Estimating the foreign-born population on a current basis

Georges Lemaitre and Cécile Thoreau

1

Introduction

For many OECD countries, there are no reliable statistics of the foreign-born population available on a current basis. Although many labour force surveys now identify the foreign-born population in their questionnaires, the estimates produced from this source are problematical. Households of foreign-born persons, particularly recent arrivals, are subject to higher non-response rates than other households and this makes it difficult to estimate year-to-year changes in the foreign-born population, which necessarily require the presence in the sample of the groups responsible for the change. In addition, even without non-response problems which could in principle be adjusted for, the magnitude of annual change in the foreign-born population is relatively small. Because labour force surveys are sample surveys, the estimates of such change would be subject to high sampling variability.

In the 2000 round of population censuses, virtually all OECD countries identified the foreign-born population. For a certain number of these countries, this is the only reliable source of data on the size of this population. National data sources, in particular population registers, may have detailed information on the foreign population, on migration flows, on naturalisations, on births of persons of foreign nationality, etc., but these sources do not always contain information on the country of birth of residents or no longer identify separately foreign-born persons who have acquired the nationality of the host country. The concept of <u>nationality</u> is the one that prevails in national statistics of the immigrant population and naturalised foreigners tend to disappear from the statistics.

With the OECD's compilation of the census-based database on the foreign-born population (see http://www.oecd.org/dataoecd/27/5/33868740.pdf), comparable statistics on the immigrant population have become available for the first time for almost all OECD countries. It would clearly be desirable to be able to update these statistics on a current basis, so as to constitute a time series. This note describes a few relatively simple methods that have been used to update the census figures in countries where current estimates do not exist, as well as an evaluation of the methods using data for countries for which current data on the foreign-born exist, but have been ignored in order to mimic the estimation procedure for countries for which new estimates are being produced. The statistics so produced are then compared to the official national series to assess their accuracy.

Two methods of estimation have been used; which one is used in a given country depends on data availability. The first imitates the standard demographic component method for updating population estimates, whereas the second estimates a parameter for combined death and outflow rates using past data and then extrapolates into the future using the most recent parameter estimate.

The component method

Methodology

We begin with two demographic identities which relate the size of the foreign and foreign-born populations, respectively, at time t_1 to that at time t_0 through the components of change in these populations:

$$S(t_1) = S(t_0) + \{I(t_0, t_1) - O(t_0, t_1)\} + \{B(t_0, t_1) - D(t_0\ t_1)\} - \{N(t_0, t_1) - L(t_0, t_1)\} \quad \text{[foreign population]}$$

$$S'(t_1) = S'(t_0) + \{I'(t_0, t_1) - O'(t_0, t_1)\} - D'(t_0\ t_1)\} \quad \text{[foreign-born population]}$$

The primes (') indicate that the quantities apply to the foreign-born population. S are the stocks, I and O the inflows and outflows, B and D births and deaths, N naturalisations and L loss or abandonment of nationality by nationals. "(t_0, t_1)" indicates that the variable in question is measured over the period from t_0 to t_1.

Dropping the time variables from the flows for simplicity, assuming that L is negligible, which is generally the case, and rearranging, we obtain for the foreign-population identity:

$$S(t_1) - S(t_0) – B + N = (I – O) – D$$

Note that the right-hand side of this identity has the form of measured change in the foreign-born population (that is, net migration flows minus deaths), except that it applies to the foreign population. To obtain the actual corresponding components for the foreign-born population, one would have to carry out the following operations to the figure on net flows minus deaths of foreigners (the right-hand side of the equation):

- add net flows of foreign-born nationals;
- subtract net flows of native-born foreigners;
- add deaths of native-born foreigners;
- subtract deaths of foreign-born nationals.

