



Project 2006/S 100-106607/EN, LOT 2
Modelling of statistical data
on migration and migrant populations

MIMOSA

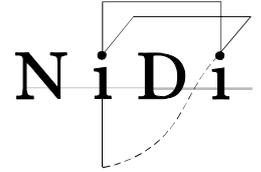
Methodology for the estimation of population
stocks by country of birth, sex and age

12 June 2008

Prepared by
Dorota Kupiszewska and Jakub Bijak
Central European Forum for Migration and Population Research
(CEFMR)
Warsaw, Poland

Main contractor:

Netherlands Interdisciplinary Demographic Institute (NIDI)
Lange Houtstraat 19, NL-2511 CV The Hague, The Netherlands
tel. +31 70 356 52 00, fax: +31 70 364 71 87
e-mail: info@nidi.nl, web: <http://www.nidi.knaw.nl/en/>



Partner responsible for this report:

Central European Forum for Migration and Population Research
ul. Twarda 51/55, PL-00818 Warsaw, Poland
tel. +48 22 697 88 34, fax: +48 22 697 88 43
e-mail: cefmr@cefmr.pan.pl, web: www.cefmr.pan.pl



Central European Forum for Migration and Population Research (CEFMR) is a research partnership of the Foundation for Population, Migration and Environment (BMU-PME) from Zurich, Institute of Geography and Spatial Organization of the Polish Academy of Sciences and IOM – International Organization for Migration. The views presented in this paper are the ones of the authors and do not necessarily reflect the opinions of the founding organisations of the CEFMR. The authors of particular sections are listed in alphabetic order.

Table of contents

Introduction	6
1. Population stocks by country of birth – current national practices	7
2. Availability of the 2002–2007 data on population stocks by country of birth for 31 countries	11
3. Potential methods of estimating population by country of birth, sex and age	14
3.1 Notation and basic concepts	14
3.2 Interpolation of 5-year into 1-year age groups	15
3.3 Interpolation of population stocks	16
3.3.1 Cohort-wise interpolation	16
3.3.2 Period-wise interpolation	18
3.4 Cohort-component projections	18
3.5 Proportional fitting methods	20
3.5.1 Proportional adjustment / decomposition	20
3.5.2 Direct proportional fitting	20
3.5.3 Iterative proportional fitting	21
3.6 Shares propagation methods	21
3.6.1 Empirical data	22
3.6.2 Period-wise shares propagation	22
3.6.3 Cohort-wise shares propagation	23
3.6.4 Mixed solutions	24
3.7 Auxiliary methods	26
3.8 Can population by country of birth be estimated from the citizenship structure? 26	
4. Country-specific recommendations of estimation methods	28
4.1 Austria	29
Availability of population stocks data	29
Proposed estimation methodology	29
4.2 Belgium	30
Availability of population stocks data	30
Proposed estimation methodology	30
4.3 Bulgaria	31
Availability of population stocks data	31
Proposed estimation methodology	31
4.4 Cyprus	32
Availability of population stocks data	32
Proposed estimation methodology	32
4.5 Czech Republic	33
Availability of population stocks data	33
Proposed estimation methodology	33
4.6 Denmark	34
Availability of population stocks data	34
Proposed estimation methodology	34
4.7 Estonia	35
Availability of population stocks data	35
Proposed estimation methodology	35

4.8 Finland	36
Availability of population stocks data.....	36
Proposed estimation methodology	36
4.9 France	37
Availability of population stocks data.....	37
Proposed estimation methodology	37
4.10 Germany	38
Availability of population stocks data.....	38
Proposed estimation methodology	39
4.11 Greece	40
Availability of population stocks data.....	40
Proposed estimation methodology	40
4.12 Hungary	41
Availability of population stocks data.....	41
Proposed estimation methodology	41
4.13 Ireland	42
Availability of population stocks data.....	42
Proposed estimation methodology	43
4.14 Italy	44
Availability of population stocks data.....	44
Proposed estimation methodology	44
4.15 Latvia	46
Availability of population stocks data.....	46
Proposed estimation methodology	46
4.16 Lithuania	47
Availability of population stocks data.....	47
Proposed estimation methodology	47
4.17 Luxembourg	48
Availability of population stocks data.....	48
Proposed estimation methodology	48
4.18 Malta	49
Availability of population stocks data.....	49
Proposed estimation methodology	49
4.19 Netherlands	50
Availability of population stocks data.....	50
Proposed estimation methodology	50
4.20 Poland	51
Availability of population stocks data.....	51
Proposed estimation methodology	51
4.21 Portugal	52
Availability of population stocks data.....	52
Proposed estimation methodology	52
4.22 Romania	53
Availability of population stocks data.....	53
Proposed estimation methodology	53

4.23 Slovakia	54
Availability of population stocks data.....	54
Proposed estimation methodology	54
4.24 Slovenia	55
Availability of population stocks data.....	55
Proposed estimation methodology	55
4.25 Spain	56
Availability of population stocks data.....	56
Proposed estimation methodology	56
4.26 Sweden	58
Availability of population stocks data.....	58
Proposed estimation methodology	58
4.27 United Kingdom	59
Availability of population stocks data.....	59
Proposed estimation methodology	59
4.28 Iceland	61
Availability of population stocks data.....	61
Proposed estimation methodology	61
4.29 Liechtenstein	62
Availability of population stocks data.....	62
Proposed estimation methodology	62
4.30 Norway	63
Availability of population stocks data.....	63
Proposed estimation methodology	63
4.31 Switzerland	64
Availability of population stocks data.....	64
Proposed estimation methodology	64
5. Summary and conclusion	65
References	68

Tables and Figures

Table 1. Sources and methods of producing annual population statistics in 31 European countries	8
Table 2. Availability of data on population by country of birth, sex and age in the JMQ, 31 countries, 2002–2007	12
Table 3. Overview of recent Census data on population by country of birth, sex and age, 31 countries	13
Table 4. Coefficients for the Karup-King interpolation formula.....	16
Table 5. Summary table: availability of data on population by country of birth and proposed estimation methods.....	66
Figure 1. Cohort-wise interpolation of population stocks: a general idea.....	17
Figure 2. Relationships between population stocks P^n , and events B , D^n and X^n on a Lexis diagram ..	19
Figure 3. Age structures of foreign-born population, Nordic EU and Lithuania	23
Figure 4. Differences between various age structures: cohort-wise and period-wise approaches	25
Figure 5. Sample age structures of foreign-born and foreign citizens, Nordic EU and Lithuania	27

Introduction

The aim of the report is to present the methodology which has been selected for the estimation of population stocks as of 1st January, for the countries for which the information is unavailable or incomplete, in a breakdown by sex, age group and broad category of country of birth. It has to be noted that according to the methodological standards, adopted also by Eurostat, this concept refers to the country according to the current territorial division. Hence, it takes into account border changes, such as those, which occurred in Central and Eastern Europe following the First and Second World Wars (Germany, Poland), or after the more recent break-ups of various countries (Czechoslovakia, Yugoslavia).

The study covers 31 European countries, of which 27 belong to the European Union (as of 1st January 2007), and further four – to EFTA (Iceland, Liechtenstein, Norway and Switzerland). The period of interest is 2002–2007. The country of birth groups of interest are persons born in the country, persons born in an EU Member State other than the country of residence and persons born outside the EU. The age groups are five-year, with the last, open-ended category being 85 years or more. As per the requirements stated in the call for tenders, the proposed estimation should be able to combine data from different sources (population census, vital statistics, specific surveys, etc.), taking into account different definitions in use in various countries, as well as the effects of various legal regulations in place, possible regularisations programmes, etc. At this stage of the project, the data that are already available are not modified (for example, in order to harmonise definitions, or for any other reason), unless in the case of inconsistencies between the sources. In the latter cases, the demographic data of Eurostat (from the DEMO domain) are given priority.

The report is structured in five sections. Firstly, in Section 1, current national practices in producing population stocks by country of birth are briefly discussed. Section 2 contains summary information on the availability of the 2002–2007 data on population stocks by country of birth for 31 countries under study, although without going into country-specific details. In Section 3, the potential methods of estimating population stocks by sex, age groups and country of birth are discussed. This methodological section presents such methods as estimation of data in single years of age from 5-year age-groups, cohort-wise and period-wise interpolation / extrapolation of population stocks or shares, cohort-component projections, proportional fitting, as well as other, auxiliary methods. Subsequently, Section 4 contains recommendations of estimation methods for each of the countries under study separately. The discussion is summarised in Section 5, containing synthetic guidelines as to the application of the proposed approaches for particular countries.

The study is based chiefly on data available in the Eurostat databases, supplemented by additional information from respective National Statistical Institutes, whenever required and feasible. A detailed information on the availability and reliability of data are assessed in Section 4 for each country under study. In the methodological part, standard demographic and algebraic techniques are discussed and the references to scientific literature are provided. Throughout the report, the abbreviation ‘NSI’ is used to denote the National Statistical Institute of a respective country, ‘JMQ’ stands for the Joint Questionnaire on International Migration Statistics (Joint Migration Questionnaire) of Eurostat, UN Statistical Division (UNSD), UN Economic Commission for Europe (UNECE), the Council of Europe and International Labour Office (ILO). ‘LFS’ depicts the Labour Force Survey.

1. Population stocks by country of birth – current national practices

The review of national practices concerning statistics on population by country of birth, sex and age in European countries indicates that the availability of these statistics is worse than in the case of statistics by citizenship. The most complete source of information are population censuses. In the last census round (2000–2002, except France – 1999 and Malta – 1995 and 2005, Ireland – 2002 and 2006) the question on the country of birth was asked in all countries that conducted traditional census. All other countries (out of 31 countries investigated in MIMOSA) prepared the statistics on population by country of birth using information from their registers of population. The only exception is Germany, where there was no traditional census in the last 20 years (the most recent one was conducted in 1987) and statistics using alternative sources were not produced within the last census round.

The availability of annual statistics by country of birth, sex and age in individual European countries depends mainly on the sources and methods of producing annual population statistics. Table 1 gives the summary of the situation in the 31 countries. It presents the sources and methods used to produce official figures concerning total population and also the figures disaggregated by citizenship or country of birth, if produced. Countries which prepared annual statistics by country of birth (for all years 2002–2007, or for some years only), or can easily do it (Belgium), are marked in bold. Countries that produce some statistics by country of birth for foreigners only are marked in italics.

In general, all countries where population statistics are based directly on the population register can provide statistics on population by country of birth. There are 11 countries in this group: Austria, Belgium, Denmark, Finland, Latvia, Lithuania, the Netherlands, Slovenia, Sweden, Iceland and Norway. The situation in Liechtenstein remains to be checked with the NSI. Most countries which rely on the component method do not produce statistics on population by country of birth. However, there are some exceptions from these general rules, as described below.

Italy is the only country that uses population register for producing population statistics but does not provide statistics on population by country of birth. This is linked to the way the statistics are generated. ISTAT (the Italian NSI) collects data from the registers (*anagrafi*) in individual municipalities through four surveys requesting aggregated data: “Municipal Resident Population by sex, year of birth and marital status” (POSAS), “Movement and calculation of resident population” (form ISTAT P.2), “Non-Italian Municipal Resident Population by year of birth and sex” (STRASA) and “Movement and calculation of resident foreign population” (form ISTAT P.3). No information on population by country of birth was requested in these surveys till the reference date 1 January 2006 inclusive. However, the question on the number of foreigners born in Italy present in the country as of 1 January 2007 appeared in form ISTAT P.3 requesting the data on the demographic balance of population for 2006. The same form was used for the collection of data for the next year. Accordingly, since the reference date 1 January 2007, ISAT publishes figures on the number of foreigners born in Italy, but no data on population by country of birth are available before that date (except census figures).

The remaining 18 countries rely on the component method to prepare statistics on total population by sex and age using the component method. Out of these, only Romania and Spain prepare also complete statistics on population by country of birth, sex and age group. Five countries – France, Germany, Ireland, Slovakia and the UK – have some data on population by country of birth but not complete.

Table 1. Sources and methods of producing annual population statistics in 31 European countries

Country	Population register	Register of foreigners/ Residence permit register	Calculations using the component method	Estimations using the component method	Other
Austria	x				
Belgium	x				
Bulgaria				x	
Cyprus				x	
Czech Republic		x	x		
Denmark	x				
Estonia				x	
Finland	x				
<i>France</i>				x ¹	x ¹
<i>Germany</i>		x ²	x		
Greece		x ³		x	
Hungary		x	x		
Ireland				x	x ⁴
<i>Italy</i> ⁵	x				
Latvia	x				
Lithuania	x	x ⁶	x		
Luxembourg				x ⁷	
Malta				x	
Netherlands	x				
Poland			x		
Portugal		x		x	
Romania				x	
Slovakia	x ⁸	x ⁷	x ⁷		
Slovenia	x	x			
Spain	x		x		
Sweden	x				
United Kingdom				x	x ⁹
Iceland	x				
Liechtenstein (?)	x				
Norway	x				
Switzerland			x		

Note: Countries that produce statistics on population by country of birth are marked in bold. Countries that produce some statistics by country of birth for foreigners only are marked in italics

¹ FR – The “rolling census” was introduced in 2004. Earlier, the component method was used.

² DE – Central Register of Foreigners is used to produce statistics on foreigners by citizenship and by country of birth.

³ GR – Residence permit register is used to produce statistics on foreigners by citizenship

⁴ IE – Quarterly National Household Survey is used to produce statistics on population by citizenship and by country of birth

⁵ IT – Since 2007 Italy publishes the total number of foreigners born in Italy

⁶ LT – Data from the Personalisation of Identity Documents Centre under the Ministry of the Interior were used for estimating foreigners by citizenship on 1.01.2006.

⁷ LU – Estimates by citizenship are adjusted taking into account the figures in the Population Register

⁸ SK – For 2004, 2006 and 2007, statistics on population by country of birth were prepared for foreigners only, based on the register on foreigners. For 2005, statistics were provided for the whole population and the “demographic statistics” (the component method?) of the NSI and the Population Register were indicated as the sources used to prepare the data.

⁹ UK – Labour Force Survey

It is not entirely sure how Romania produces its statistics on population by country of birth (complete breakdown by country of birth, sex and 5-year age group is prepared since 1.01.2006, while data by country of birth and sex only were provided also for 1.01.2004). According to the recent report by Poulain, Kupiszewska, Perrin and Gheorghiu (2007), all population statistics, including the disaggregations by citizenship and by country of birth are done using the component method. If so, Romania would be the only country that uses the component method alone to provide statistics on population by country of birth. It should be noted that in the Joint Migration Questionnaire, population register and register of foreigners are indicated as the sources of data for producing statistics on population by citizenship and by country of birth and the component method is not mentioned.

A specific situation is in Spain, where two sets of annual population data are prepared and published on the website of the Spanish NSI (INE - *Instituto Nacional de Estadística*). The first one is labelled by INE as “Population Now-Cast” and is prepared for population by sex and age. The “Population Now-Cast” is prepared using the component method, in which births and deaths are taken into account through, respectively, fertility rates and survival probabilities. Migration flows are estimated separately for foreign and Spanish nationals. The second set of population statistics contains data from the population register (*Padrón*). The total population according to both sets differ. The figures from the Population Now-Cast are considered by INE to provide the best estimates and are available in DEMO domain of the Eurostat on-line database. Estimations of population by country of birth or by citizenship are not part of the “Population Now-Cast”. Nevertheless, the statistics by country of birth, sex and age provided in the Joint Migration Questionnaire for 2005–2007 are consistent with the Population Now-Cast figures by sex and age, while the breakdown by country of birth was estimated using information from the population register. The data provided in the JMQ for 2002 originate from the population register.

Two countries that use the component method to produce annual figures on population by sex and age prepare statistics on population by country of birth for foreigners only, based on the data from the registers of foreigners. These are Germany and Slovakia. For Slovakia, these statistics are provided in the Joint Migration Questionnaire since the reference date 1 January 2004 (by individual countries of birth and sex for 2004 and additionally in disaggregation by age since 2006). The exception is 2005: the data for this year concern the whole population, not foreigners only, however the numbers include 274 thousand persons with unknown country of birth (the number of foreigners was 22 thousand).

As concerns Germany, the data on foreigners by country of birth are neither provided in the JMQ nor published in the Destatis database, but data on the number of foreigners born in Germany and born abroad can be found for some years in other places in the internet (but without more detailed information on population by individual countries of birth, sex and age; see also Section 4.10).

Another distinctive group of countries that use the component method for producing annual statistics on population by sex and age and provide also some data on population by country of birth, sex and age are Ireland and the UK. In both countries, the annual statistics by country of birth are based on surveys: the Quarterly National Household Survey in Ireland and the Labour Force Survey in the UK. The reference date of the data provided to Eurostat is not 1 January but mid-April in Ireland and mid-year average in the UK (in the UK, based on the data collected across four quarters). The Irish figures are disaggregated only by broad age groups (0–14, 15–24, 25–44, 45–64 and 65+ years) and provided for selected categories of country of birth only: persons born in the EU (EU15 for 2002–2005, EU25 for 2006), persons born in Ireland, in the UK and in the USA. In the UK, statistics are more complete but some data are also missing: data are marked as unavailable for these countries of birth for which the data were considered unreliable (due to a low number of persons born in these countries interviewed in the survey).

The last exception is France, which is different from all other countries and derives its annual estimates of population from the “rolling census”. Annual statistics on population by country of birth are not published, but there are indications that some data might be available or can be produced in future. In particular, the mid-2004 data on the number of foreigners born in France, foreigners born

abroad and persons born abroad as foreigners who acquired French citizenship, were provided in the publication *Insee Première* n°1098 (published in August 2006; see Section 4.9 for detailed references).

It should be noted that even if the countries listed above had information necessary for producing the complete statistics by country of birth, sex and age, the data are often not available for all years 2002–2007. In particular, as presented in detail in the Section 2, the data cross-tabulated by country of birth and age are not available from the Eurostat on-line database for the years 2002–2004. This is because in the Joint Questionnaire on International Migration the age dimension was introduced to the table on population by country of birth since the reference date 1 January 2005.

No annual statistics by on population by country of birth have been identified in the remaining 11 countries, i.e.: Bulgaria, Cyprus, Czech Republic, Estonia, Greece, Hungary, Luxembourg, Malta, Poland, Portugal and Switzerland. In these countries, population censuses remain the key sources of information on population distribution by sex, age and country of birth.

