11	1.11	1.04	77.6
Norway	0.99	0.99	0.75	0.72	0.85	0.88	0.95	1.15	1.09	1.03	78.1
Poland	1.56	1.87	2.47	2.35	2.12	1.71	1.33	1.08	0.96	0.92	70.9
Portugal	1.40	1.37	1.73	1.33	1.14	1.07	1.13	1.09	1.07	1.01	75.5
Russia	3.60	10.07	7.88	5.30	3.72	2.79	1.74	1.34	1.09	0.98	60.3
Slovakia	1.73	1.77	1.91	2.22	2.17	2.06	1.54	1.17	1.05	0.96	70.4
Spain	1.14	1.07	1.10	1.05	1.05	1.00	0.94	0.96	0.95	0.95	77.6
Sweden	0.92	0.70	0.61	0.71	0.78	0.95	0.98	1.10	1.12	1.05	78.7
Switzerland	0.99	0.93	0.71	0.66	0.78	0.87	0.86	0.94	1.03	1.01	79.1
Ukraine	3.06	6.87	7.13	4.76	3.38	2.62	1.80	1.35	1.07	0.96	62.3
United Kingdom	1.10.	1.22	1.03	0.93	0.91	1.03	1.04	0.99	0.96	0.95	77.2

Note: e_o is life expectancy at birth.

Table B.2. Values of the risk ratios of age-specific death probabilities of European countries and the average of Northern, Western and Southern European countries at the knots, 2006, females

	0-20	30	40	50	60	70	80	90	100	109	e_o
Austria	1.05	0.77	0.80	0.92	0.91	1.00	0.99	1.13	1.12	1.05	82.7
Belarus	2.07	3.48	2.29	2.12	2.06	2.24	1.88	1.48	1.07	0.95	75.5
Belgium	1.09	1.22	1.16	1.31	1.11	1.08	1.02	1.07	1.08	1.02	82.2
Bulgaria	2.55	2.05	1.91	1.94	1.67	2.15	1.93	1.56	1.19	1.03	76.3
Czech Republic	0.94	1.25	1.18	1.15	1.33	1.47	1.46	1.30	1.18	1.05	79.9
Denmark	0.94	0.98	1.23	1.25	1.34	1.60	1.18	1.06	1.02	0.98	80.5
Estonia	1.86	1.93	1.68	1.81	2.13	1.46	1.38	1.28	1.07	0.99	78.6
Finland	1.11	1.19	1.10	1.41	1.10	1.00	0.99	1.07	1.08	1.03	82.8
France	0.99	0.98	1.17	1.18	0.90	0.82	0.75	0.81	0.96	0.98	84.1
Germany	0.89	1.01	0.98	1.14	1.00	1.04	1.07	1.14	1.12	1.05	82.3
Hungary	1.25	1.60	2.07	2.19	1.85	1.86	1.60	1.14	1.01	0.94	77.7
Ireland	1.06	1.19	0.97	0.89	0.97	1.08	1.09	1.12	0.98	0.95	81.9
Italy	0.87	0.80	0.81	0.78	0.84	0.89	0.89	0.97	0.98	0.98	84.1
Latvia	2.13	2.08	2.22	2.89	2.35	1.96	1.57	1.46	1.19	1.05	76.5