That is, $(I' – O') – D' = (I – O) – D + \{I(FBNA) – O(FBNA)\} – \{I(NBFO) – O(NBFO)\}$
$$- \{D(FBNA) – D(NBFO)\}$$

where FBNA and NBFO refer to foreign-born nationals and native-born foreigners, respectively. Generally, information on the quantities within brackets are not available in national statistics. For the purposes of estimation, we assume that the difference between the net flows of foreign-born nationals and those of native-born foreigners is small relative to the net flows of foreigners and can be ignored to a first approximation. There are other assumptions one could make here, for example, that the net migration rates of foreign-born nationals and native-born foreigners are the same as those of nationals and foreigners, respectively. As we will see, this assumption does not appear to be so strongly violated that fairly reasonable estimates cannot be produced.[1]

Thus we have:

$$S'(t_1) - S'(t_0) = I' – O' – D' \approx S(t_1) - S(t_0) – B + N – \{D(FBNA) – D(NBFO)\}. \quad (*)$$

The stocks, naturalisations and births of foreigners on the right-hand side are generally available for countries having population registers, but the death statistics within brackets may not be. However, the number of deaths of foreigners is either generally available or can be calculated residually as:

$$D = I – O – N + B – \{S(t_1) - S(t_0)\}.$$

To obtain deaths for the quantity within brackets above, it has been assumed that age-specific mortality rates for foreign-born nationals and native-born foreigners are the same as for the population at large. The

---

[1]   There is one exception, namely Germany, for which inflows of foreign-born nationals are substantial. The reason is that this group includes so-called "ethnic Germans" (Aussiedler), who are foreign-born persons who traditionally have been granted German nationality by right upon entry into the country and who do not enter into standard statistics of inflows of foreigners. For Germany, the inflows of this group have been added to the right-hand side of (*) above.

population by age group for foreigners, foreign-born nationals and native-born foreigners, respectively, is then estimated from the labour force survey and the quantity within brackets itself estimated as follows:

$$\{D(FBNA) - D(NBFO)\} \approx D * \{D(FBNA) - D(NBFO)\}_L/D_L,$$

where the subscript L indicates that the quantities in question have been estimated using labour force survey population estimates by age group and age-specific mortality rates.

This yields finally:

$$S'(t_1) - S'(t_0) \approx S(t_1) - S(t_0) - B + N - D\{D(FBNA) - D(NBFO)\}_L/D_L .$$

If we now let $S'(t_0)$ be the actual observed data for the foreign-born population at the census date t_0, we have a method for estimating $S'(t_1)$, where $t_1 - t_0$ can be either positive or negative.

Evaluation

An evaluation was carried out for three countries for which the necessary data for the foreign population were available and which also produce current statistics on the foreign-born population. The length of the series estimated for the purpose of evaluation was limited by the availability of data on foreign births and deaths. The results are summarised in Table 1 and Charts 1a and 1b. Note that the countries concerned (Denmark, the Netherlands and Norway) vary considerably in the percentage of the foreign-born who are foreigners, from 35% for the Netherlands to 52% for Norway and 60% for Denmark. The method has been assessed by comparing the difference between the estimated value and the official statistic relative to the total official change in the foreign-born population since the 2000 reference date. Since the estimation method essentially estimates only the change in the foreign-born population since the base-year date, the agreement between the official and estimated series is assessed on the basis of a comparison of the change rather than of the total stock figures.

The results can be summarized as follows:
- For the Netherlands and Norway, the percent difference between the estimate of change in the foreign-born population since the base year and that registered by official statistics is generally less than 10 percent and yields estimates for the foreign-born stocks that follow closely the official statistics.
- The estimated stocks of the foreign born are not quite so good for Denmark, compared to official statistics. The estimates tend to be systematically higher than the official statistics, generally by about 15%, but much more so in the period just prior to the base year (2000). The absolute difference in the estimates of change varies between about 2 and 8 thousand for a stock estimate in the 250-300 thousand range..

The charts show that the estimates pick up the trend in the foreign-born population very well for Netherlands and Norway, with differences relative to official statistics of change of less than 20 per cent. The results for Denmark are less satisfactory in recent years, showing an increasing difference relative to official figures since the year 2000. The difference does not yet exceed 20 per cent, however.

On the basis of this evaluation, it was considered that the estimates for other countries would generally be of sufficient quality to publish, especially in view of the alternative, which would involve waiting until the next decennial census for an updated figure on the foreign-born population. When the latter does become available, it will be possible to update the series to ensure that the estimate at the end of the period

coincides with the census figure. The countries for which estimates were produced (from 1995 to 2004) using the component method were Belgium, the Czech Republic (starting in 1998), Germany, Luxembourg, Portugal and Switzerland (see Table3).