2. Availability of the 2002–2007 data on population stocks by country of birth for 31 countries

Annual statistics on usually resident population by country of birth, sex and age are collected by Eurostat from the NSIs via the Joint Questionnaire on International Migration, together with migration flow data. Population statistics for 37 European countries, collected through the JMQ are checked and corrected, and subsequently loaded into Eurostat’s on-line database. The data are located under the *Population and Social conditions* theme, in the *International Migration and Asylum* domain (MIGR), tables: *migr_st_popctb* (population by country of birth and sex) and *migr_st_popctba* (population by country of birth, sex and age group)¹⁰. The data for 2000–2007 come usually from the following tables in the 2000–2007 JMQs:

2000–2003 data

- Table 8 in JMQs 2000–2003 Usually resident population by country of birth and sex;

2004 data

- Table 7 in JMQ 2004 Usually resident population by country of birth and sex;

2005–2007 data

- Table 7a in JMQs 2005–2007 Usually resident population by country of birth and age, both sexes;
- Table 7b in JMQs 2005–2007 Usually resident population by country of birth and age, males;
- Table 7c in JMQs 2005–2007 Usually resident population by country of birth and age, females.

A detailed analysis of statistics on population stocks by country of birth provided by the 31 countries covered the JMQs for the reference period 2002–2007. Selected results of the analysis of the data availability for particular countries are presented in Section 4 of this report, and summarised in Table 2, providing an overview of the situation for all 31 countries. The information on the lack of data, marked as ‘not available’ in Table 2, was based on the information provided in the JMQ or on information obtained during the THESIM project¹¹. Other missing data were marked as “not provided to Eurostat”. In addition to missing data, a number of other problems were detected, for example the presence of provisional data, some country of birth categories only, broad age groups, or a different reference date than 1st January.

A quick look at Table 2 is enough to see that complete data on population by country of birth, sex and age group are not available in Eurostat for any of the 31 countries. The best situation is for Finland, the Netherlands, Slovenia and Sweden, where all data requested in the JMQ were provided for all years 2002–2007, i.e. complete data by country of birth, sex and age for 2005–2007 and by country of birth and sex for 2002–2004. No data by country of birth were provided in the JMQs by half of the countries.

Data on total population stock on 1st January, not disaggregated by country of birth, are collected by Eurostat within the framework of the annual Demographic Statistics data collection. These data, disaggregated by sex and age, are located under the *Population and social conditions theme*, in the *Demography* (DEMO) domain of the database, table *demo_pjan*. The review of the availability of these data for the years 2002–2007 revealed that for all the 31 countries covered by the project there are data on population by sex and single-year age group up to at least 85+.

¹⁰ At the time of writing of this report, the table *migr_st_popage* available from the on-line Eurostat database (<http://epp.eurostat.ec.europa.eu>) contained data for 2004–2006 but not for 2002, 2003 and 2007.

¹¹ Research project *THESIM: Towards Harmonised European Statistics on International Migration*, funded by the European Commission through the Sixth Framework Programme and executed by a research consortium led by GÉDAP (Université Catholique de Louvain).

In addition to the annual data, Eurostat also collects and disseminates statistics on population by country of birth, sex and age obtained by the countries during population censuses. In the on-line database these data are located under the *Population and social conditions theme*, in the *Census (CENS)* domain of the database, table *cens_nscbirth*. The basic information on the census reference date and census type, concerning the most recent census conducted in each of the 31 countries in the 2000/2001 round of censuses, is presented in Table 3. In the same table, the results of a review of the availability of census data on resident population by country of birth, sex and age in the Eurostat database are presented. They show that census data on population by country of birth, sex and age are available for almost all of the 31 countries, unlike annual population statistics. Census data are missing for Germany and Malta, as well as for the 2006 census conducted in Ireland.

The Eurostat's collection of data from the Labour Force Survey, which was a supplementary source for estimating population stocks by citizenship, is not useful for the estimations of population by country of birth because the available tables do not provide data disaggregated by this variable. However, it is worth to note that the estimates of population by country of birth, sex and age group provided by the UK in the Joint Migration Questionnaires are based on the Labour Force Survey.

Table 2. Availability of data on population by country of birth, sex and age in the JMQ, 31 countries, 2002–2007¹²

Country	2002	2003	2004	2005	2006	2007
Austria	-cob	-cob	-cob	-cob	-	+
Belgium	-age	x, -age	-	-	-	-
Bulgaria	-	-	-	-	-	-
Cyprus	-age, dref	na	na	na	na	na
Czech Republic	-	-cob	-cob	-cob	-cob	-
Denmark	-age	-age	-age	+	+	-
Estonia	na	na	na	na	na	na
Finland	-age	-age	-age	+	+	+
France	-	-	-	-	-	-
Germany	-	-cob	-cob	-cob	p, -cob	-
Greece	-	-	-cob	-	-	-
Hungary	-	-	-cob	-cob	-cob	-
Ireland	p, ±cob, -age, dref	p, ±cob, -age, dref	p, ±cob, -age, dref	p, ±cob, broad age, dref	p, ±cob, broad age, dref	-
Italy	-	-	-	-cob, -age	-cob, -age	-
Latvia	-cob	-age	-age	-age	-age	+
Lithuania	-cob, -age	-	-age	+	+	+
Luxembourg	-	-	-cob	-cob	-cob	-cob (na)
Malta	-	-	-	-	-	-cob
Netherlands	-age	-age	-age	+	+	+
Poland	-age, dref	na	na	na	na	-cob (na)
Portugal	-	na	na	na	-	na
Romania	-age, dref	-	-age	-cob	+	+
Slovakia	-	-	for, -age	-age	for	for
Slovenia	-age	-age	-age	+	+	+
Spain	-age	-	p, -age	+	+	+
Sweden	-age	-age	-age	+	+	+
United Kingdom	-	±cob, -age, dref	±cob, dref	±cob, dref	±cob, dref	-
Iceland	-age	-age	-	-	-	-
Liechtenstein	-	-	-	-	-	-
Norway	-age	-age	-age	+	+	-
Switzerland	na	na	na	na	na	na

+ data were provided to Eurostat; **-** data were not provided to Eurostat; **-age** no disaggregation by age; **-cob** no disaggregation by country of birth; **±cob** data provided for a few country of birth categories; **broad age** data disaggregated by broad age groups; **dref** reference date different than 1 January; **for** data provided for foreigners only; **na** data not available; **p** provisional data; **x** problems have been detected in the data sent by the NSI.

¹² The age dimension appears in the table on population by country of birth since the Joint Migration Questionnaire 2005 (data for 2005)

Table 3. Overview of recent Census data on population by country of birth, sex and age, 31 countries

Country	Reference date	Census type	Availability of data in the Eurostat database: Population by country of birth, sex and age
Austria	15.05.2001	Traditional combined with the use of registers	+
Belgium	01.10.2001	Traditional combined with the use of registers	+
Bulgaria	01.03.2001	Traditional	+
Cyprus	01.10.2001	Traditional	+
Czech Republic	01.03.2001	Traditional	+
Denmark	1.01.2001	Register-based only	+
Estonia	31.03.2000	Traditional	+
Finland	31.12.2000	Register-based only	+
France ¹³	08.03.1999	Traditional	+
Germany	25.05.1987	Traditional	-
Greece	18.03.2001	Traditional	+
Hungary	01.02.2001	Traditional	+
Ireland	28.04.2002 23.04.2006	Traditional	+ -
Italy	21.10.2001	Traditional	+
Latvia	31.03.2000	Traditional combined with the use of registers	+
Lithuania	06.04.2001	Traditional	+
Luxembourg	15.02.2001	Traditional combined with the use of registers	+
Malta	26.11.1995 27.11.2005	Traditional Traditional	- -
Netherlands	1.01.2001	Register-based only	+
Poland	20.05.2002	Traditional	+
Portugal	12.03.2001	Traditional	+
Romania	18.03.2002	Traditional	+
Slovakia	26.05.2001	Traditional	+
Slovenia	31.03.2002	Traditional combined with the use of registers	± ¹⁴
Spain	01.11.2001	Traditional combined with the use of registers	+
Sweden	01.11.1990 1.01.2001 [†]	Register-based combined with a questionnaire <i>Population register data</i>	+
United Kingdom	29.04.2001	Traditional	+
Iceland	31.12.2000 [†]	<i>Population register data</i>	+
Liechtenstein	5.12.2000	Traditional combined with the use of registers	+
Norway	3.11.2001	Register-based combined with a questionnaire	+
Switzerland	5.12.2000	Traditional combined with the use of registers	+

+ data available at the Eurostat database; - data not available at the Eurostat database; ± some disaggregated data not available due to confidentiality; † reference date of the census data in the Eurostat database.

Note: In some countries, country of birth have not been specified for a significant proportion of persons.

¹³ France: Rolling census since 2004.

¹⁴ Slovenia: Data were not published if there were very few counts in a given category

3. Potential methods of estimating population by country of birth, sex and age

The current section presents a theoretical background of methods proposed for the calculations of missing elements in the population stocks by sex, age group and country of birth group. After a brief summary of the notation used, the following methods are subsequently discussed: interpolation of 5-year into 1-year age groups, regarded as a data preparatory method (Section 3.2), followed by cohort-wise interpolation of population stocks (3.3), cohort-component projections traditionally-used in demography (3.4) and cohort-wise shares propagation (3.5). Further, Section 3.6 describes selected proportional fitting methods, which category encompasses three approaches, depending on the availability of information, namely, the proportional adjustment, direct proportional fitting and iterative proportional fitting. Section 3 concludes by presenting some auxiliary methods for dealing with the *Unknown* categories, and for estimation of missing elements of age distributions (3.7).

3.1 Notation and basic concepts

Throughout the report, the notation used for population variables follows a common convention presented below. In all cases, the superscript n indicates one of the three broad groups of country of birth: persons born in the country, persons born in an EU Member State other than the country of residence and persons born outside the EU, abbreviated as N , EU and nEU , respectively, where EU refers to 27 countries, thus to the composition of the European Union as of 1st January 2007. An abbreviation FOR is used for all foreign-born population. For the transparency of presentation, the country index is skipped, as all calculations presented in the report are always country-specific. The variables in question are as follows:

Stock variables:

- $P^n(x, t)$ - Population, in the age of x years on 1st January, year t .
 $P^n(x, c)$ - Population born in the country of birth group n , in the age of x years at the census date c .

Event variables:

- $B(t)$ - Births during calendar year t ;
 $D^n(x, t)$ - Deaths of persons aged x years, born in the country of birth group n , during calendar year t ;
 $I^n(x, t)$ - Registered immigration of persons born in the country of birth group n , aged x years, during calendar year t ;
 $E^n(x, t)$ - Registered emigration of persons born in the country of birth group n , aged x years, during calendar year t ;
 $R^n(x, t)$ - Outcome of the regularisation of the status of formerly irregular residents aged x , in year t , born in the country of birth group n ;
 $S^n(x, t)$ - Statistical adjustment (official correction) concerning the size of population born in the country of birth group n , aged x , in year t , due to other reasons than regularisations;

In all cases, unless noted otherwise, age is reported in years reached during a given calendar year, and thus the events in question (deaths, migrations, etc.) correspond to parallelograms with vertical sides on the Lexis diagram. An illustration of the relevant concepts on a Lexis plane is shown in Figure 2, in Section 3.4.

Whenever necessary, the index denoting sex is added as an additional subscript $g \in \{m, f\}$ for males and females, respectively, e.g. $P_f^n(x, t)$ refers to female population stock, and $D_m^n(x, t)$ to deaths among males. In order to distinguish five-year age groups, an additional left-hand side subscript '5' is added.

For example, ${}_5P_m^n(x, t)$ refers to male population born in the country of birth group n , which was in the age of $[x, x+5)$ years on 1st January of year t . The same principle applies to almost all event variables (D, I, E, R, S), with a clear exception of B .

In some instances, for clarity of presentation, the summation of a particular variable over a given index is indicated by an asterisk in a respective place, e.g. $P^*(x, t) = \sum_n P^n(x, t)$ denotes all persons aged x , in year t , irrespective of their country of birth, and $P^n(*, t) = \sum_x P^n(x, t)$ - all persons born in the country of birth group n , without respect to their age.

3.2 Interpolation of 5-year into 1-year age groups

Among the preparatory steps for the estimation of missing data, the most frequent problem concerns disaggregation of five-year age groups of population (or events) into single years. This has to be performed in order to enable cohort-wise interpolations or cohort-component projections with yearly steps, as described in Sections 3.3 and 3.4.

If auxiliary information is available from a different source (e.g., from a census, from the previous or next year, etc.), the population size or the number of events can be disaggregated using a ‘Prorating’ method (Shryock et al., 1993: 5-61), whereby the relative distribution from the auxiliary source is imposed on the data in question. The obtained distribution might need to be further adjusted to marginal totals, by means of proportional fitting procedures, described in Section 3.5.

If the data on population stocks by sex, broad country of birth group and 5-year age group ${}_5P^n(x, t)$ are available and the stocks by sex and 1-year age group $P^*(x, t)$ are also known, then, assuming no other information about the distribution by single years, we can estimate the missing distributions for particular country of birth groups proportionally, that is as: $P^n(x+i, t) = {}_5P^n(x, t) \cdot P^*(x+i, t) / {}_5P^*(x, t)$. This is an example of the application of the direct proportional fitting method described in Section 3.5.2.

If none of the above information is available, the proposed methodology can use the well-known interpolative four-term third-difference solution of Karup and King (Shryock et al., 1993: 5-65). For each five-year group, the disaggregation into fifths is done via applying multiplicative coefficients to the global value of this group and the neighbouring ones. Different multipliers are used for the first group, the middle groups and the last group, as set forth in Table 4. For example, if we want to split a middle five-year group with population N_i into five single-year groups n_1, n_2, n_3, n_4, n_5 , then: $n_1 = 0.064 N_{i-1} + 0.152 N_i - 0.016 N_{i+1}$, $n_2 = 0.008 N_{i-1} + 0.224 N_i - 0.032 N_{i+1}$, etc. When Karup-King multipliers are used, the condition $N_i = n_1 + n_2 + n_3 + n_4 + n_5$ is automatically fulfilled.

As an alternative to the Karup-King interpolation, the six-term fifth-difference interpolative formulas of Sprague or Beers can be applied, which however use information from more surrounding groups. Methodological details can be found in Shryock et al. (1993: 5-65–5-71). In our case, the Karup-King interpolation is recommended for the sake of simplicity.

Table 4. Coefficients for the Karup-King interpolation formula

	First group, N_0			Middle groups, N_i			Last group, N_K		
	N_0	N_1	N_2	N_{i-1}	N_i	N_{i+1}	N_{K-2}	N_{K-1}	N_K
First fifth	+0.344	-0.208	+0.064	+0.064	+0.152	-0.016	-0.016	+0.112	+0.104
Second fifth	+0.248	-0.056	+0.008	+0.008	+0.224	-0.032	-0.032	+0.104	+0.128
Third fifth	+0.176	+0.048	-0.024	-0.024	+0.248	-0.024	-0.024	+0.048	+0.176
Fourth fifth	+0.128	+0.104	-0.032	-0.032	+0.224	+0.008	+0.008	-0.056	+0.248
Last fifth	+0.104	+0.112	-0.016	-0.016	+0.152	+0.064	+0.064	-0.208	+0.344

Source: Shryock et al. (1993: Table C-1, p. 5-69).

For variables depicting non-vital events, in particular for migration, the estimates for particular **cohorts** can be obtained from two neighbouring period-age estimates yielded by the Karup-King formula, split equally by half. For the first cohort, we can assume that a half of such events concern the cohort born during a given year t , while for the last, open-ended cohort, we can add up the period-age estimate for the open-ended group and a half of the events concerning the age group immediately preceding the last one. The underlying rationale is an assumption that non-vital events for the members of a given cohort are equally spread throughout the calendar year (see also Figure 2 in Section 3.4). In any case, the estimates for the eldest cohorts would be anyway close to zero for all practical migration-related applications.

Regardless of the method, if the disaggregation is performed on data broken down by sex or country of birth, the final estimate might need to be obtained by proportional fitting methods (described in Section 3.5), in order to ensure the summation to available marginal totals.

3.3 Interpolation of population stocks

3.3.1 Cohort-wise interpolation

Given information on the age structures of the population for two non-adjacent moments of time, a simple idea to obtain the missing figures for in-between moments would be to apply interpolation techniques. In this case, we propose cohort-wise interpolation for all cohorts apart from the youngest and oldest one, which are discussed separately. Overall, this method requires much less information on input than the cohort-component projections presented in the next section, but it requires information about population both before and after the moment for which the estimates are to be done. The interpolative approach is recommended for the cases where (a) the span between two points with available data is not wide (say, two-three years), and (b) no information on the distribution of deaths and migratory events by country of birth is available.

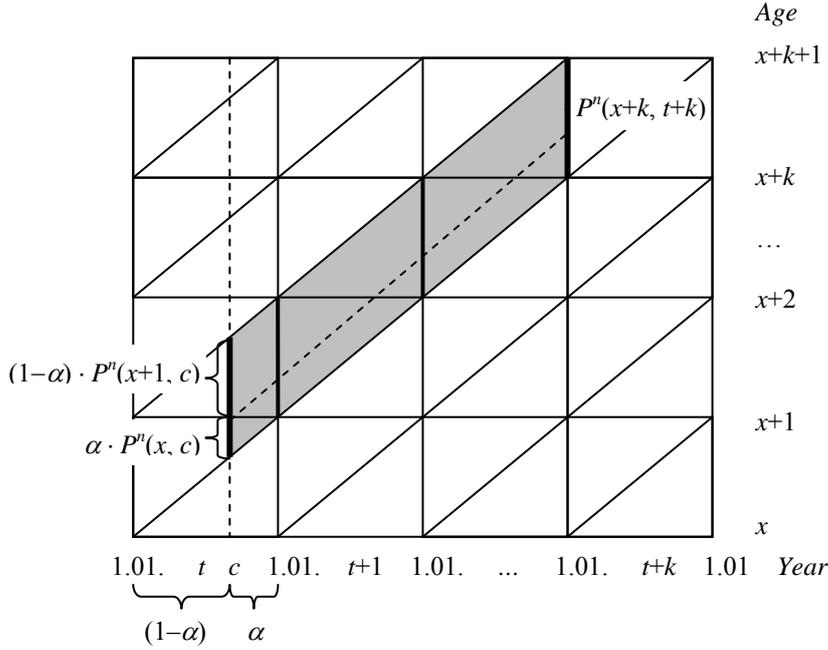
In practical applications listed further in Section 4, the situation is very often such that data are available for year t from the census conducted at time c ($t \leq c < t+1$), and for the 1st January of the year $t+k$, not immediately following the census. Such situations can be put in a general framework illustrated on a Lexis diagram in Figure 1, where α denotes the fraction of a year remaining after the census until 31st December. Figure 1 and the methodology proposed below cover also the situations when data come from other sources than the census, and the situations when the reference date of the data for year t is 1 January. In the latter case it suffices to set $\alpha = 1$.