	0-20	30	40	50	60	70	80	90	100	109	e_o
Lithuania	2.03	2.91	2.51	2.38	2.11	1.76	1.55	1.39	1.15	1.03	77.1
Netherlands	0.97	1.22	1.12	1.20	1.08	1.09	1.07	1.16	1.10	1.04	81.9
Norway	1.04	0.89	0.91	0.98	1.11	0.97	0.96	1.15	1.10	1.04	82.7
Poland	1.39	1.10	1.50	1.63	1.56	1.50	1.35	1.20	1.06	0.99	79.6
Portugal	1.31	1.22	1.06	1.10	1.05	1.04	1.10	1.19	1.09	1.02	82.2
Russia	3.11	5.82	4.18	2.95	2.41	2.57	2.00	1.53	1.18	1.03	73.2
Slovakia	1.42	0.83	1.47	1.31	1.47	1.72	1.71	1.29	1.12	1.00	78.4
Spain	0.98	0.74	1.05	0.85	0.73	0.82	0.89	1.02	1.01	1.00	84.1
Sweden	0.91	1.10	0.89	0.86	1.04	1.14	0.91	1.05	1.10	1.05	82.9
Switzerland	1.04	0.83	0.84	0.90	0.71	0.89	0.79	0.95	1.02	1.01	84.0
Ukraine	2.69	4.66	3.77	2.55	2.28	2.60	2.02	1.61	1.23	1.05	73.8
United Kingdom	1.10	1.28	1.16	1.21	1.06	1.27	1.12	1.04	0.98	0.96	81.5

Note: e_o is life expectancy at birth.

Table B.3. Estimated values of coefficient φ of the partial adjustment model, males

	0-20	30	40	50	60	70	80	90	100	109
Austria	0.9555	0.9629	0.9568	0.9726	0.9819	0.9732	0.9798	0.9880	0.9926	0.9974
Belarus	0.9559	0.9671	0.9708	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Belgium	0.9569	0.9799	0.9750	0.9777	0.9749	0.9743	0.9800	0.9809	0.9955	0.9992
Bulgaria	0.9570	0.9840	0.9777	1.0000	1.0000	0.9939	0.9951	0.9969	1.0000	1.0000
Czech Republic	0.9595	0.9680	0.9864	0.9840	0.9891	0.9801	0.9833	0.9880	0.9960	0.9989
Denmark	0.9637	0.9799	0.9764	0.9803	0.9832	0.9847	0.9901	0.9933	0.9966	0.9989
Estonia	0.9713	0.9546	0.9706	0.9914	0.9993	0.9947	0.9918	0.9913	0.9924	0.9956
Finland	0.9632	0.9767	0.9712	0.9666	0.9659	0.9762	0.9785	0.9843	0.9948	0.9989
France	0.9596	0.9875	0.9808	0.9787	0.9844	0.9733	0.9792	0.9826	0.9945	0.9986
Germany	0.9415	0.9702	0.9711	0.9794	0.9812	0.9737	0.9795	0.9865	0.9966	1.0000
Hungary	0.9574	0.9823	0.9966	1.0000	1.0000	0.9908	0.9889	0.9819	0.9859	0.9888
Ireland	0.9592	0.9829	0.9341	0.9563	0.9703	0.9755	0.9784	0.9826	0.9944	0.9979
Italy	0.9508	0.9817	0.9696	0.9662	0.9733	0.9742	0.9807	0.9857	0.9927	0.9972
Latvia	0.9725	0.9662	0.9781	0.9940	1.0000	0.9951	0.9917	0.9851	0.9939	0.9952