Parametric method

Methodology

The starting point of this method is, as in the previous case, the demographic equality relating the figure for the foreign-born population at time t_1 to that at time t_0 plus the components of change in this population over the period (t_0,t_1):

$$S'(t_1) = S'(t_0) + \{I'(t_0, t_1) - O'(t_0, t_1)\} - D'(t_0\ t_1)$$

Here S' represents the stock of foreign-born persons and I', O' and D' the inflows, outflows and deaths, respectively, of foreign-born persons. This method is used for countries where there are no data on deaths or outflows of foreign-born persons.

One first approximates the inflows of the foreign-born by the inflows of foreigners, which amounts to assuming that the adjustment made to the inflows of foreigners to obtain those of the foreign-born (namely adding in the inflows of foreign-born nationals less those of native-born foreigners) is negligible relative to the total inflows of foreigners. One then assumes that the outflows and the deaths of foreign-born persons during a given year are a fraction α of the stock of the foreign born in the previous year, where α. is assumed to be constant over the entire estimation period (between two consecutive censuses):

$$O'(t_0, t_1) + D'(t_0\ t_1) = \alpha\ S'(t_0)$$

This leads to the general formula $S'(t+1) = S'(t) + I'(t,t+1) - \alpha\ S'(t)$. The foreign-born figures are estimated starting at time t=0 (census figure) and proceeding recursively, with α being chosen to ensure agreement with the subsequent census figure for the foreign-born population. The value for α is then used to extrapolate the process into the future. This method thus requires data on the foreign-born population for two consecutive censuses.[2]

.

Evaluation

An evaluation was carried out using data for four countries that have annual data on the foreign-born population, namely Australia, Denmark, Finland and Sweden. The estimation method was applied assuming that data only existed in census years, estimating under these conditions and the results compared to the published national figures for the foreign-born population. For Australia, the census data years were taken to be 1991, 1996 and 2001 and for the other countries 1995 and 2000.

The results are given in Table 2 and Charts 2a and 2b.

- For all countries the percent difference in the estimates of change tends to be large in the year(s) immediately following the base year and to decrease thereafter to small values. This suggests that the errors in estimation are of fixed size and decline in relative importance as the change in the

---

[2] Census figures are normally for a mid-year reference date, whereas flow data are for calendar years. The inflows from mid-year t to mid-year t+1 were estimated by taking the average of the flows for year t and year t+1.

foreign–born population becomes larger. In recent years the agreement between the estimated and official series is quite good in all countries except Australia, where the error is large and appears to be increasing. On the whole, the charts of the stock figures for the other three countries indicate that the estimated series appear to be tracking reasonably well with the official ones. For intercensal estimates, this is hardly surprising because the end-points are fixed. However, for postcensal years, where only the starting point is fixed, the tracking remains good.

- Note that the estimates for Denmark for the parametric method tend to agree more closely with the official series than those produced using the component method, which makes use of much more external information. This suggests that the recent empirical behaviour of the official series may be a better guide to its future evolution than an estimate which uses more information but incorporates a number of assumptions.

Despite the risks revealed by the recent behaviour of the Australian estimates, the parametric method has been applied to data for Canada, Ireland, New Zealand, the United Kingdom and the United States. The estimates generated are illustrated in table 3.

Conclusion

Despite the risks in estimating the foreign-born population revealed by the behaviour of the Danish data for the component method and by the Australian data for the parametric method, it was decided to proceed with the estimation for countries for which it was possible to do so. The need for current data on the foreign-born population was deemed to outweigh the possible estimation risks.