For cohorts already existent at the census date c , the interpolation can follow various patterns, but an arithmetic and geometric pattern of growth (Rowland, 2003: 46–64; Calot and Sardon, 2003: 23–32) will be the most frequent choices. As noted by Rowland (2006: 50), “under arithmetic growth, successive population totals differ from one another by a constant amount [, while] under geometric growth they differ by a constant ratio”. For short-period interpolations, both approaches should yield

similar results, although this is an empirical issue, and there is no convincing argument in favour of either of them. Hence, a selection of appropriate methods should rely on case-specific judgements.

It has to be noted that the cohort aged x completed years on 1 January $t+k$ was split at the census date between two cohorts: the younger one (aged x completed years) and the older (aged $x+1$), as shown in Figure 1. Therefore, the interpolative estimate of $P^n(x, t+i)$ depends on $P^n(x, c)$, $P^n(x+1, c)$ and $P^n(x, t+k)$.

Figure 1. Cohort-wise interpolation of population stocks: a general idea



Source: Own elaboration

Given the above, the formula for an interpolative estimate of population sizes belonging to a particular age group $x+i$ and Country of birth group n , assuming linear pattern of change, is as follows:

$$P^n(x+i, t+i) = (k-i) / (k-1+\alpha) \cdot [\alpha \cdot P^n(x, c) + (1-\alpha) \cdot P^n(x+1, c)] + (i-1+\alpha) / (k-1+\alpha) \cdot P^n(x+k, t+k), \quad (1a)$$

while for the geometric change:

$$P^n(x+i, t+i) = \{[\alpha \cdot P^n(x, c) + (1-\alpha) \cdot P^n(x+1, c)]^{k-i} \cdot P^n(x+k, t+k)^{i-1+\alpha}\}^{1/(k-1+\alpha)}. \quad (1b)$$

In order to ensure consistency of the results and summation of the age-specific estimates to the marginal totals by sex or country of birth group, whenever available, the estimates have to be adjusted by the means of iterative proportional fitting, presented in Section 3.5.

For the youngest and oldest cohorts, the interpolation as proposed above is not possible due to reaching either minimum (zero) age towards the beginning, or maximum age towards the end of the interpolation period. In such cases we suggest to assume the country of birth structure of these cohorts as available from the later date for the youngest cohort(s) and from the earlier date for the oldest cohort(s).

The framework presented above can be easily generalised to a much less frequent situation with interpolation between two censuses – in such case, a fraction β of a year between the 1st January of the year of the second census and the second census date, c' , should be additionally accounted for.

It should be noted that an identical solution as shown above in (1a), or in (1b) can be used for extrapolating cohort sizes *beyond* the available data points, in whichever direction. In either case, it would suffice to put an appropriate integer $i \leq 0$ for the backward extrapolation (in particular, following the example from Figure 1, set $i = 0$ to obtain values for the beginning of the census year), or $i > k$ for the forward extrapolation.

Noteworthy, the methods discussed above resemble to some extent the ones presented in the *Human Mortality Database Methods Protocol* (Wilmoth et al., 2005: 15–32), with the exception of the oldest age groups, where the quoted study suggests more sophisticated extinct cohort and survivor ratios approaches. Direct application of the methods proposed by Wilmoth et al. would be, however, difficult. This is not because of computational reasons, but rather due to the lack of yearly estimates of deaths and migratory events broken down by country of birth groups, which has been listed as a pre-condition for selecting cohort-wise interpolation method at the very beginning of the current section.

3.3.2 Period-wise interpolation

As an alternative to the cohort-wise interpolation, a similar period-wise method can be proposed, especially for those age groups (in principle, the youngest ones, as shown further in Section 3.6.1), for which such methodology can yield age structures being closer to reality than the ones produced by the cohort-wise method. This feature is primarily attributable to the presence of highly age-selective migration flows in several (especially younger productive) age groups.

In the period-wise approach, interpolation is done across the calendar years, thus “horizontally” in the convention of Figure 1. The interpolation formulas (1a) and (1b) then become:

$$P^n(x, t+i) = (k-i) / (k-1+\alpha) \cdot P^n(x, c) + (i-1+\alpha) / (k-1+\alpha) \cdot P^n(x, t+k), \text{ or} \quad (2a)$$

$$P^n(x, t+i) = [P^n(x, c)^{k-i} \cdot P^n(x, t+k)^{i-1+\alpha}]^{1 / (k-1+\alpha)}, \quad (2b)$$

assuming respectively arithmetic or geometric pattern of change. Again, to ensure consistency of the resulting estimates with the marginal totals by sex or country of birth group, whenever available, the figures obtained from (2a) and (2b) need to be adjusted, for example using the iterative proportional fitting, described in Section 3.5.

3.4 Cohort-component projections

Let us denote by $X^n(x, t)$ a sum of all event variables *not* related to the natural change of population stocks (i.e. all but births and deaths), thus:

$$X^n(x, t) = I^n(x, t) - E^n(x, t) + S^n(x, t) + R^n(x, t). \quad (3)$$

The population accounting equations for each broad country of birth groups n are:

$$P^n(0, t+1) = B(t) - D^n(0, t) + X^n(0, t), \text{ for } n = N; \quad (4a')$$

$$P^n(0, t+1) = D^n(0, t) + X^n(0, t), \text{ for } n \neq N; \quad (4a'')$$

$$P^n(x, t+1) = P^n(x-1, t) - D^n(x, t) + X^n(x, t), \text{ for } x \in \{1, 2, \dots, x_{\max}-1\}; \quad (4b)$$

$$P^n(x_{\max}, t+1) = [P^n(x_{\max}-1, t) + P^n(x_{\max}, t)] - D^n(x_{\max}, t) + X^n(x_{\max}, t). \quad (4c)$$

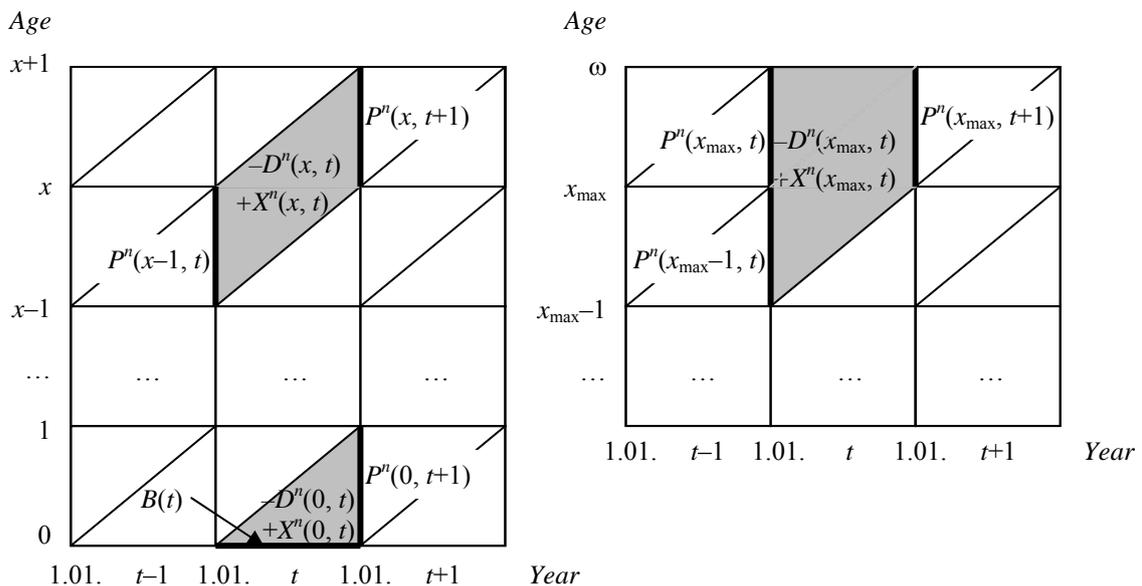
In (4c), x_{\max} stands for the highest (open-ended) age group for which information is available. Note also that deaths and other event variables in age group x_{\max} refer to the trapezoid on the Lexis diagram rather than to a parallelogram, while for age group 0 – to a right triangle (see Figure 2). Under the assumptions presented above, the projection is made following the equations (4a'), (4a''), (4b) and (4c) for consecutive years, on the basis of information available for single-year age groups, decomposed from the five-year groups, if needed.

Note that the content of variable B might differ between the countries. Theoretically, all births change population stock of the country. However, this is not the case in the countries where the legal rule specifies that children born to mothers who are not resident in this country are not registered as residents and are not added to population stocks. In such cases (probably most of the countries), only children born to mothers who are residents of the given country should be counted. In practice, this should not be a problem, because usually published statistics probably cover only those births which are included in the population balance of total population.

If the breakdown by country of birth group of all variables referring to vital and migratory events can be assumed proportional to the country of birth structure of the population at the beginning of each year, then the projection methodology can be *de facto* simplified to proportional decomposition, whereby the distribution by country of birth of the considered cohort in the previous year would directly apply to all cohorts except the first and the last one in each year. In particular, this situation applies if the following two conditions hold:

1. We may assume that the distribution of deaths and migration flows by country of birth group is the same as the country of birth composition of the population;
2. There was no regularisation, or it may be ignored.

Figure 2. Relationships between population stocks P^n , and events B , D^n and X^n on a Lexis diagram



Source: own elaboration

In such cases, the projection equation (4b) simplifies to a proportional decomposition of $P^*(x, t)$ by country of birth group (assuming that $P^*(x, t)$ is known). The estimations can be performed using the formulas:

$${}_5P^n(x, t) = {}^{\text{DEMO}}_5P^*(x, t) \cdot {}_5P^n(x-1, t-1) / {}_5P^*(x-1, t-1). \quad (5)$$

or, if available data refer to the census year, using the formulas given in the next section (on cohort-wise shares propagation).

The first and the last cohort may be disaggregated using the country of birth structure of the first and last age group in the previous year or at the census date. In such cases, the following formulas apply:

$${}_5P^n(0, t) = {}^{\text{DEMO}}{}_5P^*(0, t) \cdot {}_5P^n(0, t-1) / {}_5P^*(0, t-1), \text{ or:} \quad (6a)$$

$${}_5P^n(x_{\text{max}}, t) = {}^{\text{DEMO}}{}_5P^*(x_{\text{max}}, t) \cdot {}_5P^n(x_{\text{max}}, t-1) / {}_5P^*(x_{\text{max}}, t-1). \quad (6b)$$

3.5 Proportional fitting methods

The general task is to estimate $P_g^n(x, t)$, i.e. the elements of a three-dimensional cube (with the dimensions being sex, age and country of birth). The choice of a particular proportional fitting method depends on which marginal information (cube's edges or faces) is known and if an initial estimate of the cube elements are available. Below, the examples of possible situations are given. The formulas have been given for population by single years of age but the analogical formulas apply to population by 5-year age group. For more information on multi-proportional techniques see for example the studies of Willekens (1977, 1982), Willekens et al (1981), Rees (1994) and Norman (1999). Note that proportional fitting methods presented below are known under various names in the scientific literature.

In addition to a potential application as the main estimation method, proportional fitting may be used, in almost all the countries for which estimations are needed, as the final stage of the estimation procedure, in order to adjust the initial estimates to known aggregates or marginal totals. The initial estimate might be obtained for example using interpolation or projection, or assumed to be the same as at some different time (e.g. the same as at the census date). Such an initial estimate has to be subsequently adjusted for example to the known total population size by age and sex.

3.5.1 Proportional adjustment / decomposition

Among the proportional methods, the simplest one can be applied to situations, when a population can be directly disaggregated by a variable (sex, age or country of birth), according to the pattern observed in an auxiliary source. In general, the idea is the same as in the case of the Prorating' method (Shryock et al., 1993: 5-61), mentioned in Section 3.2.

For example, if the aggregates $P^*(x, t)$ and an initial estimate of the country of birth structure $P''^n(x, t)$ is known, then the final estimate $P''^n_g(x, t)$ may be obtained using the formula:

$$P''^n_g(x, t) = P''^n(x, t) \cdot P_g^*(x, t) / P_g^*(x, t). \quad (7)$$

If information about the country of birth structures is available from the census only (or from some earlier year) and total population by sex and age is also available, then the proportional adjustment/decomposition method might be considered as an alternative to projection, with the shares calculated using one of the three shares propagation method described in Section 3.6.

3.5.2 Direct proportional fitting

The estimation problem becomes slightly more complicated, if one wants to estimate $P_g^n(x, t)$, but does not have any initial estimate of it. One possible situation is that at least some parts of the data cube (faces and/or edges) are available and provide coherent information (sum up to the same totals).

In such cases, the most straightforward solution is provided by a direct proportional fitting method, whereby the missing elements (i.e. the inside of the cube) can be obtained by taking simple proportions to all available marginal totals.

For example, let the available data consist of known $P_g^*(x, t)$ and $P^n(*, t)$, i.e. the age-sex face and the country of birth edge of the age-sex-country of birth cube. Then, the sought-for $P_g^n(x, t)$ can be estimated as:

$$P_g^n(x, t) = P_g^*(x, t) \cdot P^n(*, t) / P^*(*, t) \quad (8)$$

In practical applications discussed in Section 4, this option will be used rather infrequently, because there usually are some initial estimates of the population structures, for example from the census. Willekens et al (1981: 97) noted that general formulas of a form akin to (8) for a one face – one edge problem, as well as similar closed-form solutions for the three edges or two faces problems are solutions of the entropy-maximisation problems in research tasks aimed at reconstructing the elements of a three-dimensional arrays, given the available marginal sums.

3.5.3 Iterative proportional fitting

In a general case, a closed-form solution (8) may not exist due to possible incoherence between the data at hand. Such problems call for a multi-step iterative proportional fitting (IPF) method, whereby the solutions are sought step-wise, through iterative adjustments of the consecutive approximations to marginal totals available from the faces or edges of the data cube. In particular, the method can be used for adjusting the existing joint preliminary distributions to known marginal distributions.

For example, let the initial estimate of the country of birth structure $P_g'^n(x, t)$ be known, so as the sex-age face and the country of birth edge of the data cube, respectively $P_g^*(x, t)$ and $P^n(*, t)$. By the IPF algorithm, the initial estimates are iteratively corrected by proportional adjustment. An additional superscript (k) in $P_g^{(k)n}(x, t)$ denotes the iteration step (for $k \geq 1$). The starting value $k = 1$ defines also the initial estimate of the joint sex-age-country of birth distribution, $P_g^{(1)n}(x, t) = P_g'^n(x, t)$.

$$P_g^{(2k)n}(x, t) = P_g^{(2k-1)n}(x, t) \cdot P_g^*(x, t) / P_g^{(2k-1)*}(x, t); \quad (9a)$$

$$P_g^{(2k+1)n}(x, t) = P_g^{(2k)n}(x, t) \cdot P^n(*, t) / P^{(2k)n}(*, t). \quad (9b)$$

The procedure defined by (9a) and (9b) is repeated iteratively till some convergence criterion is achieved. For example, the estimates yielded by subsequent steps should differ by no more than by an arbitrarily-selected small number ε . More details of the method have been discussed by Willekens (1982: 69–71), Willekens et al (1981), Rees (1994) and Norman (1999).

Although the IPF method is purely mechanical, its main advantage is that it does not require any additional information (such as data on vital events or migration) or excessive labour resources, and the obtained results (in terms of joint distributions by all variables under study) are automatically coherent with marginal distributions of particular variables. Moreover, under some not very strong assumptions, the IPF estimates can be interpreted from a statistical viewpoint as joint probability distributions obtained using the maximum likelihood or entropy maximisation methods (Bishop, Fienberg and Holland, 1975: 83–97; after: Willekens, 1982: 70–71 and Norman, 1999: 2).

3.6 Shares propagation methods

In most cases, too much information on the age-sex-country of birth distribution of the components of population change is missing, which renders projections too dubious with respect to the number of assumptions that need to be made. In practice, in such instances the only reliable information comes

from the population census and from annual population stocks available for example in the DEMO domain of the Eurostat on-line database. Three alternative procedures are proposed, described in detail in Sections 3.6.2 – 3.6.4. In Section 3.6.1, the empirical data illustrating the problem are presented.

3.6.1 Empirical data

To illustrate selected patterns observed in various age structures of the foreign-born populations, the current section presents two empirical examples, concerning three Nordic EU countries (Denmark, Finland and Sweden), being a net immigration area, as well as Lithuania, a net emigration country. For these examples, information on the country of birth structures was available either from the registers, as in the EU Nordic countries (data for 2002 and 2007 either downloaded from the NSI websites or available from the respective JMQs), or, in the case of Lithuania, from the population census for 2001 (Eurostat) and population register for 2006 (JMQ). In each case, the five-year time span between the observations enables tracking the earlier cohort in time (‘cohort-wise ageing’¹⁵), and comparing the resulting age structure with the initial one, as well as with the one actually observed five years later.

A comparison of various age structures of foreign-born population for the Nordic EU and Lithuania is presented in Figure 3. For the Nordic EU countries, the comparison reveals an interesting property: the actual age distribution observed in 2007 can be seen as a composition of: 2002 patterns for the youngest age groups (until ca. 20 years), ‘aged’ 2002 patterns for the oldest, much less mobile groups (over 40 years), and a middle part (20–35 years of age), exceeding both 2002 patterns (original and ‘aged’), due to the presence of highly age-selective immigration flows. In this middle part, the shares observed in 2007 are closer to those observed in 2002 in the same age groups than to the shares observed in 2002 in the same cohorts. The picture for Lithuania is a little bit different, but also here the actual structure observed in 2006 for the youngest age groups is closer to the original 2001 structure than to the ‘aged’ one. Unlike in the Nordic countries, for the middle age groups the 2006 structure is closer to the “aged” 2001. It is worth noting that the patterns observed for Lithuania are still to some extent a legacy of the Soviet times, with relatively high shares of foreign-born persons in older age groups, to a large extent concerning people born in other republics of the former USSR, whereas the youngest age groups exhibit entirely different characteristics, with much less foreign-born.

Such empirical features of the mentioned data examples (Nordic EU and Lithuania) are the main reason, while next to the cohort-wise propagation of structures, its period-wise equivalent has to be considered, especially for the net emigration countries.. The same rationale applies to the interpolation methods presented in Section 3.1.