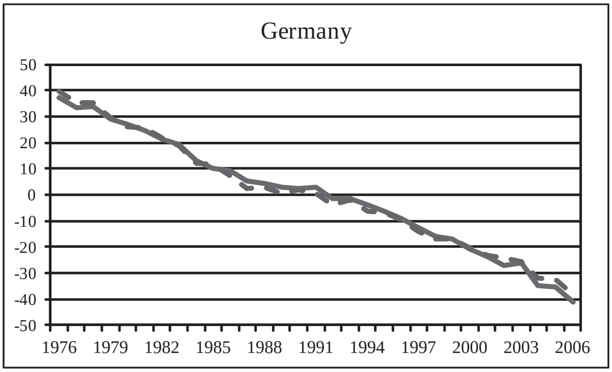
	0-20	30	40	50	60	70	80	90	100	109
Lithuania	0.9725	0.9665	0.9799	1.0000	1.0000	0.9988	0.9934	0.9963	0.9994	0.9982
Netherlands	0.9639	0.9865	0.9712	0.9770	0.9756	0.9810	0.9898	0.9953	0.9996	1.0000
Norway	0.9665	0.9627	0.9700	0.9711	0.9755	0.9769	0.9861	0.9900	0.9965	0.9994
Poland	0.9699	0.9877	0.9925	0.9958	0.9937	0.9910	0.9898	0.9907	0.9928	0.9954
Portugal	0.9399	0.9795	0.9707	0.9746	0.9743	0.9712	0.9821	0.9856	0.9923	0.9969
Russia	0.9702	0.9929	0.9980	1.0000	1.0000	1.0000	0.9949	0.9959	1.0000	1.0000
Slovakia	0.9610	0.9730	0.9829	0.9941	0.9997	0.9922	0.9933	0.9889	0.9949	0.9967
Spain	0.9558	0.9841	0.9819	0.9832	0.9832	0.9784	0.9782	0.9886	0.9952	0.9991
Sweden	0.9601	0.9572	0.9599	0.9695	0.9794	0.9787	0.9851	0.9902	0.9976	1.0000
Switzerland	0.9616	0.9728	0.9760	0.9746	0.9753	0.9742	0.9806	0.9849	0.9932	0.9976
Ukraine	0.9616	0.9797	1.0000	1.0000	1.0000	1.0000	0.9983	0.9966	0.9984	0.9988
United Kingdom	0.9634	1.0000.	0.9891	0.9728	0.9731	0.9752	0.9811	0.9858	0.9915	0.9962
Convergence scenario	0.9546	0.9857	0.9794	0.9773	0.9794	0.9756	0.9811	0.9871	0.9949	0.9987
Acceleration scenario	0.9116	0.9715	0.9588	0.9548	0.9642	0.9517	0.9622	0.9747	0.9899	0.9974

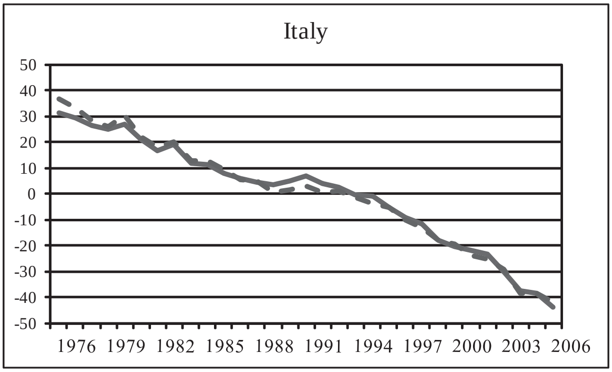
Table B.4. Estimated values of coefficient φ of the partial adjustment model, females