Table 1. Foreign-born population: official statistics and estimates, 1995-2004.
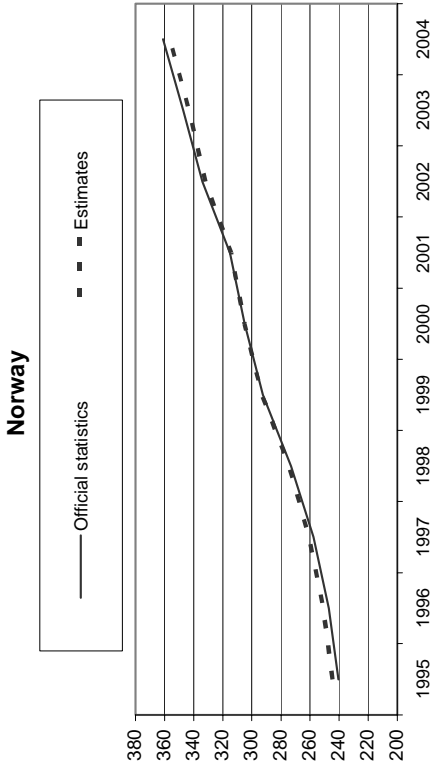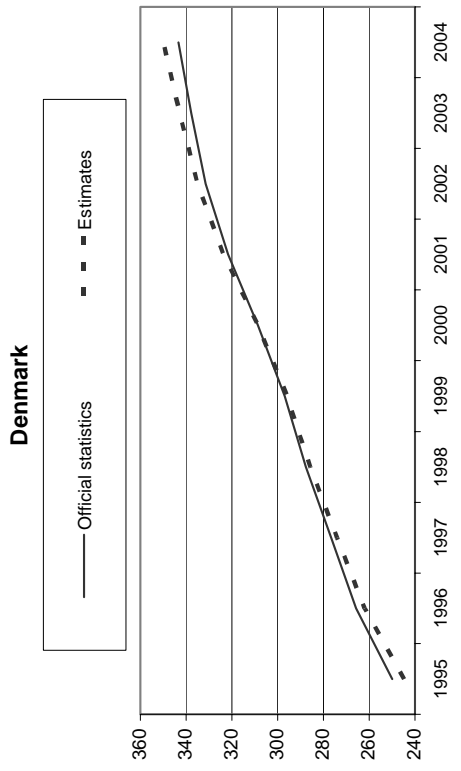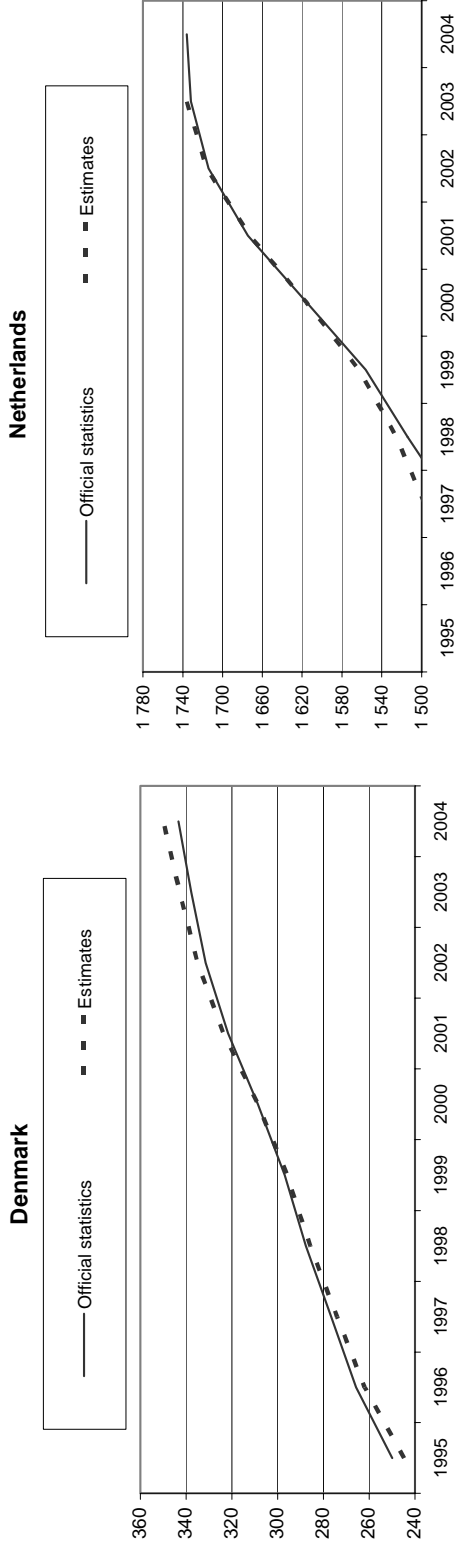Thousands and percentages