3.6.2 Period-wise shares propagation

The period-wise shares propagation is the simplest of the three proposed shares propagation methods. It is expected that it might be superior to the cohort-wise shares propagation in the situations when population structures are highly influenced by short term migration flows of foreigners concentrating in selected age groups. In such cases, the breakdown of population by country of birth can be estimated using formula (7) and the country of birth structure from the from the previous year (or from the census). Formula (7) then becomes:

$${}_5P_g^n(x, t) = {}_5P_g^*(x, t) \cdot {}_5P_g^n(x, c) / {}_5P_g^*(x, c), \text{ if } t \text{ is the census year;} \quad (10a)$$

$${}_5P_g^n(x, t) = {}_5P_g^*(x, t) \cdot {}_5P_g^n(x, t-1) / {}_5P_g^*(x, t-1), \text{ for subsequent years,} \quad (10b)$$

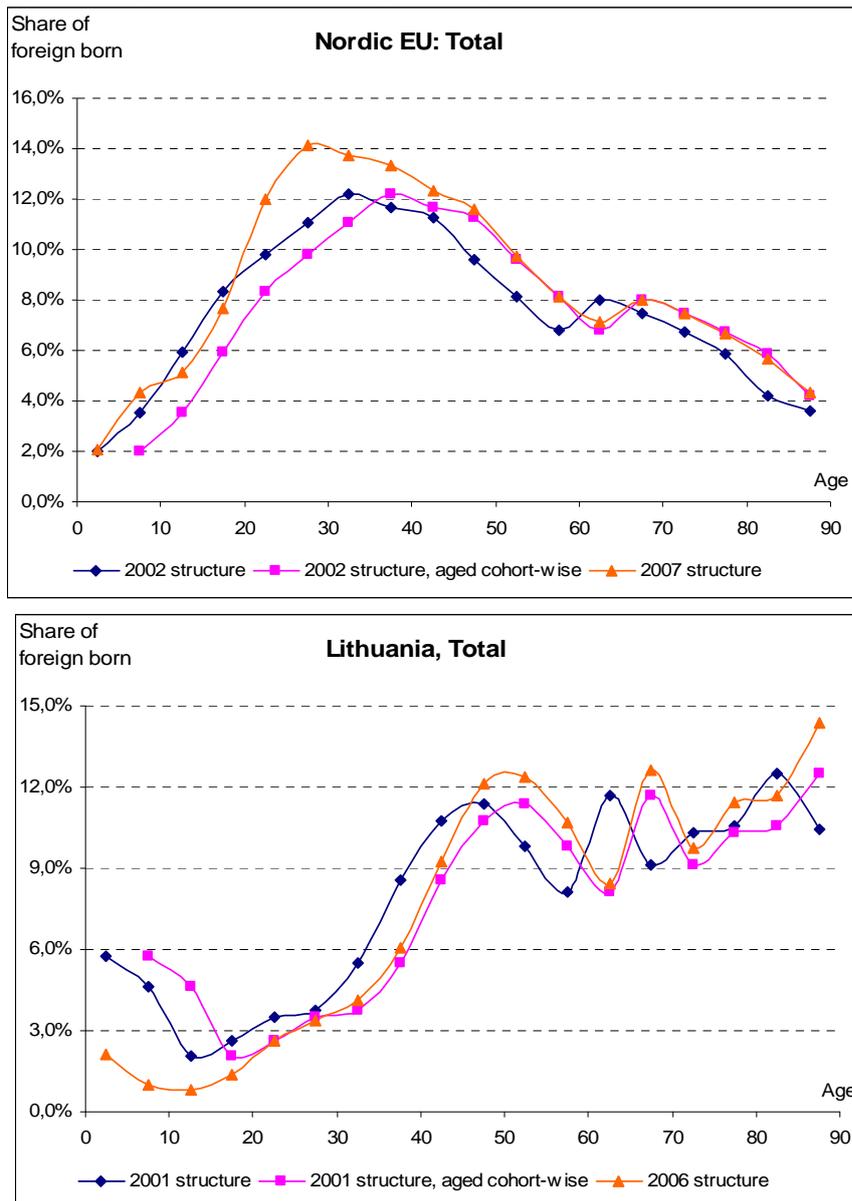
¹⁵ Cohort-wise aging corresponds in this case to the assumption that mortality of foreign-born persons is the same as for native-born and that either there is no net migration or net migration rate of foreign-born is the same as for native-born.

or simply:

$${}_5P_g^n(x, t) = {}_5P_g^*(x, t) \cdot {}_5P_g^n(x, c) / {}_5P_g^*(x, c), \text{ for all years.} \quad (10)$$

The estimates are automatically consistent with the aggregated population by sex and age, so no further adjustment is necessary. Unlike in the cohort-shares propagation and weighted shares method, the estimates are directly made for 5-year age groups.

Figure 3. Age structures of foreign-born population, Nordic EU and Lithuania



Source: own elaboration based on the Eurostat / NSI data

3.6.3 Cohort-wise shares propagation

For the cohort-wise propagation of weights, the proposed algorithm is as follows:

1. For the census population, apply the structure by country of birth, taken from each five-year age group, to the respective single-year age groups (i.e. from age group 0–4 to single ages 0, 1, ..., 4; from 5–9 to 5, 6, ..., 9 etc.). Let $w^n(x, c) = P^n(x, c) / P(x, c)$ denote the age-specific shares of country of birth group n in the census.

2. Further, set α as a fraction of the calendar year before the census date. It is implicitly assumed that the census population in single-year age groups can be divided between ‘older’ and ‘younger’ cohorts using the α and $(1-\alpha)$ partition.

3. For the census date, use the following formula to calculate the share of country of birth group n in the cohort that was aged x years on 1st January of the census year:

$$w^n(x+\alpha, c) = [(1-\alpha) P^n(x, c) + \alpha P^n(x+1, c)] / [(1-\alpha) P^*(x, c) + \alpha \cdot P^*(x+1, c)], \text{ for } x < x_{\max}; \quad (11a)$$

$$w^n(x_{\max}+\alpha, c) = P^n(x_{\max}, c) / P^*(x_{\max}, c). \quad (11b)$$

4. For the 1st January of the census year assume

$$w^n(x, t) = w^n(x+\alpha, c). \quad (12)$$

5. For the 1st January of the year following the census year ($t > c$), assume in turn:

$$w^n(x, t) = w^n(x-1+\alpha, c), \text{ for } 0 < x < x_{\max}; \quad (13a)$$

$$w^n(x_{\max}, t) = [P^n(x_{\max}-1, c) (1-\alpha) + P^n(x_{\max}, c)] / [P^*(x_{\max}-1, c) (1-\alpha) + P^*(x_{\max}, c)], \text{ and} \quad (13b)$$

$$w^n(0, t) = w^n(0, c) \text{ for the youngest age group.} \quad (13c)$$

6. For subsequent years calculate:

$$w^n(x, t) = w^n(x-1, t-1), \text{ for } x = 1, \dots, x_{\max}-1; \quad (14a)$$

$$w^n(x_{\max}, t) = [P^n(x_{\max}-1, t-1) + P^n(x_{\max}, t-1)] / [P^*(x_{\max}-1, t-1) + P^*(x_{\max}, t-1)]; \quad (14b)$$

$$w^n(0, t) = w^n(0, t-1). \quad (14c)$$

7. Calculate populations for all years using the above shares and total populations (available e.g. from DEMO), as

$$P^n(x, t) = P^*(x, t) w^n(x, t). \quad (15)$$

Finally, aggregate single-year age groups into five-year ones.

3.6.4 Mixed solutions

Given the empirical distributions of population by country of birth and age groups, such as the ones presented in Section 3.6.1, one option to combine the advantages of the two approaches would be to apply the period-wise propagation (10) for younger age groups ($x < x^0$), and the cohort-wise propagation (11) – (15) thereafter (for $x \geq x^0$). Selection of the division point, x^0 would be country-specific, although in principle it would most likely be below the most mobile age groups. For instance, concerning the examples presented in Section 3.6.1, for the Nordic EU in 2002–2007, the period-wise propagation could be applied up to $x^0 = 35$ years of age, while for Lithuania in 2001–2006 only up to $x^0 = 15$ years. In these cases the errors (differences) would be smaller while using the period-wise method than with the cohort-wise method, although, as seen in Figure 3, these errors would be still be relatively large in the case of Lithuania and the middle age groups in the Nordic countries.

Another theoretical possibility would be to perform a ‘weighted’ propagation of shares, with age-specific weights $\delta(x)$ given to the cohort-wise method and $1 - \delta(x)$ to the period-wise approach, for all

age groups except the youngest one. Using notation from the previous section, this would yield an iterative procedure for subsequent post-census years, $t > c$:¹⁶

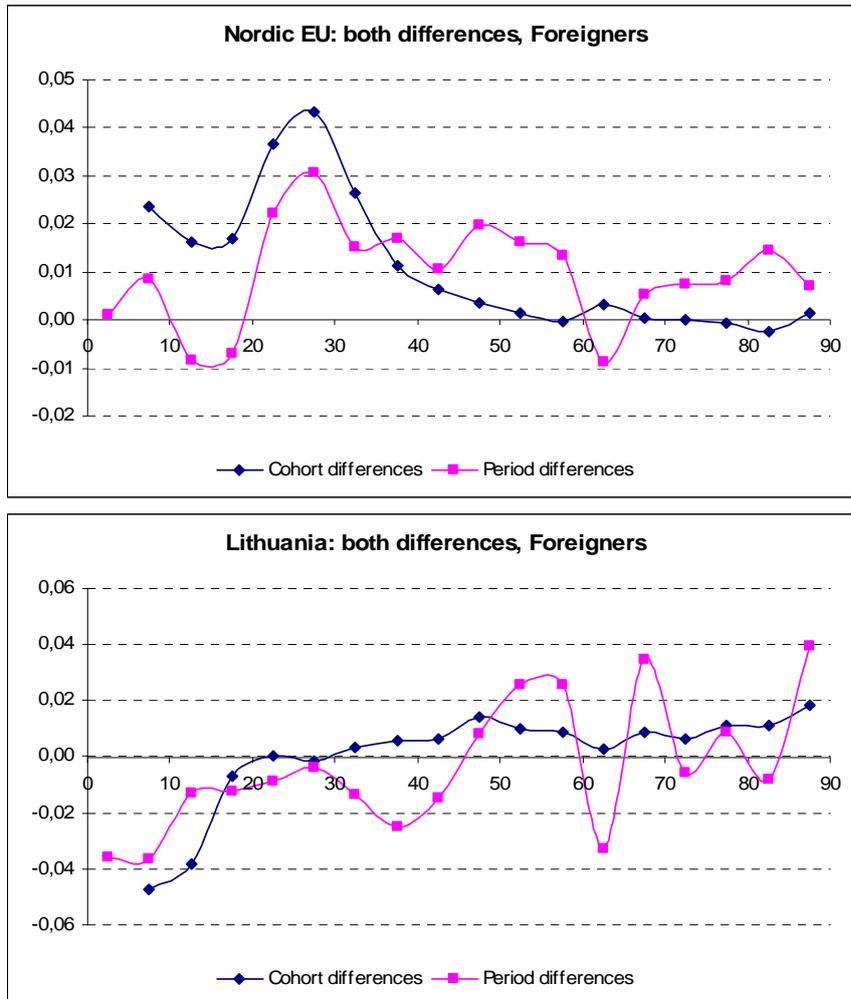
$$w^n(x, t) = \delta(x) \cdot w^n(x-1, t-1) + [1 - \delta(x)] \cdot w^n(x, t-1), \text{ and} \tag{16a}$$

$$w^n(0, t) = w^n(0, t-1). \tag{16b}$$

Ultimately, the procedure would end with an (iterative) proportional adjustment to all marginal totals, in order to ensure the internal consistency of the outcome.

The ‘weighted’ solution, however, although more general than the previous one, requires very strong judgemental assumptions about the shape of the weighting function, $\delta(x)$. Its estimation may be not straightforward: for the presented examples (Nordic EU and Lithuania), several attempts to propose various shapes of $\delta(x)$ have been undertaken, all without success. The distances between the observed patterns and the ‘propagated’ ones (under various metrics), and the ‘correction factors’ to be applied to the “old” shares, do not seem to have any universal, general form, even defined in very rough terms. For example, consider the differences between the observed and propagated patterns presented in Figure 4: the functions for Nordic EU (net immigration area) differ very much from the ones for Lithuania (net emigration country), not only with respect to the sign, but also to the overall shape.

Figure 4. Differences between various age structures: cohort-wise and period-wise approaches



Source: own elaboration based on the Eurostat / NSI data

¹⁶ The census year modifications, omitted here for the sake of transparency, would be analogous to the ones presented in Sections 3.6.2 and 3.6.3.

Considering each case separately, the differences derived using the cohort-wise method differ from the period-wise ones (Figure 4) and it is not clear how to estimate $\delta(x)$ for the countries for which no information from two points in time is available, even knowing, whether they are net migration ‘gainers’ or ‘losers’.

To conclude, the search for an ‘optimal’ functional form of $\delta(x)$ still remains a research challenge. For the purpose of MIMOSA, it is recommended to use a simplified solution proposed before, thus, assuming a step-wise $\delta(x) = 0$ for $x < x^0$ and $\delta(x) = 1$ for $x \geq x^0$.

3.7 Auxiliary methods

Among the auxiliary methods, the foremost one is the decomposition of the *Unknown* category wherever it appears (i.e., with respect to age, country of birth, or even sex, as in the case of Greece in 2005). The universal solution proposed in such cases is a proportional disaggregation: population belonging to the *Unknown* category is broken down proportionally to the existing, well defined categories (country of birth groups, age groups, etc.) and the resulting parts are attached to these categories. For example, if total population P consists of n well-defined groups P_1, \dots, P_n , and the *Unknown* category, P_{unk} , such that $P = \sum_i P_i + P_{unk}$, where $i = 1, \dots, n$, the following corrections apply:

$$P'_j = P_j + P_{unk} \cdot P_j / \sum_i P_i = P_j (1 + P_{unk} / \sum_i P_i), \text{ for all } j, \text{ with } i = 1, \dots, n. \quad (17)$$

If some elements of age structures are missing (e.g., tails of respective age distributions, or breakdown into five-year groups given the availability of broader ones), we may either use a structure from a different year or fit a mathematical function to available data. For example, we can assume that foreign population stocks are a double-exponential function of age, as originally proposed for the intensity of migration flows by A. Rogers and L. J. Castro (Rogers and Castro, 1981; Castro and Rogers, 1983). The number of foreign population aged x , $\phi(x)$, would then be given by the following equation:

$$\phi(x) = c + a_1 \cdot \exp(-\alpha_1 \cdot x) + a_2 \cdot \exp\{-\alpha_2 \cdot (x - \mu_2) - \exp[-\lambda_2 \cdot (x - \mu_2)]\}. \quad (18)$$

The parameters $c, a_1, \alpha_1, a_2, \alpha_2, \lambda_2$ and μ_2 can be estimated separately for each sex, for example using the ordinary least squares method (OLS) on the basis of the data for the available age groups (for example, below 65 years of age). Technically, the calculations can be done in a spreadsheet (e.g. MS Excel) using a solver-like tool, controlling for sensitivity to the choice of initial input values of the algorithm. Based on the obtained parameter estimates, formula (18) yields approximations of $\phi(x)$ for the remaining age groups. The last, open-ended group (85+) can be obtained by subtraction of all other figures from the total. To avoid negative numbers in the 85+ category, appropriate constraints should be set during the estimation procedure.

In either case, when adjustment to broader age groups is needed in order to ensure summation to respective totals (e.g. for functional age groups), it can be done via a proportional fitting method presented in Section 3.5.

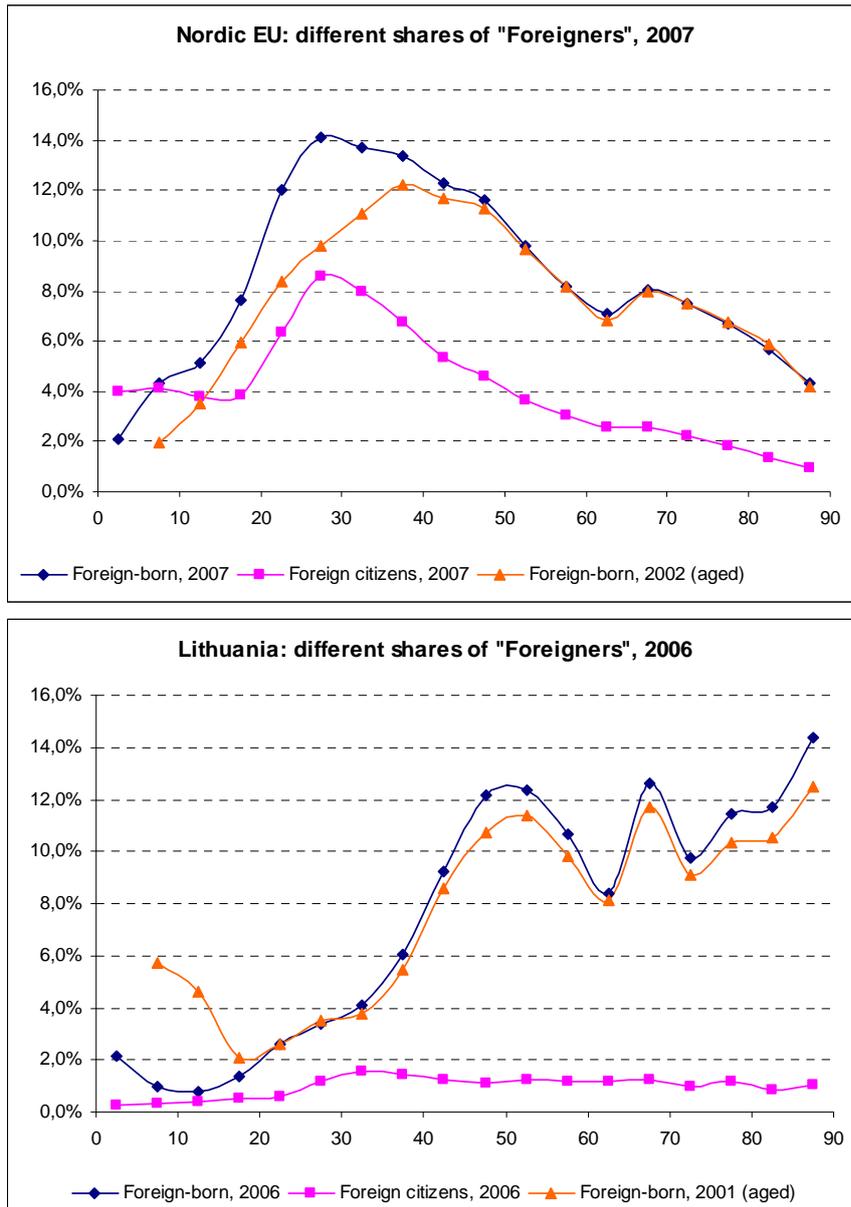
3.8 Can population by country of birth be estimated from the citizenship structure?

The last method considered with regard to the estimation of population by sex, age and country of birth groups was to apply structures by citizenship, available either from the original Eurostat / NSI data, or from earlier MIMOSA calculations. However, in general case, this idea might work only for these countries, for which there are (almost) no acquisitions of citizenship, and there have been no changes of borders in the last century. Otherwise, the structures by citizenship and country of birth

might differ much more than the structure by country of birth and the respective one recorded for example in the census five years earlier.