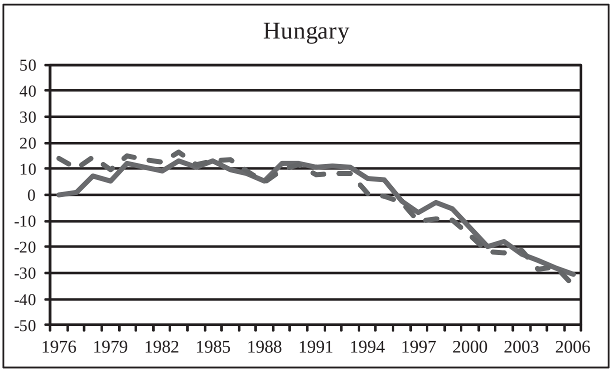
	0-20	30	40	50	60	70	80	90	100	109
Austria	0.9470	0.9338	0.9637	0.9576	0.9746	0.9690	0.9712	0.9858	0.9930	0.9984
Belarus	0.9472	0.9342	0.9647	0.9636	0.9922	0.9998	0.9960	1.0000	1.0000	1.0000
Belgium	0.9548	0.9301	0.9755	0.9777	0.9740	0.9735	0.9721	0.9755	0.9910	0.9973
Bulgaria	0.9531	0.9783	0.9892	1.0000	0.9920	0.9861	0.9900	0.9949	1.0000	1.0000
Czech Republic	0.9557	0.9462	0.9474	0.9796	0.9866	0.9820	0.9844	0.9869	0.9945	0.9982
Denmark	0.9713	0.9137	0.9539	0.9677	0.9825	0.9895	0.9872	0.9877	0.9922	0.9965
Estonia	0.9804	0.8776	0.9410	0.9618	0.9812	0.9840	0.9833	0.9906	0.9951	0.9983
Finland	0.9704	0.9136	0.9534	0.9872	0.9730	0.9679	0.9735	0.9833	0.9958	1.0000
France	0.9543	0.9785	0.9797	0.9784	0.9843	0.9683	0.9710	0.9782	0.9909	0.9975
Germany	0.9425	0.9748	0.9689	0.9777	0.9769	0.9702	0.9746	0.9834	0.9947	0.9992
Hungary	0.9526	0.9745	0.9815	0.9961	0.9905	0.9863	0.9853	0.9828	0.9906	0.9945
Ireland	0.9483	0.9377	0.9353	0.9453	0.9561	0.9722	0.9705	0.9850	0.9906	0.9959
Italy	0.9449	0.9730	0.9707	0.9671	0.9685	0.9715	0.9715	0.9804	0.9908	0.9971
Latvia	0.9790	0.8894	0.9752	0.9875	0.9907	0.9918	0.9866	0.9911	1.0000	1.0000

	0-20	30	40	50	60	70	80	90	100	109
Lithuania	0.9790	0.8895	0.9753	0.9878	0.9918	0.9905	0.9875	0.9977	1.0000	1.0000
Netherlands	0.9607	0.9818	0.9839	0.9846	0.9880	0.9812	0.9812	0.9901	0.9971	1.0000
Norway	0.9536	0.9328	0.9579	0.9735	0.9828	0.9805	0.9777	0.9893	0.9966	1.0000
Poland	0.9608	0.9701	0.9836	0.9904	0.9846	0.9855	0.9861	0.9894	0.9949	0.9982
Portugal	0.9278	0.9310	0.9557	0.9679	0.9687	0.9654	0.9735	0.9799	0.9914	0.9973
Russia	0.9610	0.9707	0.9849	0.9934	0.9980	0.9995	0.9947	1.0000	1.0000	1.0000
Slovakia	0.9530	0.9395	0.9606	0.9765	0.9889	0.9860	0.9859	0.9861	0.9953	0.9978
Spain	0.9470	0.9511	0.9792	0.9624	0.9637	0.9633	0.9712	0.9827	0.9943	0.9994
Sweden	0.9506	0.9414	0.9326	0.9620	0.9852	0.9798	0.9756	0.9867	0.9977	1.0000
Switzerland	0.9537	0.9354	0.9451	0.9584	0.9671	0.9702	0.9733	0.9790	0.9918	0.9975
Ukraine	0.9534	0.9406	0.9638	0.9841	0.9989	0.9990	0.9942	0.9995	1.0000	1.0000
United Kingdom	0.9617	0.9554	0.9778	0.9719	0.9795	0.9806	0.9800	0.9830	0.9918	0.9970
Convergence scenario	0.9537	0.9817	0.9790	0.9755	0.9781	0.9734	0.9749	0.9829	0.9932	0.9983
Acceleration scenario	0.9057	0.9648	0.9576	0.9517	0.9563	0.9481	0.9499	0.9659	0.9865	0.9966

Figure B.1. Time dependent parameter k_t of the Lee-Carter model







Solid line: Men; dashed line: Women.