Component method	1995	1996	1997	1998	1999	Base year 2000	2001	2002	2003	2004
Denmark										
Official statistics	250	266	277	288	297	*309*	322	332	338	343
Estimates	245	262	274	286	296	*309*	324	335	343	350
Change in official statistic since base year (C)	- 59	- 43	- 32	- 21	- 12	0	13	23	29	35
Change in estimate since base year (D)	- 64	- 47	- 35	- 23	- 13	0	15	26	34	41
Percent difference [100*(D/C-1)]	9	9	9	9	9	0	14	15	17	19
Netherlands										
Official statistics	1 407	1 434	1 469	1 514	1 556	*1 615*	1 675	1 714	1 732	1 736
Estimates			1 497	1 525	1 563	*1 615*	1 672	1 716	1 736	
Change in official statistic since base year (C)	- 208	- 182	- 146	- 101	- 59	0	59	99	116	121
Change in estimate since base year (D)				- 90	- 53	0	57	100	121	
Percent difference [100*(D/C-1)]				- 11	- 11	0	- 4	2	4	
Norway										
Official statistics	240	247	258	273	292	*305*	315	334	347	361
Estimates	245	251	261	274	293	*305*	314	332	343	356
Change in official statistic since base year (C)	- 65	- 58	- 47	- 32	- 13	0	10	29	42	56
Change in estimate since base year (D)	- 60	- 54	- 44	- 31	- 12	0	9	27	38	51
Percent difference [100*(D/C-1)]	- 7	- 7	- 6	- 3	- 5	0	- 9	- 7	- 10	- 9

Chart 1a. Percent difference between change in the estimate of the foreign-born population since the base year (2000) and the change recorded by official statistics
Component method

Denmark

Norway

Netherlands
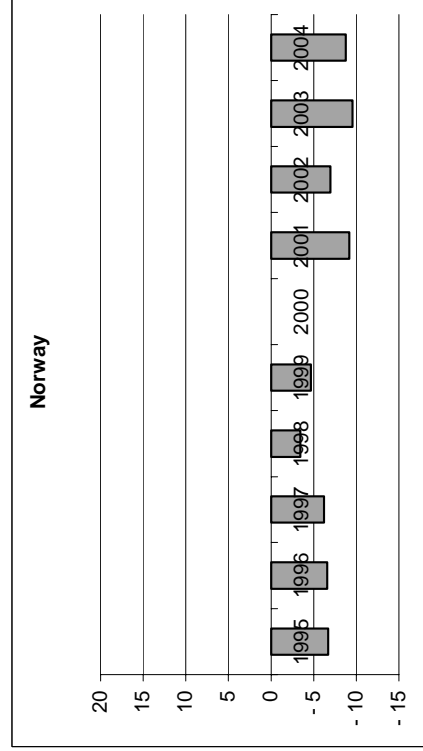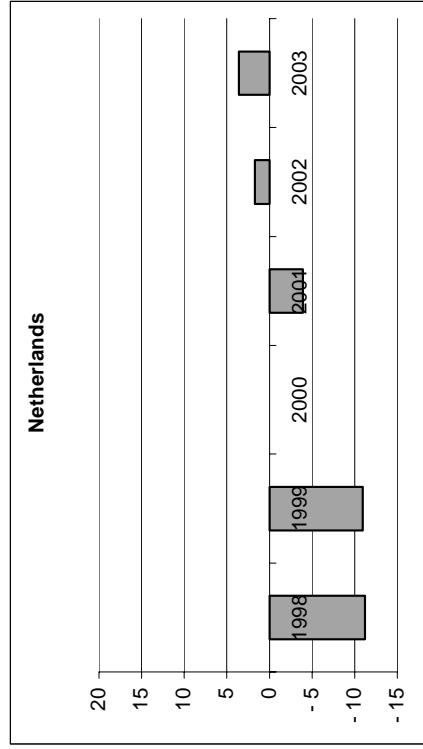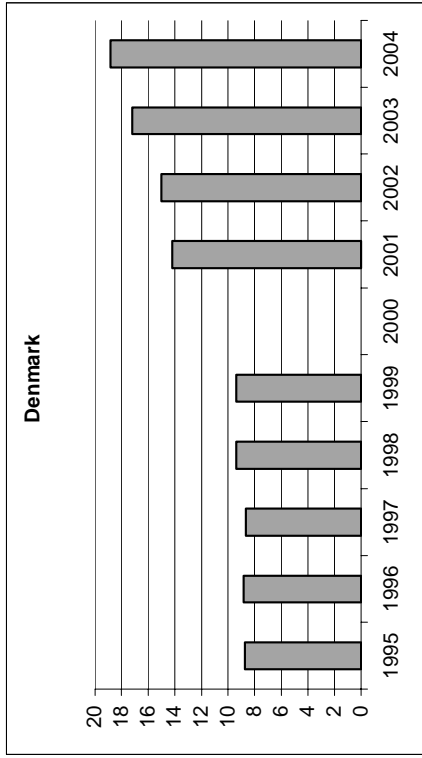
**Chart 1b. Official statistics and estimates of the foreign-born population for 3 OECD countries
(Estimates produced according to the component method)**