Respective empirical examples for the Nordic EU and for Lithuania are provided in Figure 5. Differences for the Nordic countries can be likely attributed to citizenship acquisitions, while for Lithuania – to the legacy of the Soviet times, coupled with a naturalisation of non-Lithuanian Soviet residents of Lithuania performed in 1989. To sum up, it is recommended that in a general case census-based information should have precedence before likely inferior solutions, such as using auxiliary structures (e.g. by citizenship) as proxies.

Figure 5. Sample age structures of foreign-born and foreign citizens, Nordic EU and Lithuania



Source: own elaboration based on the Eurostat / NSI data

4. Country-specific recommendations of estimation methods

In the light of the overview of data availability presented in Section 2 it is clear, that complete data on population by country of birth, sex and age group are not available in Eurostat for any of the 31 European countries covered by MIMOSA. Therefore, some estimations may be needed for all the countries, unless the data can be downloaded from the NSI websites or obtained by contacting the NSI.

In the current section, the availability of data, both in Eurostat and in the NSI, and the proposed estimation methodologies are presented for each country separately. In each case, before doing any calculations, it is important to try to get missing data and to perform data validation and internal consistency checks (the already spotted inconsistencies in the Eurostat data have been described in the relevant country sections). In addition to the use of the proposed estimation method, there might be also a need to deal with the *Unknown* category, usually by distributing it proportionally among the well specified categories, as described in Section 3.7.

In all cases where several methods could be alternatively applied, preference is given to more straightforward ones, and definitely to the ones having less judgemental elements, thus less potential sources of error. This approach conforms to the *Occam's razor* principle, stating that 'entities are not to be multiplied beyond necessity'¹⁷, which in this case means that the proposed models should not be more sophisticated than necessary, given various possible sources of error. For example, given complete data on stocks by age and country of birth group for two moments of time (e.g. from subsequent population censuses), if the data on migration flows and deaths are not available by country of birth and/or age, and require estimation, then it is recommended to calculate the intermediate values using cohort-wise or period-wise interpolation rather than projection. In the former case, the only source of possible error is the composition of population as such, whereas in the latter, judgemental assumptions on the relevant distributions of all components of the balance equation are likely to result in higher uncertainty of the ultimate results, which within the deterministic framework of the project is impossible to assess.

Subsections 4.1 through 4.31 present a detailed country-specific analysis of data availability and proposed estimation methods for all 31 countries under study. 27 countries of the European Union (as of 1st January 2007) are presented in alphabetic order, followed by four countries of the EFTA (Iceland, Liechtenstein, Norway and Switzerland).

¹⁷ After: "Ockham's razor", in: *Encyclopædia Britannica Online*, <http://www.britannica.com/eb/article-9056716>, accessed on 21st May 2007.

4.1 Austria

Availability of population stocks data

The JMQ data on population stocks by country of birth in Austria are complete only for 2007. The total population figure for 2007 in the JMQ is the same as in DEMO. For the years 2002–2006 the breakdown by country of birth is missing. The only data by country of birth, apart from those for 2007 are the data from the census conducted in January 2001. However, the Austrian NSI might also have data on population by country of birth, sex and age for 31.03.2006, because information on country of birth is kept in Austrian population register since 2006 (Statistics Austria, 2007) and, since 31.12.2003, stock data have been routinely produced at the end of each quarter (Lebhart 2005). Complete annual data by sex and age are available from DEMO.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	+	+	+	+	+	+
Total population by sex and age	+	+	+	+	+	+
Population by country of birth, sex and age	-	-	-	-	-	+

+ data were provided to Eurostat; - data were not provided to Eurostat.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	15.01.2001
Population by country of birth, sex and age	+

Proposed estimation methodology

Austrian data on population stock by country of birth, sex and age for 2007 do not need any estimation and they can be directly aggregated into the three country of birth categories of interest (*N*, *EU*, *nEU*). For 2006, it should be checked if the data could be obtained from the Austrian NSI (for 31.03.2006).

For 2002–2005 (or 2002–2006, if the data for 2006 cannot be obtained from the NSI), the cohort-wise interpolation method is recommended, based on the data from the Census and 2006 (2007) data. The estimates should be adjusted proportionally to the 2002–2006 population by sex and age available in DEMO.

4.2 Belgium

Availability of population stocks data

The JMQ data on population stock by country of birth, sex and age for Belgium have been delivered only for 2002 and 2003, but no aggregates have been provided for 2002 and some inconsistencies have been detected in the 2003 data.

The NSI has a complete set of data from the population register needed for the calculation of population stocks by sex, age and country of birth for all years 2002–2007, but no such statistics were published on the NSI website.

The census data on population by country of birth, sex and age (as of 1.10.2001), as well as annual data in DEMO on population by sex and age are complete.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	-	x	-	-	-	-
Total population by sex and age	-	x, -age	-	-	-	-
Population by country of birth, sex and age	-age	x, -age	-	-	-	-

+ data were provided to Eurostat; - data were not provided to Eurostat; **-age** no disaggregation by age; **x** problems detected in the data sent by the NSI.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date 1.10.2001
 Population by country of birth, sex and age +

Proposed estimation methodology

For 2002–2006, the aggregated data on population by country of birth (Belgium, EU27, outside EU27, Unknown), sex and 5-year age group have been prepared by Nicolas Perrin (UCL) using the Population Register – Statistics Belgium data. It is proposed to assign the persons with unknown country of birth to the category “Born outside EU27”.

It is recommended to contact the NSI for the required 2007 data. If the data cannot be obtained, data for 2006 and the shares propagation method can be used make the estimates for 2007.

4.3 Bulgaria

Availability of population stocks data

For Bulgaria, the annual data on population by country of birth, sex and age for 2002–2007 are not available. The only information comes from the census of 1st March 2001, provided in the 2005 JMQ, and from annual data on population by age and sex, available in the DEMO database. It should be noted that the annual population numbers are prepared by the NSI using the component method, based on the initial figures from the 2001 census as well as annual statistics on vital events and internal migration. International migration is not taken into account due to the lack (or poor accessibility) of data.

The 2001 data for Bulgaria available from the Eurostat on-line database in the table on population by country of birth, sex and age refer to the Census results (as of 1.02.2001).

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	-	-	-	-	-	-
Total population by sex and age	-	-	-	-	-	-
Population by country of birth, sex and age	-	-	-	-	-	-

- data were not provided to Eurostat.

Note: The table on population by country of birth, sex and age provided by Bulgaria in the JMQ in 2005 contains population data from the 2001 census.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	1.3.2001
Population by country of birth, sex and age	+

Proposed estimation methodology

As the only information on the breakdown by country of birth is from the census, only rough estimation is possible. Taking into account data availability, either cohort-wise or period-wise shares propagation based on census figures is recommended. These shares would be applied to the known annual figures on population by sex and age from DEMO. Should one wish to make a projection, then in order to be consistent with the NSI figures on total population, the projection equations should include births and deaths, but ignore international migration. Therefore, projection would require assumptions about the breakdown of deaths by country of birth. Under the assumption that this breakdown is the same as the breakdown of total population, the projection would simplify to the cohort-wise shares propagation.

4.4 Cyprus

Availability of population stocks data

Data on the population of Cyprus by country of birth, sex and age requested in the JMQ have been provided for 2002 only and in fact refer to the unadjusted census results¹⁸. According to THESIM, no annual data on population by country of birth are available and no such data were found on the NSI website. The only data that can be used for making the estimates are thus those from the census of 1st October 2001, as well as the stocks by sex and single-year age group available from DEMO.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	Census	-	-	-	-	-
Total population by sex and age	Census	-	-	-	-	-
Population by country of birth, sex and age	Census, -age	na	na	na	na	na

- data were not provided to Eurostat; **-age** no disaggregation by age; **Census** data for the Census reference date.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	1.10.2001
Population by country of birth, sex and age	+

Proposed estimation methodology

Considering the significant immigration to Cyprus (net migration above 20 persons per 1000 population in 2004) and taking into account that the only available information are the detailed data from the 2001 Census, we think that the reliable breakdown of total population of Cyprus by country of birth is not possible. However, some very rough estimates can be made, for example for the purpose of using them to produce rough estimates for the whole EU.

For making the rough estimates, either cohort-wise or period-wise shares propagation based on census figures is recommended. These shares would be applied to the known annual figures on population by sex and age from DEMO. Projection would require assumptions about the breakdown of deaths and migration by country of birth. Under the assumption that this breakdown is the same as the breakdown of total population, the projection would simplify to the cohort-wise shares propagation.

¹⁸ Only basic census results (population by sex, age, residence type and district) have been adjusted for under-enumeration after the post-enumeration survey. The census covered the Government controlled area of Cyprus.

4.5 Czech Republic

Availability of population stocks data

For the Czech Republic, the annual data on population by country of birth, sex and age for 2002–2007 are not available. The only information for making the estimations comes from the census of 1st March 2001 and from annual data on population by single-year age group and sex, available in the DEMO database. No additional data have been identified on the NSI website.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	-	+	+	+	+	-
Total population by sex and age	-	+	+	+	+	-
Population by country of birth, sex and age	-	-	-	-	-	-

+ data provided to Eurostat; - data were not provided to Eurostat.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	1.3.2001
Population by country of birth, sex and age	+

Proposed estimation methodology

As the only information on the breakdown by country of birth is from the census, only rough estimation is possible. Taking into account data availability, either cohort-wise or period-wise shares propagation based on census figures is recommended. These shares would be applied to the known annual figures on population by sex and age from DEMO. Projection would require assumptions about the breakdown of deaths and migration by country of birth. Under the assumption that this breakdown is the same as the breakdown of total population, the projection would simplify to the cohort-wise shares propagation.

4.6 Denmark

Availability of population stocks data

The JMQ data on population stocks in Denmark for the years 2005 and 2006 are complete, as indicated in the table below. The data are consistent between the JMQ and the DEMO database. The 2007 data available in the *migr_popctba* table of the Eurostat on-line database, originating from the JMQ 2007, refer in fact to year 2006. For the period 2002–2004, the JMQ tables contain complete data by country of birth and sex. The disaggregation by age was not requested by Eurostat at that time.

The complete data on men and women by country of birth and 5-year age group (up to 110+) for 2002–2004 and for 2007 (as well as for other years) can be downloaded from the NSI website <http://www.statbank.dk/BEF5>.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	+	+	+	+	+	-
Total population by sex and age	+	+	+	+	+	-
Population by country of birth, sex and age	-age	-age	-age	+	+	-

+ data were provided to Eurostat; - data were not provided to Eurostat; -age no disaggregation by age.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	1.1.2001
Population by country of birth, sex and age	+

Proposed estimation methodology

Danish data on population stock by country of birth, sex and age for 2005–2006 are available in the JMQ. Data for 2002–2004 and 2007 can be downloaded from the NSI website. The data do not need any estimation and can be directly aggregated into the three country of birth categories of interest (*N*, *EU*, *nEU*).

4.7 Estonia

Availability of population stocks data

Annual data on population by country of birth, sex and age are not available. The only information are data from the Census (as of 31st March 2000), and annual data on population by age and sex available from the DEMO database. In the census, two sets of population figures were produced: *de facto* and *de jure* (permanent), *de facto* population excluding persons temporarily absent and including those temporarily present in the country. Disaggregations by country of birth, delivered to Eurostat, were published for the *de jure* population.

Annual population numbers (by sex and age) are prepared by the NSI using the cohort-component method, based on the figures from the 2000 census and annual vital and internal migration statistics. International migration is not taken into account due to the poor quality of data.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	-	-	-	-	-	-
Total population by sex and age	-	-	-	-	-	-
Population by country of birth, sex and age	na	na	na	na	na	na

- data not provided to Eurostat; **na** data not available.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	31.3.2000
Population by country of birth, sex and age	+

Proposed estimation methodology

As the only information on the breakdown by country of birth is from the census, only rough estimation is possible. Taking into account data availability, either cohort-wise or period-wise shares propagation based on census figures is recommended. These shares would be applied to the known annual figures on population by sex and age from DEMO. Should one wish to make a projection, then in order to be consistent with the NSI figures on total population, the projection equations should include births and deaths, but ignore international migration. Therefore, projection would require assumptions about the breakdown of deaths by country of birth. Under the assumption that this breakdown is the same as the breakdown of total population, the projection would simplify to the cohort-wise shares propagation.

4.8 Finland

Availability of population stocks data

The JMQ data on population stocks in Finland for the years 2005–2007 are complete, as indicated in the table below. The data are consistent between the JMQ and the DEMO database. For the period 2002–2004, the JMQ tables contain complete data by country of birth and sex. The disaggregation by age was not requested by Eurostat at that time.

For 2002–2004, data are available from "Country of birth according to age and gender by region 1990–2007", from: http://pxweb2.stat.fi/database/StatFin/vrm/vaerak/vaerak_en.asp, though only up to the age group 75+. The remaining data should be available from the NSI upon request.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	+	+	+	+	+	+
Total population by sex and age	+	+	+	+	+	+
Population by country of birth, sex and age	-age	-age	-age	+	+	+

+ data were provided to Eurostat; -age no disaggregation by age.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	31.12.2000
Population by country of birth, sex and age	+

Proposed estimation methodology

Finish data on population stock by country of birth, sex and age for 2005–2007 do not need any estimation and they can be directly aggregated into the three country of birth categories of interest (*N*, *EU*, *nEU*).

For 2002–2004, the data on population by country of birth, sex and 5-year age group up to 75+ should be downloaded from the NSI website and aggregated as required. The missing data for the age groups 75–79, 80–84 and 85+ can be requested from the NSI. If data cannot be obtained, the cohort-wise interpolation method (based on the data from the Census and the 2005 data) is recommended to obtain an initial estimate of the population structure by country of birth, sex and age. It should be followed by proportional fitting to the known aggregated population in the age 75+ by country of birth and sex.

4.9 France

Availability of population stocks data

French data on population by country of birth, sex and age are available only from the last traditional population census¹⁹ (as of 8th March 1999). Annual estimates are available for population distributions by sex and age (available from DEMO), but annual data by country of birth were not published. There are some indications that such data could be available from the rolling census which was introduced in 2004. In particular, the mid-2004 data on the number of foreigners born abroad and the number of foreigners born in France were provided in the publication *Insee Première* n°1098, available from <http://www.insee.fr/fr/ffc/ipweb/ip1098/ip1098.html>)

The number of foreigners and French nationals is available from the census, as well as from the INSEE estimates for 1.01.2005, the latter available from

http://www.insee.fr/fr/ffc/chifcle_fiche.asp?ref_id=NATTEF02131&tab_id=339

A cross-tabulation of population by nationality (French and foreigners) and country of birth (born in France and abroad) is available from the census. These data are accessible at:

http://www.insee.fr/fr/ffc/chifcle_fiche.asp?ref_id=NATCCI02123&tab_id=426

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	-	-	-	-	-	-
Total population by sex and age	-	-	-	-	-	-
Population by country of birth, sex and age	-	-	-	-	-	-

- data were not provided to Eurostat.

Note: The tables on population by country of birth, sex and age provided by France in the JMQ in 2000–2005 contain population data from the 1999 Census.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	8.3.1999
Population by country of birth, sex and age	+

Proposed estimation methodology

It is recommended to make an attempt to obtain some estimates from the French NSI (INSEE). Such estimates should be available at least for mid-2004, based on the rolling census. If no data can be obtained, then the only some very rough estimates of the number of foreign born and native-born population can be made. It is proposed to do it separately for foreigners and French citizens. For foreigners, the published mid-2004 data on the number of foreigners born abroad and in France can be used to obtain rough estimates for 2005–2007. Estimates for 2002–2004 would be obtained by interpolation of data from census and data for mid-2004. For French citizens, the estimates would be based on the information from census combined with the data or estimates of the number of nationals. Projection is not recommended because available data on migration flows are very incomplete.

¹⁹ Information on population stocks provided in JMQ refers to the 1999 census. In 2004, a “rolling census” was introduced. Each year till 2008, the results of the rolling census will be used to adjust population estimates for the previous years (since 1999). The final estimates will be available after the full 5-year cycle (2004–2008).

4.10 Germany

Availability of population stocks data

No data on the breakdown of total population by country of birth have been identified, either in the JMQ or on the NSI website. The most recent census was in 1987, so census data would not be appropriate as a basis for making direct estimates for 2002–2007.

Information on the country of birth is collected in the Central register of Foreigners (*Ausländerzentralregister*). Statistics on foreigners by country of birth based on the register of foreigners are not published in the Destatis database, but some data on the number of foreigners born in Germany and born abroad can be found in other places in the internet. In particular, the following data have been identified:

31.12.2002: http://www.einbuergern.de/Pressespiegel/stat_ausl_bev.htm

31.12.2004: <http://www.migration-asyl.de/public1/auftr/home.nsf/url/AF37E55EBD7314B0C125703B002CB1A4?OpenDocument>

Most probably, the German NSI has data on foreigners disaggregated by country of birth (foreign-born, native-born) for other years as well. No data on the breakdown of German citizens by country of birth have been identified.

When making the estimates, it should be remembered that annual data on foreigners come from two different sources. The component method (*Bevölkerungsfortschreibung*), based on the last traditional German census of 25th May 1987²⁰, is used to produce annual figures on total population, total nationals and total foreigners, as well as nationals and foreigners by sex and age. Central Register on Foreigners (CRF) is used to produce disaggregation of foreigners by citizenship. The resulting total numbers and age structures of foreigners differ between both sources due to different definitions: “Generally, numbers of foreigners from the [CRF are] smaller than those from the current population statistics because the Register covers only foreigners living in Germany not just on a temporary basis. After the register was adjusted in the period 2000 to 2004, the differences have considerably increased. It is therefore not possible to directly compare the results for the foreign population from the two sources.” (<http://www.destatis.de/basis/e/bevoe/bevoetxt.htm>)

An auxiliary source of information is a recent publication from OECD “A profile of Immigrant Population in the 21st century” (OECD 2008). It contains some data on the population of Germany by country of birth, with information on native-born population based on the LFS 1998-2002 and information on foreign-born based on the Microcensus 2005.