List of NIDI books

83. Joop de Beer, *Transparency in population forecasting: Methods for fitting and projecting fertility, mortality and migration*. 2011, pp. 256
82. Nico van Nimwegen, *Demography Monitor 2008. Demographic trends, socioeconomic impacts and policy implications in the European Union.* 2010, pp. 161.
81. Judith P.M. Soons, *Love, life and happiness: A study of partner relationships and well-being in young adulthood.* 2009, pp. 175.
80. Nico van Nimwegen and Liesbeth Heering, *Bevolkingsvraagstukken in Nederland anno 2009: Van groei naar krinp. Een demografische omslag in beeld. (Population issues in the Netherlands, 2009: From population growth to decline. Perspectives on a demographic turning point).* 2009, pp. 240.
79. Anne Elisabeth van Putten, The role of intergenerational transfers in gendered labour patterns, 2009, pp. 215.
78. Kène Henkens, Harry van Dalen and Hanna van Solinge, De vervagende grens tussen werk en pensioen: over doorwerkers, doorstarters en herintreders. (*The fading line between work and pension)*, 2009, pp. 129.
77. Pearl A. Dykstra, *Ageing, intergenerational solidarity and age-specific vulnerabilities*, 2008. pp. 167.
76. Tineke Fokkema, Susan ter Bekke and Pearl A. Dykstra. S*olidarity between parents and their adult children in Europe.* 2008. pp. 125.
75. Harry van Dalen and Kène Henkens. Weg uit Nederland: emigratie aan het begin van de 21e eeuw. *(Leaving the Netherlands: Emigration at the start of the 21 century)*. 2008, pp. 134.
74. Harry van Dalen, Kène Henkens and Joop Schippers. *Oudere werknemers door de lens van de werkgever. (Older employees through the eyes of the employer)* 2007, pp. 122. € 11,50.
73. Harry van Dalen, Kène Henkens, Wilma Henderikse and Joop Schippers, *Dealing with an ageing labour force: What do European employers expect and do?* 2006, pp. 55, € 10.
72. Nico van Nimwegen and Gijs Beets (eds.), *Social situation observatory. Demography monitor 2005. Demographic trends, socioeconomic impacts and policy implications in the European Union*. 2006, pp. 375, € 35.
71. N. van Nimwegen and I. Esveldt (eds.), *Bevolkingsvraagstukken in Nederland anno 2006: de grote stad in demografisch perspectief. (Population issues in the Netherlands, 2006: The big city in demographic perspective).* 2006, pp. 340, € 25.
70. Hanna van Solinge, *Changing tracks. Studies on life after early retirement in the Netherlands*. 2006, pp. 157, € 15.

A NIDI book (1-74) can be ordered by remitting the amount due, plus postage and administrative costs (€ 5,00), to bank account number 45.83.68.687 (ABN-AMRO, The Hague) in the name of NIDI-KNAW, The Hague, mentioning the relevant report number with reference to the SWIFT-code: ABNANL2A and the IBAN-code: NL56ABNA0458368687. The address of the ABN-AMRO is P.O. BOX 90, 1000 AB in Amsterdam. If you wish to order more than one report, please telephone us (+31-70-3565200) as the editions are limited.

Report 75 etc. can be ordered at Amsterdam University Press, Herengracht, 1016 BG Amsterdam, info@aup.nl, www.aup.nl.

(Subject to changes)

The accuracy of a new forecast is unknown. One can only judge the methods used and assumptions made by the forecaster. Thus transparency of forecasts is important. This requires that the forecaster argues the choice of methods and assumptions.

Population projections are based on assumptions about future changes in fertility, mortality and migration. This book describes methods that can be used for making these assumptions and discusses which choices have to be made by the forecaster.

BOOK 83

The *Netherlands Interdisciplinary Demographic Institute* (NIDI) is an institute of the *Royal Netherlands Academy of Arts and Sciences* (KNAW)

This book shows how international comparisons can be helpful in making population projections. For international migration international comparisons can contribute to improve estimates of the size and direction of immigration and emigration flows.

For projecting fertility international comparisons can be used to assess in what direction the future level and age pattern of fertility may move. Projections of future life expectancy can be based on a comparison with the highest level of life expectancy in the world. Moreover estimating an average trend in life expectancy over a group of countries will make projections more stable.

ISSN 0922-7210
ISBN 978-90-6984-637-8

9 789069 846378