Denmark

Netherlands

Norway

9

Table 2. Foreign-born population: official statistics and estimates
Thousands and percentages
Parametric method

Australia

	Base year 1991	1992	1993	1994	1995	Base year 1996	1997	1998	1999	2000	Base year 2001	2002	2003	2004
Official statistics	3965	4028	4054	4085	4164	4259	4316	4335	4373	4418	4482	4566	4655	4751
Estimates	3965	4049	4110	4151	4198	4259	4308	4346	4383	4427	4482	4535	4580	4636
Change in official statistic since base year (C)	0	63	89	119	199	293	57	76	115	159	223	84	173	269
Change in estimate since base year (D)	0	84	145	186	232	293	50	88	125	169	223	53	98	154
Percent difference [100*(D/C-1)]	0	33	63	56	17	0	-13	15	9	6	0	-37	-44	-43

Denmark

	Base year 1995	1996	1997	1998	1999	Base year 2000	2001	2002	2003	2004
Official statistics	250	266	277	288	297	309	322	332	338	343
Estimates	250	269	280	290	299	309	320	331	338	343
Change in official statistic since base year (C)	0	16	27	38	47	59	13	23	29	35
Change in estimate since base year (D)	0	19	31	40	49	59	12	22	29	35
Percent difference [100*(D/C-1)]	0	18	14	6	5	0	-11	-2	1	0

Finland

	Base year 1995	1996	1997	1998	1999	Base year 2000	2001	2002	2003	2004
Official statistics	106	111	118	125	131	136	145	152	159	166
Estimates	106	112	118	124	130	136	144	152	159	167
Change in official statistic since base year (C)	0	5	12	19	25	30	9	16	23	30
Change in estimate since base year (D)	0	6	11	18	24	30	8	16	23	30
Percent difference [100*(D/C-1)]	0	16	-2	-6	-5	0	-14	-1	0	1

Sweden

	Base year 1995	1996	1997	1998	1999	Base year 2000	2001	2002	2003	2004
Official statistics	936	944	954	969	982	1004	1028	1053	1078	1100
Estimates	936	948	959	973	987	1004	1025	1049	1074	1098
Change in official statistic since base year (C)	0	8	18	33	46	68	24	50	74	96
Change in estimate since base year (D)	0	12	23	37	51	68	22	45	70	94
Percent difference [100*(D/C-1)]	0	58	27	12	11	0	-11	-9	-6	-2