²⁰ Data from the 1987 population census are not publicly available (sources checked: Eurostat, NSI’s webpage, UNSD’s Demographic Yearbook Special Census Topics). On 5th December 2001, a register-based test census was conducted (a sample of 1.2% of the population), but no detailed data from this survey are available. German data in the 2001-round census part of the Eurostat database (total population 81 465 344) are marked as not reliable and do not mention the source or reference date. The NSI was not able to provide any explanation, so these data should be probably removed from the database. In future, register-based surveys will replace traditional census. Only such data which can not be obtained from administrative registers will be collected directly from respondents.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	-	+	+	+	p	-
Total population by sex and age	-	+	+	+	p	-
Population by country of birth, sex and age	-	-	-	-	-	-

+ data were provided to Eurostat; - data not provided to Eurostat; p provisional data.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	25.05.1987
Population by country of birth, sex and age	- (see footnote 20)

Proposed estimation methodology

Taking into account very poor availability of data, probably either no or only very rough estimates of population by country of birth can be produced. In particular, the breakdown of foreign-born into born in the EU and born outside the EU is not possible.

It is recommended to begin by contacting the NSI. It should be checked if the data on the breakdown of German citizens by country of birth is available. Moreover, the annual data on the number of foreigners by country of birth, originating from the register of foreigners should be requested. It is assumed further that only the latter data can be obtained.

It is proposed to make the estimates for foreigners and German citizens separately. The methodology should be consistent with methods used by the NSI. Therefore, the total number of foreigners should be assumed to be the same as calculated by the NSI using the component method. The breakdown by country of birth can be based on the breakdown observed in the register of foreigners.

A reliable estimation of the breakdown of German citizens is not possible unless some data are obtained from the NSI. If no data are obtained, only very rough estimates can be made, for example by assuming that the share of foreign born and native born is the same as observed for nationals in some other country (for example Austrian in Austria or French in France).

The data published recently by OECD may be used as an auxiliary source of information.

4.11 Greece

Availability of population stocks data

Data on population of Greece by country of birth, sex and age are available only from the last population census (as of the 18th of March 2001). The annual estimates of population are available for population by sex and age (available from DEMO), but not by country of birth. No additional data have been identified on the NSI website.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	-	-	+	-	-	-
Total population by sex and age	-	-	+	-	-	-
Population by country of birth, sex and age	-	-	-	-	-	-

- data were not provided to Eurostat.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	18.3.2001
Population by country of birth, sex and age	+

Proposed estimation methodology

Considering the significant immigration to Greece and taking into account that the only available information are the detailed data from the 2001 Census, the results of which are very likely obscured by the 2001 and 2006 regularisations of irregular workers (cf. Cangiano, 2008), we think that the reliable breakdown of total population of Greece by country of birth is not possible. However, some very rough estimates can be made, for example for the purpose of using them to produce rough estimates for the whole EU.

For making the rough estimates, either cohort-wise or period-wise shares propagation based on census figures is recommended. These shares would be applied to the known annual figures on population by sex and age from DEMO. Projection is not recommended because data on migration flows are not available.

4.12 Hungary

Availability of population stocks data

Data on population of Hungary by country of birth, sex and age are available only from the last population census (as of the 1st of February 2001). The annual estimates of population are available for population by sex and age (available from DEMO), but not by country of birth. No additional data have been identified on the NSI website.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	-	-	+	+	+	-
Total population by sex and age	-	-	+	+	+	-
Population by country of birth, sex and age	-	-	-	-	-	-

+ data were provided to Eurostat; - data were not provided to Eurostat.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	1.02.2001
Population by country of birth, sex and age	+

Proposed estimation methodology

As the only information on the breakdown by country of birth is from the census, only rough estimation is possible. Taking into account data availability, either cohort-wise or period-wise shares propagation based on census figures is recommended. These shares would be applied to the known annual figures on population by sex and age from DEMO. Projection would require assumptions about the breakdown of deaths and migration by country of birth. Under the assumption that this breakdown is the same as the breakdown of total population, the projection would simplify to the cohort-wise shares propagation.

4.13 Ireland

Availability of population stocks data

The JMQ data on population stocks in Ireland for the period 2002–2006 are available, although very patchy, as indicated in the table below. First of all, the reference date is April, and not 1st January of a given year, and the figures are universally labelled as ‘provisional’. Secondly, the data are provided for selected categories of country of birth only: for persons born in the EU (EU15 for 2002–2005, EU25 for 2006), born in Ireland, the UK and the USA. Thirdly, the figures are disaggregated only by broad age groups: 0–14, 15–24, 25–44, 45–64 and 65+ years. The data provided in the JMQ originate from the Quarterly National Household Survey (QNHS). According to the documentation of the QNHS data for 2005 available at <http://issda.ucd.ie/documentation/cso/qnhs2005-q2codebook.pdf>, there should be information on the number of persons born in the EU accession states, which would allow to calculate the number of persons born in the EU25 countries.

In the DEMO domain of the NewCronos database there are detailed data on population by sex and single-year age groups for 1st January each, as required in MIMOSA. The difference in the reference dates might explain the observed discrepancy between the JMQ and DEMO aggregated figures on total population.

The census data on population by sex, country of birth and 5-year age group are available for both 2002 and 2006 census. The 2002 census data (as of the 28th of April 2002) can be downloaded from the Eurostat on-line database and from the NSI website, while the 2006 Census data (as of the 23rd of April 2006) from the NSI website only:

<http://beyond2020.cso.ie/Census/TableViewer/tableView.aspx?ReportId=5029> (Table 32)

http://www.cso.ie/census/census2006results/volume_4/vol_4_2006_complete.pdf (the report “Census 2006 Volume 4 – Usual Residence, Migration, Birthplaces, Nationalities”).

The census data refer to population usually resident and present on census night²¹. As concerns the 2006 census, the published data allow for the calculation of the number of persons born in the EU (probably EU25) by sex and age group, but information by age on the number of persons born in EU27 or in Romania and Bulgaria is missing. However, in another table there are data on the number of persons born in Romania by sex (Table 29, available at:

<http://beyond2020.cso.ie/Census/TableViewer/tableView.aspx?ReportId=5025>)

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	p, dref	p, dref	p, dref	p, dref	p, dref	-
Total population by sex and age	p, broad age, dref	p, broad age, dref	-			
Population by country of birth, sex and age	p, ±cob -age, dref	p, ±cob -age, dref	p, ±cob -age, dref	p, ±cob broad age, dref	p, ±cob broad age, dref	-

+ data provided to Eurostat; - data were not provided to Eurostat; **p** provisional data; **dref** reference date different than 1st January (= 1st April); **broad age** data disaggregated by broad age group; **±cob** data provided for a few country of birth categories.

²¹ It can be noted that in the censuses, some demographic characteristics have been also collected for persons usually resident but temporarily absent (for less than 3 months), but tables on population by country of birth have been published for population usually resident and present on census night. “From 2007 onwards a new concept of usual residence will be used, i.e. all persons usually resident and present in the State on census night *plus* absent persons who are usually resident in Ireland but are temporarily away from home *and outside the state* on census night.” (information from the NSO publication “Population and Migration Estimates April 2007”, available at <http://www.cso.ie/releasespublications/documents/population/current/popmig.pdf>).

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	28.04.2002, 23.04.2006
Population by country of birth, sex and age	2002: + 2006: -

Proposed estimation methodology

The estimation of population by country of birth, sex and age for Ireland requires combination of data from two Censuses, the QNHS data and data from DEMO. It is recommended to contact the NSI for the missing information, for example for the 2007 data and QNHS data on the number of persons born in the EU27. If no additional information can be obtained, the number of EU27 citizens may be estimated by applying the ratio of EU27/EU25 or EU27/EU15 calculated from the 2002 Census to the available data on the number of persons born in EU25 or EU15 countries. The initial estimates by sex and country of birth for 2003–2004 and by sex, country of birth and broad age for 2005–2006 can be done by interpolating the QNHS data (in order to move from mid-April to 1 January). The initial estimates by sex, country of birth and 5-year age group may be prepared by interpolation between two censuses. The two sets of estimates should be combined and then adjusted to the DEMO population by sex and age.

For 2002 and 2007 the estimates may be produced by, respectively, backward or forward shares propagation from the 2002 and 2006 Census

4.14 Italy

Availability of population stocks data

For Italy, the annual data on population by country of birth, sex and age for 2002–2007 are not available. The only information for making the estimations for 2002–2006 comes from the census of the 21st of October 2001 and from annual data on population by single-year age group and sex, available in the DEMO database.

Information on country of birth is collected in the Population Registers (*anagrafi*), but till recently was not used in population statistics. The number of foreigners born in Italy, by sex (but not by age) is available annually since 31.12.2006 (but not for earlier years). The figure for 31.12.2006 can be found at: http://demo.istat.it/str2006/index_e.html. These data are based on the information from the annual survey “Movement and calculation of foreign population (modello P.3), which is one of the surveys used by ISTAT to collect data from the Population Register offices in the Italian communities. The form used to collect the data on the demographic balance of foreign population for the years 2006 and 2007 are available at:

http://www.istat.it/strumenti/rispondenti/indagini/demosanitarie/2007/mod_p3.pdf and <http://www.istat.it/strumenti/rispondenti/indagini/demosanitarie/2008/p3.pdf>.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	-	-	-	+	+	-
Total population by sex and age	-	-	-	-	-	-
Population by country of birth, sex and age	-	-	-	-	-	-

+ data provided to Eurostat; - data were not provided to Eurostat.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	21.10.2001
Population by country of birth, sex and age	+

Proposed estimation methodology

Considering the significant immigration to Italy and taking into account that the only available information are (i) the detailed data from the 2001 Census, so before the significant regularisation that took place in 2002 and (ii) the number of foreigners born in Italy on 1.01.2007, we think that the reliable breakdown of total population of Italy by country of birth is not possible. In particular, there is no indication about the proportion of foreign-born among Italian nationals. However, some very rough estimates of population by country of birth can be made, for example for the purpose of using them to produce rough estimates for the whole EU.

When making the rough estimates, it is recommended to begin with the estimations for 2007, where some partial information exists, and then proceed to estimations for 2002–2006 using cohort-wise or period-wise interpolation between the census and 2007.

For 2007, the number of foreigners born abroad can be directly calculated as the difference between total foreigners and foreigners born in Italy published by the NSI. The proportion of those born in the

EU can be assumed to be either the same as the proportion of EU citizens among foreigners or as the proportion of foreigners born in the EU in the census, if the latter data are available.

The breakdown of Italian citizens by country of birth can be estimated only very roughly. First, the shares of native-born, foreign-born in EU and born outside EU should be estimated, for example by cohort-wise or period-wise shares propagation from the census (if available), or assuming the shares “borrowed” from another country. These shares would be applied to the total number of Italians available from the JMQ. Finally, the 2007 estimates for foreigners and Italians would be combined.

Projection would require assumptions about the breakdown of deaths and migration by country of birth. Under the assumption that this breakdown is the same as the breakdown of total population, the projection would simplify to the cohort-wise shares propagation.

4.15 Latvia

Availability of population stocks data

The JMQ data on population stocks by country of birth in Latvia for 2007 are complete, as indicated in the table below. The data are consistent between the JMQ and the DEMO database. For the period 2003–2006, the JMQ tables contain complete data by country of birth and sex. For 2002, data by country of birth were not provided. Complete data by country of birth, sex and age are available from the census, as of 31.3.2000. There are also complete annual data by sex and single-year age group (available in DEMO). No additional data have been identified on the NSI website.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	+	+	+	+	+	+
Total population by sex and age	+	+	+	+	+	+
Population by country of birth, sex and age	-	-age	-age	-age	-age	+

+ data were provided to Eurostat; - data were not provided to Eurostat; **-age** no disaggregation by age.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	31.3.2000
Population by country of birth, sex and age	+

Proposed estimation methodology

Latvian data on population stock by country of birth, sex and age for 2007 do not need any estimation and they can be directly aggregated into the three country of birth categories of interest (*N*, *EU*, *nEU*).

For the 2003–2006, the cohort-wise interpolation method (based on the data from the Census and the 2007 data) is recommended to obtain an initial estimate of the population structure by country of birth, sex and age. It should be followed by the iterative proportional fitting, with two faces of migration cube being known: the population by country of birth and sex face and population by sex and age face.

For 2002, it is proposed to use the cohort-wise interpolation based on census data and the estimates obtained for 2003. The initial estimates should be adjusted to the known population figures by sex and age.

4.16 Lithuania

Availability of population stocks data

The JMQ data on population stocks by country of birth in Lithuania for 2005–2007 are complete, as indicated in the table below. The data are consistent between the JMQ and the DEMO database. For 2004, the JMQ table contains complete data by country of birth and sex (the disaggregation by age was not requested by Eurostat at that time). For 2002 and 2003, data by country of birth were not provided. Complete data by country of birth, sex and age are available from the census, as of 6.04.2001. There are also complete annual data by sex and single-year age group (available in DEMO). No additional data have been identified on the NSI website.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	+	-	+	+	+	+
Total population by sex and age	-	-	-	+	+	+
Population by country of birth, sex and age	-	-	-age	+	+	+

+ data were provided to Eurostat; - data were not provided to Eurostat; -age no disaggregation by age.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	31.3.2000
Population by country of birth, sex and age	+

Proposed estimation methodology

Lithuanian data on population stock by country of birth, sex and age for 2005–2007 do not need any estimation and they can be directly aggregated into the three country of birth categories of interest (*N*, *EU*, *nEU*).

For the 2004, the cohort-wise interpolation method (based on the data from the Census and the 2005 data) is recommended to obtain an initial estimate of the population structure by country of birth, sex and age. It should be followed by the iterative proportional fitting, with two faces of migration cube being known: the population by country of birth and sex face and population by sex and age face.

For 2002 and 2003, it is proposed to use the cohort-wise interpolation based on census data and the estimates obtained for 2004. The initial estimates should be adjusted to the known population figures by sex and age.

4.17 Luxembourg

Availability of population stocks data

Data on population of Luxembourg by country of birth, sex and age are available only from the last population census (as of the 15th of February 2001). The annual estimates of population are available for population by sex and age (available from DEMO), but not by country of birth. No additional data have been identified on the NSI website.

It should be noted, that the NSI has recently revised its population figures. The data on total population and population by citizenship previously provided in the JMQ are not consistent with the revised figures and with the data available from DEMO.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	-	-	-	-	-	-
Total population by sex and age	-	-	-	-	-	-
Population by country of birth, sex and age	-	-	-	-	-	-

- data were not provided to Eurostat.

Note: The table on population by country of birth and sex provided by Luxembourg in the JMQ in 2002 contains population data from the 2001 Census.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	15.02.2001
Population by country of birth, sex and age	+

Proposed estimation methodology

As the only information on the breakdown by country of birth is from the census, only rough estimation is possible. Taking into account data availability, either cohort-wise or period-wise shares propagation based on census figures is recommended. These shares would be applied to the known annual figures on population by sex and age from DEMO.

Projection would require assumptions about the breakdown of deaths and migration by country of birth. The regularisation of status of irregular immigrants took place in Luxembourg in 2001, with about 2,900 applications and ca. 1,180 persons granted residence permits by the end of 2002 (Levinson, 2005 and <http://www.gouvernement.lu/dossiers/justice/sanspapiers>). It can be assumed that the latter figure, being a sum of $R^*(*, 2001) + R^*(*, 2002)$ are already included in the immigration totals: $I^*(*, 2001)$ and $I^*(*, 2002)$. This assumption may be justified by the fact that the annual population accounts published by the NSI do not include regularisation as a separate item, while immigration and emigration figures are the same as given in the JMQ.

4.18 Malta

Availability of population stocks data

For Malta, there are no annual data on population by country of birth, sex and age. Information about country of birth was collected during the censuses conducted in November 1995 and 2005. For 2005, the following data were published in the “Census Population and Housing Report 2005” (available at: http://www.nso.gov.mt/docs/Census2005_Voll.pdf): population by sex and country of birth (in Table 13) and population by age group and country of birth (in Table 14). Data were published for the most important countries and aggregates, in particular there are data for on the number of persons born in Malta and in the EU, but not in the EU27. The data by age were provided for 10-year age groups. The results from the 1995 Census include a table on population by country of birth, sex and age, probably with the similar groupings as 2005 data, however the contents of the table is not available on the Internet (for a list of tables with the results of the 1995 Census, available for purchase from the NSI, see: <http://www.nso.gov.mt/site/page.aspx?pageid=216>).

The annual data on population by sex and age are available in DEMO.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	-	-	-	-	-	-
Total population by sex and age	-	-	-	-	-	-
Population by country of birth, sex and age	-	-	-	-	-	-

+ data were provided to Eurostat; - data not provided to Eurostat.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	26.11.1995, 27.11.2005
Population by country of birth, sex and age	-

Proposed estimation methodology

It is recommended to contact the NSI for the 1995 Census results, and for more detailed (e.g. 5-year age groups) 2005 results. The estimates for 2002–2005 can be done by cohort-wise interpolation between two censuses, followed by proportional adjustment to the total population by sex and age available from DEMO. The exact procedure depends on what data can be received from the NSI. For 1st January 2006, the estimations can be done using the total population from DEMO and the shares by country of birth from the 2005 Census (as of 27 November 2005). For 2007, it is recommended to make the estimations using the cohort-wise shares propagation method.

4.19 Netherlands

Availability of population stocks data

The JMQ data on population stocks in the Netherlands for the years 2005–2007 are complete, as indicated in the table below. The data are consistent between the JMQ and the DEMO database. The 2005 data available in the *migr_popctb* table of the Eurostat on-line database (accessed on 18.04.2008) refer in fact to year 2006. For the period 2002–2004, the JMQ tables contain complete data by country of birth and sex. The disaggregation by age was not requested by Eurostat at that time. Complete data for 2002–2004 should be available from the Dutch NSI. The data available for download from the NSI website are disaggregated by 15-year age group (instead of required 5-year groups) and include data on the number of persons born in the Netherlands, born abroad and born in the EU on 1.01.1996–2007 (before using the data, it should be checked with the NSI how many EU countries were included in the EU category). These data are located at:

<http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=70648NED&D1=0-2&D2=0-3&D3=a&D4=0&D5=a&HDR=T,G1&STB=G2,G3,G4&VW=T>.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	+	+	+	+	+	+
Total population by sex and age	+	+	+	+	+	+
Population by country of birth, sex and age	-age	-age	-age	+	+	+

+ data were provided to Eurostat; -age no disaggregation by age.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	1.01.2001
Population by country of birth, sex and age	+

Proposed estimation methodology

Dutch data on population stock by country of birth, sex and age for 2005–2007 do not need any estimation and they can be directly aggregated into the three country of birth categories of interest (*N*, *EU*, *nEU*).

For the 2002–2004 data, it is recommended to contact the NSI. If the data cannot be obtained, the cohort-wise interpolation method (based on the data from the Census and the 2005 data) is recommended to obtain an initial estimate of the population structure by country of birth, sex and age. It should be followed by the iterative proportional fitting to the known information for 15-year age groups (downloaded from the NSI website).

4.20 Poland

Availability of population stocks data

Data on population of Poland by country of birth, sex and age are available only from the last population census²² (as of the 20th of May 2002). The annual estimates of population²³ are available for population by sex and age (available from DEMO), but not by country of birth. No additional data have been identified on the NSI website.

The Polish census figures include information about 583.5 thousand persons with unknown country of birth, a majority of whom were most likely persons born in Poland, temporarily absent at the time of the Census, and enumerated from the population register (in such cases no detail information on several variables, including country of birth, was available).

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	Census	-	-	+	+	+
Total population by sex and age	Census	-	-	+	+	+
Population by country of birth, sex and age	Census	na	na	na	na	na

+ data provided to Eurostat; - data not provided to Eurostat; **Census** data for the Census reference date; **na** data not available

Note: The table on population by country of birth and sex provided by Poland in the JMQ in 2003 and 2004 contain population data from the 2002 Census.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	20.05.2002
Population by country of birth, sex and age	+

Proposed estimation methodology

As the only information on the breakdown by country of birth is from the census, only rough estimation is possible. Taking into account data availability, either cohort-wise or period-wise shares propagation based on census figures is recommended. These shares would be applied to the known annual figures on population by sex and age from DEMO. Projection would require assumptions about the breakdown of deaths and migration by country of birth. Under the assumption that this breakdown is the same as the breakdown of total population, the projection would simplify to the cohort-wise shares propagation.

²² The census information chiefly refers to permanent population. Population estimates have been produced also for alternative definitions (e.g. usual residents), but very few statistics based on the latter have been published. An attempt could be made to make rough estimates of annual population by sex, age and citizenship taking the available census data on usually resident population as the starting point, but the obtained estimates would be inconsistent with official annual population statistics which refer to persons permanently resident in Poland.

²³ The annual population stocks are obtained with the component method, using information on population on 1st January of the previous year, registered births, deaths, and international migration flows. They do not include foreigners who stay in Poland more than a year but on a limited duration permit (temporary residence), nor exclude nationals living abroad for more than a year who keep their permanent address in Poland.

4.21 Portugal

Availability of population stocks data

Data on population of Portugal by country of birth, sex and age are available only from the last population census (as of the 12th of March 2001). The annual estimates are available for population by sex and age (available from DEMO), but not by country of birth. The annual figures are calculated by the NSI using the component method, but it is not clear how net migration is estimated, but not all data used for the estimation are available: data on immigration flows provided to Eurostat concern residence permits issued to foreigners, while the estimates of net migration take into account also the information on return migration of nationals obtained from the Labour force Survey.

Information from the census refers to usually resident population, but some figures on present population have also been produced. However, in 2001–2003, the most recent regularisation programme for formerly-irregular immigrants was held in Portugal (Cangiano, 2008), which casts some doubts on the use of the census figures for the purpose of the estimation of 2002–2007 stocks, although the census data are the only ones available in the required breakdown by sex, age and country of birth (see tables below).

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	-	-	-	-	-	-
Total population by sex and age	-	-	-	-	-	-
Population by country of birth, sex and age	-	na	na	na	-	na

- data not provided to Eurostat; **na** data not available.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	12.03.2001
Population by country of birth, sex and age	+

Proposed estimation methodology

Considering the significant immigration to Portugal and taking into account that the only available information are the detailed data from the 2001 Census, so before the significant regularisation that took place in 2001–2003, we think that the reliable breakdown of total population of Portugal by country of birth is not possible. However, some very rough estimates can be made, for example for the purpose of using them to produce rough estimates for the whole EU.

For making such rough estimates, either cohort-wise or period-wise shares propagation based on census figures is recommended. These shares would be applied to the known annual figures on population by sex and age from DEMO. Projection would require assumptions about the breakdown of deaths and migration by country of birth. Under the assumption that this breakdown is the same as the breakdown of total population, the projection would simplify to the cohort-wise shares propagation.

4.22 Romania

Availability of population stocks data

For Romania, the data on population stocks by country of birth, age and sex are available for the 2002 census (as of 18th March), as well as for 1st January 2006 and 2007. For 2004, the data provided in the JMQ are disaggregated by country of birth and sex but not by age. For 1st January 2002, 2003 and 2005 the breakdown by country of birth was not provided. The 2005 data available in the *migr_popctb* table of the Eurostat on-line database (accessed on 18.04.2008) refer in fact to year 2006.

It is worth noting that the census data disseminated in publications and delivered to Eurostat refer to usually resident population, while another set of figures based on the census results concerns permanently resident population. The latter is used as a starting point for estimating annual population of Romania in the years following the census²⁴. In DEMO, data on population stocks by sex and age for 2002–2007 are complete.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	Census	-	+	+	+	+
Total population by sex and age	Census	-	+	+	+	+
Population by country of birth, sex and age	Census	-	-age	-	+	+

+ data provided to Eurostat; - data not provided to Eurostat; **Census** data for the Census reference date; **na** data not available

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	18.03.2002
Population by country of birth, sex and age	+

Proposed estimation methodology

Romanian data on population stock by country of birth, sex and age for 2006 and 2007 do not need any estimation and they can be directly aggregated into the three country of birth categories of interest (*N*, *EU*, *nEU*).

For 2002–2004 the estimations should take into account the observation that the numbers of foreign-born provided by the NSI for 1st January 2004 are identical to the respective numbers from the 2002 census. Therefore, the numbers by sex and age for the two foreign-born groups on 1st January 2002, 2003 and 2004 can be assumed the same as in the census 2002. The sex and age specific numbers on native-born can be calculated as a difference between the corresponding numbers on total population from DEMO and the above assumed numbers on foreign-born.

For 2005, the cohort-wise interpolation method (based on the 2006 data and the estimated figures for 2004) is recommended to obtain an initial estimate of the population by country of birth, sex and age. It should be adjusted to the known population figures by sex and age.

²⁴ In the cohort-component method, migration data used to produce the yearly population estimates consist of immigration data on foreigners who received permanent residence status and data on permanent emigration of nationals. Immigration of nationals and emigration of foreigners are not taken into account due to the lack of data.

4.23 Slovakia

Availability of population stocks data

In Slovakia, complete data on population by country of birth, sex and age are available from the census only (as of 26th of May 2001). Data on population by country of birth and sex but not disaggregated by age were also provided in the JMQ for 2005, however the numbers included 274,354 persons with unknown country of birth. The JMQ data on population by country of birth and sex for 2004, 2006 and 2007 concern foreigners only (no breakdown by age for 2004). Annual data on population by sex and age are complete and available in DEMO. The 2005 figure concerning total population of Slovakia published in the *migr_popctb* table of the Eurostat on-line database (accessed on 18.04.2008) is not consistent with DEMO, because it refers to the total number of foreigners only. Moreover, it refers to 2006, not 2005.

There is also a problem with the consistency of definitions: the numbers on total population by sex and age refer to permanent residents of Slovakia only, calculated using the component method departing from the population census. Data on foreigners come from the register of foreigners and refer to foreigners with long-term residence permits (temporary or permanent).

Statistics on stocks of foreigners (residence permit data) are available from the website of the Border and Aliens Police: www.minv.sk/uhcp/new (Yearbooks 2001–2007 of the Border and Aliens Police). However, no breakdown by country of birth is provided.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	-	-	-	+	+	+
Total population by sex and age	-	-	-	+	+	+
Population by country of birth, sex and age	-	-	for, -age	-age, i	for	for

+ data were provided to Eurostat; - data not provided to Eurostat; **-age** no disaggregation by age; **for** data provided for foreigners only; **i** inconsistencies in data. Note: In JMQ2006, wrong reference date is given: the data concern 1st January 2006, not 2005 (source: Infostat).

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Reference date	26.05.2001
Population by country of birth, sex and age	+

Proposed estimation methodology

Taking into account the availability of data, only rough estimation is possible. It is recommended to adopt the follow procedure. First, make the estimates for 2004–2007, separately for nationals and foreigners. In order to do this, estimate the breakdown of foreigners by country of birth for 2005 by interpolating data for 2004 and 2006. Then, estimate the breakdown of nationals in 2005, using the estimates for foreigners and the data concerning total population available in the JMQ. Estimate the shares for nationals: for 2002–2004 by interpolating between the census and 2005, and for 2006 and 2007 by propagating shares from 2005. Apply the shares to the estimates of the number of foreigners and nationals obtained in an earlier stage of the MIMOSA project.

Due to inconsistencies in flow and stock data (mixture of data referring to permanent or long term migration), as well as to the need to break down deaths and migration events data by country of birth, projection is not recommended as the estimation method.

4.24 Slovenia

Availability of population stocks data

The JMQ data on population stocks in Slovenia for the years 2005–2007 are complete, as indicated in the table below. The data are consistent between the JMQ and the DEMO database. The 2005 data available in the *migr_popctb* table of the Eurostat on-line database (accessed on 18.04.2008) refer in fact to year 2006. For the period 2002–2004, the JMQ tables contain complete data by country of birth and sex. The disaggregation by age was not requested by Eurostat at that time. Complete data for 2002–2004 should be available from the Slovenian NSI. Data for 2002 and 2003, although with incomplete sets of countries, may be downloaded the NSI website:

data for 2002: <http://www.stat.si/doc/pub/rr798-2003/3/T03-12-01.htm>

data for 2003: <http://www.stat.si/doc/pub/rr816-2004/p-3/t03-09-02.xls>

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	+	+	+	+	+	+
Total population by sex and age	+	+	+	+	+	+
Population by country of birth, sex and age	-age	-age	-age	+	+	+

+ data were provided to Eurostat; -age no disaggregation by age.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	31.03.2002
Population by country of birth, sex and age	+

Proposed estimation methodology

Slovenian data on population stock by country of birth, sex and age for 2005–2007 do not need any estimation and they can be directly aggregated into the three country of birth categories of interest (*N*, *EU*, *nEU*).

For the 2002–2004 data, it is recommended to contact the NSI. If the data cannot be obtained, the data downloaded from the NSI website may be used for 2002 and 2003, supplemented with the data from the JMQs and from the Census (to fill the small gaps in the data). For 2004, the cohort-wise interpolation method (based on the data for 2003 and 2005) is recommended to obtain an initial estimate of the population structure by country of birth, sex and age. It should be followed by the iterative proportional fitting, with two faces of migration cube being known: the population by country of birth and sex face and population by sex and age face.

4.25 Spain

Availability of population stocks data

A summary of the availability of the JMQ data on population of Spain by country of birth, sex and age is given in the table below. The data for the period 2005–2007 are complete and consistent with the data in DEMO. No data were provided in the JMQ for 2003, while 2002 and 2004 data are disaggregated by country of birth and sex but not by age. 2002 and 2004 figures are not consistent with DEMO. The 2002 population stock data provided in the JMQ and available from the *migr_popctab* table of the Eurostat on-line database originate directly from the population register (*Padrón*), and – unlike the 2005–2007 data – were not adjusted to the Spanish “Population Now-Cast” figures (the latter is considered by INE to provide the best estimates of the population). The 2005 data available in the *migr_popctb* table of the Eurostat on-line database (accessed on 18.04.2008) refer in fact to year 2006.

Register-based data on population by country of birth may be found on the INE website at the following locations. Population by country of birth, sex and year:

<http://www.ine.es/jaxi/tabla.do?path=/t20/e245/p08/11/&file=01006.px&type=pcaxis&L=1>

The table on foreign population by age, sex and country of birth for 2002:

<http://www.ine.es/jaxi/menu.do?type=pcaxis&path=/t20/e245/p04//a2002&file=pcaxis>

Various population data are also published in *Statistical yearbooks*, available at: http://www.ine.es/en/prodyser/pubweb/anuarios_mnu_en.htm.

For example, 2004 data foreign population by country of birth and age group (register-based) are available in the 2005 yearbook:

http://www.ine.es/prodyser/pubweb/anuario05/anu05_2demografia.pdf

The methodology of producing annual estimates of population by sex and age is described in: http://www.ine.es/en/daco/daco43/epoba/metodo_en.pdf. The results of estimation may be found at:

<http://www1.ine.es/jaxi/menu.do?type=db&divi=EPOB&his=0>

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	+	-	p	+	+	+
Total population by sex and age	+	-	p	+	+	+
Population by country of birth, sex and age	-age	-age	p, -age	+	+	+

+ data were provided to Eurostat; -age no disaggregation by age; p provisional data.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Reference date	1.01.2001
Population by country of birth, sex and age	+

Proposed estimation methodology

Spanish data on population stock by country of birth, sex and age for 2005–2007 do not need any estimation and they can be directly aggregated into the three country of birth categories of interest (*N*, *EU*, *nEU*).

For 2002–2004, it is recommended to contact the NSI for the missing information (population register data on population by country of birth, sex and age may be available). If no additional information can be obtained, other than that available on the NSI website, then the following procedure is proposed. The estimation of population by country of birth can be done by proportional decomposition of total population from DEMO, using information on the shares of the three country of birth categories available from the population register (from the NSI website). The initial estimates of population by country of birth, sex and age would be done by interpolation between census and 2005. Iterative proportional fitting method can be used to adjust these initial estimates to the population by sex and age known from DEMO and population by country of birth as estimated earlier. An alternative method, requiring more work, would be to make the estimates by age separately for foreigners and nationals. In that case, information on the age structure of foreigners published on the NSI website would be used.

4.26 Sweden

Availability of population stocks data

The JMQ data on population stocks in Sweden for the years 2005–2007 are complete, as indicated in the table below. The data are consistent between the JMQ and the DEMO database. For the period 2002–2004, the JMQ tables contain complete data by country of birth and sex. The disaggregation by age was not requested by Eurostat at that time. Complete data on foreign-born by country of birth, sex and single-year age group for 2000–2007 can be downloaded from the Swedish NSI website:

<http://www.ssd.scb.se/databaser/makro/produkt.asp?produktid=BE0101&lang=2>

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	+	+	+	+	+	+
Total population by sex and age	+	+	+	+	+	+
Population by country of birth, sex and age	-age	-age	-age	+	+	+

+ data were provided to Eurostat; **-age** no disaggregation by age.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Reference date 1.01.2001
 Population by country of birth, sex and age +
 Note: Data were taken from population register

Proposed estimation methodology

Swedish data on population stock by country of birth, sex and age for 2005–2007 do not need any estimation and they can be directly aggregated into the three country of birth categories of interest (*N*, *EU*, *nEU*).

For 2002–2004, data on foreign-born should be downloaded from the NSI website and aggregated to 5-year age groups and required country of birth groups. The number of persons born in Sweden would be calculated as a difference between the figures on total population from DEMO and the above calculated numbers for foreign-born.

4.27 United Kingdom

Availability of population stocks data

The annual statistics on population by country of birth, sex and age, provided in the JMQs, are prepared using the results of the Labour Force Survey conducted in four consecutive quarters of the year. In early 2008, revised data for 2004–2008 were sent to Eurostat. Data by country of birth are not complete: data are marked as unavailable for these countries of birth for which the data were considered unreliable (due to a low number of persons born in these countries interviewed in the survey). In particular, information on the number of persons born in several EU27 countries is missing. The JMQ data were also provided for 2003, however without the breakdown by age (age was not requested in the JMQ at that time). No data were provided for 2002 and 2007.

LFS data on the working-age population (16–64 for men and 16–59 for women) by country of birth (UK, EU25, outside EU-25) for the 2nd quarter of 2002–2006 were published in the ONS report “Labour market summary for the UK household population by country of birth” available at:

<http://www.statistics.gov.uk/articles/nojournal/Birthcountry.pdf>

Complete data on population by country of birth, sex and age are available from the population census, as of 29.04.2001. The annual data on population by sex and age in DEMO are also complete.

Data on population stocks are not consistent between the sources (DEMO and LFS data) and the difference cannot be explained by the different reference dates. For example, the total population on 1 January according to DEMO was 59.4 million in 2003 and 59.7 in 2004, while the figure from the LFS given in the JMQ for 2003 was 58.4 million.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	-	dref	dref	dref	dref	-
Total population by sex and age	-	dref	dref	dref	dref	-
Population by country of birth, sex and age	-	±cob, -age, dref	±cob, dref	±cob, dref	±cob, dref	-

- data not provided to Eurostat; **-age** no disaggregation by age; **±cob** data are missing for some country of birth categories; **dref** reference date different than 1st January (here: an average over four quarters);.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	29.04.2001
Population by country of birth, sex and age	+

Proposed estimation methodology

It is recommended to contact the NSI for 2002, 2003 and 2007 data. If no additional data cannot be obtained, the following procedure for obtaining rough estimates is proposed.

- For 2005 and 2006: disaggregation of population from DEMO using the average sex- and age-specific shares by country of birth calculated from 2004 and 2005 JMQ data (when making the estimates for 1.01.2005) and from 2005 and 2006 JMQ data (when making the estimates for 2006).
- For 2007: disaggregation of population from DEMO using the shares by country of birth taken from the LFS data for 2006.

- For 2004: Initial step – estimations for mid-2003 by applying the age structures from mid-2004 to the mid-2003 population by country of birth from the JMQ. Final step - disaggregation of 2004 population from DEMO using the average sex- and age-specific shares by country of birth calculated from 2003 and 2004 JMQ data (or estimates).
- For 2003: Initial step – as for 2004. Final step - disaggregation of 2003 population from DEMO using the sex- and age-specific shares by country of birth obtained in the first step.
- For 2002 – disaggregation of 2002 population from DEMO using the average sex- and age-specific shares by country of birth calculated from census data and estimates for 2003.

4.28 Iceland

Availability of population stocks data

The JMQ population stocks data by country of birth and sex for Iceland are available only for the years 2002 and 2003, as indicated in the table below. For that period, the data are consistent between the JMQ and the DEMO database. In Iceland, censuses are not conducted since the National Registry was established in the 1950s and the information in the “Census 2001 round” part of the Eurostat database is register-based and concerns population stocks as of 1st January 2001.

The NSI website contains complete yearly information on the register-based population of Iceland, distributed jointly by country of birth, sex and single-year age group, under the following address:

<http://www.statice.is/Statistics/Population/Citizenship-and-country-of-birth>.

The data are downloadable in the Excel, html or xml format, or as delimited text files.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	+	+	-	-	-	-
Total population by sex and age	+	+	-	-	-	-
Population by country of birth, sex and age	-age	-age	-	-	-	-

+ data provided to Eurostat; - data not provided to Eurostat.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Reference date 31.12.2000
 Population by country of birth, sex and age +

Proposed estimation methodology

All 2002–2007 data should be downloaded from the NSI website. No estimations are required and the data should be directly aggregated into the three citizenship categories of interest (*N*, *EU*, *nEU*). For some years, persons born in Slovenia might be included in the “former Yugoslav” aggregate and therefore counted as born outside EU, but the number of such persons marginal.

4.29 Liechtenstein

Availability of population stocks data

Liechtenstein did not provide the Joint Migration Questionnaire data to Eurostat. The only available information includes yearly population by sex and age in the DEMO domain of NewCronos, as well as the complete census information on population stocks by country of birth, sex and age as of 5th December 2000.

Annual information on population by country of birth might be available from the NSI, (data from the population register). However, no such data are published on the websites where other population figures are available (http://www.llv.li/pdf-llv-avw-statistik-bevoelkerung_31.12.2005).

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	-	-	-	-	-	-
Total population by sex and age	-	-	-	-	-	-
Population by country of birth, sex and age	-	-	-	-	-	-

- data not provided to Eurostat.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Reference date	5.12.2000
Population by country of birth, sex and age	+

Proposed estimation methodology

It is recommended to contact the NSI for all 2002–2007 data. If the data cannot be obtained then only rough estimation is possible. Taking into account data availability either cohort-wise or period-wise shares propagation based on census figures is recommended. These shares would be applied to the known annual figures on population by sex and age from DEMO. The projection method is not a good option as it would require assumptions about the breakdown of deaths and migration by country of birth. Under the assumption that this breakdown is the same as the breakdown of total population, the projection would simplify to the cohort-wise shares propagation.

4.30 Norway

Availability of population stocks data

The JMQ data on population stocks in Norway for the years 2005 and 2006 are complete, as indicated in the table below. The data are consistent between the JMQ and the DEMO database. For the period 2002–2004, the JMQ tables contain complete data by country of birth and sex. The disaggregation by age was not requested by Eurostat at that time. The complete data for 2002–2004 and for 2007 should be available from the Norwegian NSI. On the NSI website there are numerous articles containing various (incomplete) data on population by country of birth for various years.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	+	+	+	+	+	-
Total population by sex and age	+	+	+	+	+	-
Population by country of birth, sex and age	-age	-age	-age	+	+	-

+ data were provided to Eurostat; - data were not provided to Eurostat; -age no disaggregation by age.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	3.11.2001
Population by country of birth, sex and age	+

Proposed estimation methodology

Norwegian data on population stock by country of birth, sex and age for 2005–2006 do not need any estimation and they can be directly aggregated into the three country of birth categories of interest (*N*, *EU*, *nEU*).

For the 2002–2004 data, it is recommended to contact the NSI. If the data cannot be obtained, the cohort-wise interpolation method (based on the data from the Census and the 2005 data) is recommended to obtain an initial estimate of the population structure by country of birth, sex and age. The final estimates will be obtained by iterative proportional fitting to the marginal totals available in the JMQ.

For 2007, the complete data should be available from the NSI. If the data cannot be obtained then either cohort-wise shares propagation or proportional decomposition using the country of birth structure of given age group should be used. These estimates should be combined with the data available from the NSI website and DEMO.

4.31 Switzerland

Availability of population stocks data

In Switzerland, the annual data on population by country of birth, sex and age for 2002–2007 are not available. The only information for making the estimations comes from the census of the 5th of December 2000 and from annual data on population by single-year age group and sex, available in the DEMO database.

Some data on population by country of birth may be available from the Swiss Labour Force Survey (*Schweizerische Arbeitskräfteerhebung*, SAKE), however such data are not available from the NSI website.

Availability of the 2002–2007 data in the Joint Migration Questionnaire

Population stocks	2002	2003	2004	2005	2006	2007
Total population	-	-	-	-	-	-
Total population by sex and age	-	-	-	-	-	-
Population by country of birth, sex and age	na	na	na	na	na	na

- data not provided to Eurostat; **na** data not available.

Availability of Census data on population by country of birth, sex and age in the Eurostat database

Census date	5.12..2000
Population by country of birth, sex and age	+

Proposed estimation methodology

As the only information on the breakdown by country of birth is from the census, only rough estimation is possible. Taking into account data availability, either cohort-wise or period-wise shares propagation based on census figures is recommended. These shares would be applied to the known annual figures on population by sex and age from DEMO. Projection would require assumptions about the breakdown of deaths and migration by country of birth. Under the assumption that this breakdown is the same as the breakdown of total population, the projection would simplify to the cohort-wise shares propagation.

5. Summary and conclusion

As it can be seen from the country-specific review of problems with data on population stocks by country of birth, sex and age group, there is no universal solution for estimating the missing pieces of information in the 31 countries under study. Nevertheless, depending on the availability of data at hand, either in Eurostat / JMQ, or in the National Statistical Institutes, several alternative estimation procedures can be proposed. These procedures are roughly summarised in Table 5.

In brief, four group of countries may be distinguished:

1. All data can be obtained by combining data from Eurostat and the NSI. Some estimations may be required only when there are *Unknown* categories, otherwise the data may be directly aggregated, after performing data validation. Countries belonging to this group: Belgium (possibly with the exception of 2007), Denmark, Sweden and Iceland (4 countries).
2. Data on population by country of birth, sex and age are available from the Census and some later year(s). It is recommended to perform the estimation for the years in between using interpolation (cohort-wise or period-wise) combined with proportional fitting to population by sex and age from DEMO. Countries belonging to this group: Austria, Finland, Latvia, Lithuania, the Netherlands, Romania, Slovenia, Spain, Norway (9 countries).
3. Annual data by country of birth are not available and the only source of data on population by country of birth, sex and age is the census. Rough estimates of the missing data may be obtained using cohort-wise or period-wise shares propagation from the Census, combined with the data on population by sex and age from DEMO. Countries in this group: Bulgaria, Cyprus, Czech Republic, Estonia, Greece, Hungary, Luxembourg, Liechtenstein (?), Poland, Portugal and Switzerland (11 countries). The estimates obtained using this method are considered by us to be less reliable than those obtained by interpolation. Moreover, in the countries where there was significant regularisation or net migration was very large, only some very rough estimates can be made (for example for the purpose of using them to produce rough estimates for the whole EU).
4. Incomplete data are available for some years in addition to the census data (except Germany, where census data are missing). The estimation method is country-specific, depending on the availability of data. Countries belonging to this group: France, Germany, Ireland, Italy, Malta, Slovakia and the United Kingdom (7 countries).

The current report constitutes a departure point for the application of the methods proposed above to fill the gaps in the Eurostat data on 31 European populations broken down by broad categories of country of birth, sex and five-year age group. The formulated recommendations are aimed to serve as a reference for the subsequent stage of the current research project, consisting in the estimation of population stocks as of 1st January, for the countries with incomplete information. It has to be noted that the suggestions made in the current report may be subject to changes in the implementation phase, should new data be discovered, or upon methodological advancements made within the project.

Table 5. Summary table: availability of data on population by country of birth and proposed estimation methods

Country	Data availability ²⁵	Proposed estimation methods
Austria	2007 in the JMQ	Interpolation between the Census and 2007, followed by proportional adjustment to DEMO
Belgium	2002–2006 OK 2007 might be available from the NSI	Direct aggregation (2007: shares propagation from 2006)
Bulgaria	No annual data	Shares propagation from the Census
Cyprus	No annual data	<i>Shares propagation from the Census</i>
Czech Republic	No annual data	Shares propagation from the Census
Denmark	2005–2006 in the JMQ 2002–2004 and 2007 from the NSI website	2005–2006: direct aggregation 2002–2004, 2007: download from the NSI website and aggregation
Estonia	No annual data	Shares propagation from the Census
Finland	2005–2007 in the JMQ 2002–2004 on the NSI website up to 75+	2005–2007: Direct aggregation 2002–2004: download from the NSI website and aggregation; missing age groups: interpolation + proportional fitting
France	No annual data. Some data from the “rolling census” might be available (e.g. the number of foreigners born abroad and born in France is available for mid-2004)	<i>Estimations separately for foreigners and French citizens. Foreigners: interpolation and shares propagation using the data from the Census and mid-2004 data. Nationals: based on the data from the Census (or additional information to be obtained from the NSI)</i>
Germany	No annual data in the JMQ. Two source of population statistics. Data for foreigners by country of birth probably available from the Register of foreigners (2002 and 2004 – published in the internet). No census data (last traditional census in 1987). Some OECD data from the LFS and Microcensus.	Estimations separately for foreigners and German citizens. Foreigners: total by sex and age from the NSI (component method), breakdown by country of birth – based on the population register (data or estimates from 2002 and 2004). Nationals: request to the NSI or use shares from another country. OECD data may also be used
Greece	No annual data	<i>Shares propagation from the Census</i>
Hungary	No annual data	Shares propagation from the Census
Ireland	2002–2006: QNHS (mid-April) data in the JMQ, for selected countries of birth and broad age groups Census data for 2002 (Eurostat) and 2006 (NSI website)	2003–2006: Combination of estimates obtained through interpolation between censuses and interpolation of QNHS data, followed by the adjustment to DEMO. 2002 and 2007: shares propagation from 2002 and 2006 Census.
Italy	No annual data. The total number of foreigners born in Italy known for 2007 only	<i>2007: estimates separately for foreigners and Italian nationals. 2002–2006: interpolation between the Census and 2007</i>
Latvia	2007 in the JMQ 2003–2005 in the JMQ but without age No JMQ data for 2002	2007: direct aggregation. 2003–2006: interpolation between the census and 2007, followed by IPF to the available JMQ data. 2002: interpolation between the Census and 2003, followed by proportional adjustment
Lithuania	2005–2007 in the JMQ 2004 in the JMQ but without age No JMQ data for 2002–2003.	2005–2007: direct aggregation. 2004: interpolation between the census and 2005, followed by IPF to the available JMQ data. 2002–2003: interpolation between the Census and 2004, followed by proportional adjustment
Luxembourg	No annual data	Shares propagation from the Census

Note: Only rough estimates can be made in a number of countries. The estimates that might be most problematic in this respect are marked in italics.

²⁵ Unless noted otherwise, complete census data are available for the year 2001 or 2002. 2002–2007 data on population by sex and age are complete and available in DEMO for all the countries.

Table 5. Summary table (continued)

Malta	No annual data. Censuses in 1995 and 2005. No census data in Eurostat. Selected census data by country of birth published on the NSI website (for 2005) or in the census report (1995).	Request to the NSI for the census results. 2002–2005: interpolation between two censuses followed by proportional adjustment to DEMO. 2006; estimates based on shares by country of birth from 2005 census (directly or by cohort-wise or period-wise shares propagation)
Netherlands	2005–2007 in the JMQ 2002–2004 in the JMQ but without age 2002–2004 data on the NSI website by 15-year age groups	2005–2007: direct aggregation. 2002–2004: request to the NSI or interpolation between the census and 2005, followed by IPF to the data available from the NSI website.
Poland	No annual data	Shares propagation from the Census
Portugal	No annual data	<i>Shares propagation from the Census</i>
Romania	2006–2007 in the JMQ 2004 in the JMQ but without age No JMQ data for 2002–2003.	2006–2007: direct aggregation. 2002–2004: number of foreign-born as in the census and 2004 JMQ data. 2005: interpolation between 2004 and 2006, followed by proportional adjustment
Slovakia	2006–2007 data on foreigners in the JMQ 2005 – data in the JMQ but without age 2004 – data on foreigners in the JMQ but without age	Estimates separately for Slovak citizens and foreigners using a combination of methods and sources (see Section 4.23)
Slovenia	2005–2007 in the JMQ 2002 and 2004 in the JMQ but without age Incomplete (but useful) data for 2002–2003 from the NSI website	2005–2007: direct aggregation. 2002–2004: request to the NSI or: 2002–2003- estimates using data from the website, supplemented by information from the JMQs and from the Census 2004: interpolation between 2003 and 2005, followed by IPF to the available JMQ data.
Spain	Two types of population statistics 2005–2007 in the JMQ (Now-Cast/DEMO data) 2002–2004 in the JMQ but without age (provisional for 2004, data from <i>Padrón</i> for 2002) <i>Padrón</i> data on population by country of birth and sex available at the NSI website	2005–2007: direct aggregation. 2002–2004: request to the NSI or: Breakdown of the total population from DEMO by country of birth using the shares from <i>Padrón</i> . Interpolation between Census and 2005 to obtain initial estimates by country of birth, sex and age. Adjustment to marginal totals (by age and sex from DEMO and by country of birth and sex as estimated earlier)
Sweden	2005–2007 in the JMQ 2002–2004 in the JMQ but without age 2002–2004 data on foreign-born from the NSI website	2005–2007: direct aggregation. 2002–2004: download and aggregate data for foreign-born. Calculate “Born in Sweden” as a difference
United Kingdom	2003–2006: LFS data (4 quarters average) in the JMQ, country of birth not complete, no age for 2003 no JMQ data for 2002 and 2007	Request to the NSI for the 2002, 2003 and 2007 data and/or: proportional decomposition of population from DEMO using the shares from the JMQ/LFS data (when estimating 2003–2007) or average shares from the Census and 2003 (when estimating 2002). See Section 4.27
Iceland	No JMQ data; 2002–2007 data on the NSI website	Download from the NSI website and aggregation
Liechtenstein	No JMQ data, but the NSI may have information from the population register	2002–2007: check the availability of data with the NSI, or: shares propagation from the Census
Norway	2005–2006 in the JMQ 2002–2004 in the JMQ but without age No JMQ data for 2007 Various pieces of information on the NSI website	2005–2006: direct aggregation 2007 – JMQ data were probably delayed, check with Eurostat 2002–2004: request to the NSI, or: interpolation between the census and 2005, followed by IPF to the data available from the JMQs
Switzerland	No annual data	Shares propagation from the Census

Note: Only rough estimates can be made in a number of countries. The estimates that might be most problematic in this respect are marked in italics.

References

- Bishop, Y. M. M., Fienberg, E. F., and Holland, P. W. (1975). *Discrete multivariate analysis*. Cambridge, MA: MIT Press.
- Calot, G., and Sardon, J.-P. (2003). *Methodology for the calculation of Eurostat's demographic indicators*. Eurostat Working Papers and Studies, Population and Social Conditions 3/2003/F/n° 26. Luxembourg: Eurostat.
- Cangiano, A. (2008). Foreign Migrants in Southern European Countries: Evaluation of Recent Data. In: J. Raymer, F. Willekens (eds.), *International Migration in Europe: Data, Models and Estimates*. Chichester: John Wiley, pp. 89–114.
- Cantisani, G., and Greco, V. (2006). Country Report: Cyprus. In: M. Poulain, N. Perrin and A. Singleton (eds.), *THESIM: Towards Harmonised European Statistics on International Migration*, Louvain-la-Neuve: Presses Universitaires de Louvain, pp. 491–498.
- Castro, L. J., and Rogers, A. (1983). Patterns of Family Migration: Two Methodological Approaches. *Environment and Planning A*, 15 (2): 237–254.
- Krūma, K. (2007). Lithuanian nationality: Trump card to independence and its current challenges. In: R. Bauböck, B. Perchinig and W. Sievers (eds), *Citizenship Policies in the New Europe*. Amsterdam: Amsterdam University Press, pp. 89–109.
- Lebhardt, G. (2005) Concepts, Methods and definitions. Austria's migration statistics in the new population register. [in:] *28th CEIES seminar: Migration statistics – Social and economic impacts with respect to the labour market*, Eurostat, pp. 47–54.
- Levinson, A. (2005). *The Regularisation of Unauthorized Migrants: Literature Survey and Country Case Studies. Regularisation programmes in Luxembourg*. Oxford: Centre on Migration, Policy and Society, University of Oxford.
- Norman, P. (1999). Putting Iterative Proportional Fitting on the Researcher's Desk. Working Paper 99/03. Leeds: School of Geography, University of Leeds.
- NSI [National Statistical Institute] (2006). *2005 population and demographic processes*. Sofia: National Statistical Institute of the Republic of Bulgaria.
- OECD (2008). *A profile of Immigrant Populations in the 21st Century*. OECD Publishing.
- ONS [Office for National Statistics] (2004). *Population Estimates. A short guide to population estimates*. London: Office of National Statistics.
- Perrin, N., and Poulain, M. (2006). Country Report: Luxembourg. In: M. Poulain, N. Perrin and A. Singleton (eds.), *THESIM:...* [op. cit.], pp. 519–527.
- Poulain, M., Kupiszewska, D., Perrin, N., and Gheorghiu, D. (2007). Romania. Report on the situation of international migration statistics. Unpublished manuscript.
- Rees, P. (1994). Estimating and projecting the populations of urban communities. *Environment and Planning A*, 26 (11): 1671–1697.
- Rogers, A., and Castro, L. J. (1981). *Model Migration Schedules*. IIASA Report RR-81-30. Laxenburg: International Institute for Applied Systems Analysis.
- Rowland, D. T. (2006). *Demographic methods and concepts*. Oxford: Oxford University Press.
- Shryock, H. S., Siegel, J. S., and Associates (1993). Interpolation: Selected General Methods. In: D. J. Bogue, E. E. Arriaga, D. L. Anderton, and G. W. Rumsey (eds.), *Readings in Population Research Methodology. Volume 1: Basic Tools*. Chicago, IL: Social Development Center / UN Population Fund, pp. 5-48–5-72.
- Statistics Austria (2007) Standard-Dokumentation Metainformationen (Definitionen, Erläuterungen, Methoden, Qualität) zur Statistik des Bevölkerungsstandes.
www.statistik.at/web_de/wcmsprod/groups/gd/documents/stdok/029355.pdf

- Trimikliniotis, N. (2007). Nationality and citizenship in Cyprus since 1945: Communal citizenship, gendered nationality and the adventures of a post-colonial subject in a divided country. In: R. Bauböck, B. Perchinig and W. Sievers (eds), *Citizenship Policies in the New Europe*. Amsterdam: Amsterdam University Press, pp. 263–291.
- Willekens, F. (1977). *The Recovery of Detailed Migration Patterns from Aggregate Data: An Entropy Maximizing Approach*. IIASA Report RR-81-30. Laxenburg: International Institute for Applied Systems Analysis.
- Willekens, F. (1982). Multidimensional Population Analysis with Incomplete Data. In: K. C. Land and A. Rogers (eds.), *Multidimensional Mathematical Demography*. New York: Academic Press, pp. 43–111.
- Willekens F. J., Pór A, and Raquillet R, (1981). Entropy, multiproportional, and quadratic techniques for inferring patterns of migration from aggregate data. In: A. Rogers (ed.), *Advances in Multiregional Demography*. Research Report RR-81-6. Laxenburg: International Institute for Applied Systems Analysis, pp. 83–124.
- Wilmoth, J. R., Andreev, K., Jdanov, D., Gleit, D. A., with the assistance of C. Boe, M. Bubenheim, D. Philipov, V. Shkolnikov and P. Vachon (2005). *Methods Protocol for the Human Mortality Database, version 4*. Berkeley: Department of Demography, University of California. Available from: <http://www.mortality.org/Public/Docs/MethodsProtocol.pdf> (17th May 2007).
-