Chart 2a. Percent difference between change in the estimate of the foreign-born population since the base year(s) and the change recorded by official statistics.
(Estimates produced using the parametric method)
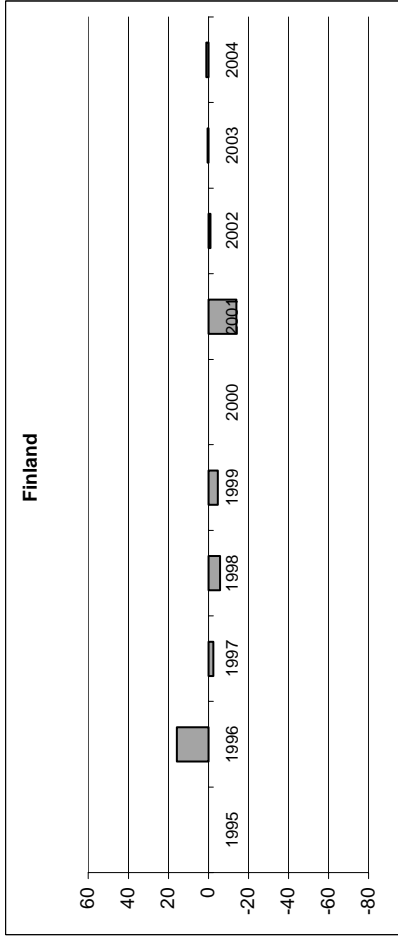

Australia

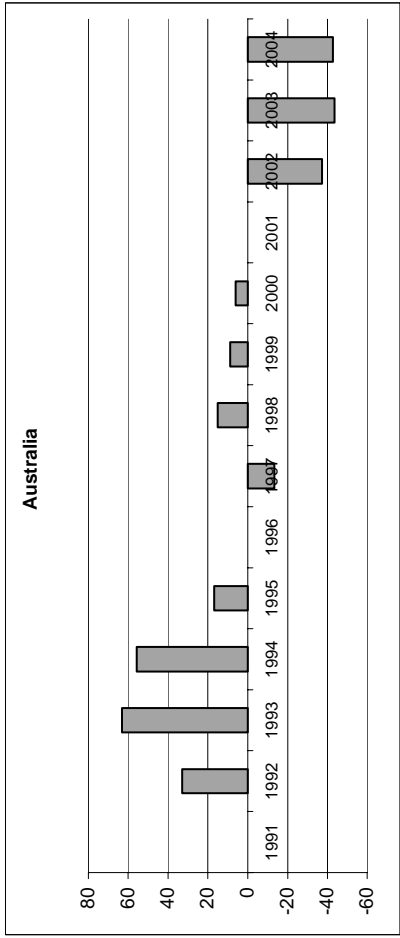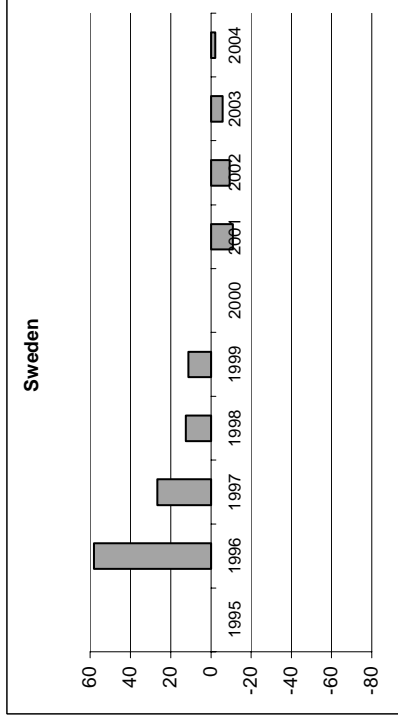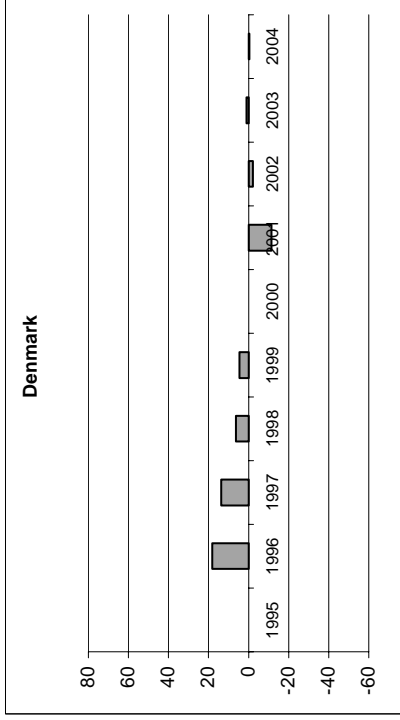

Denmark

Finland

Sweden

Chart 2b. Official statistics and estimates of the foreign-born population for 4 OECD countries.
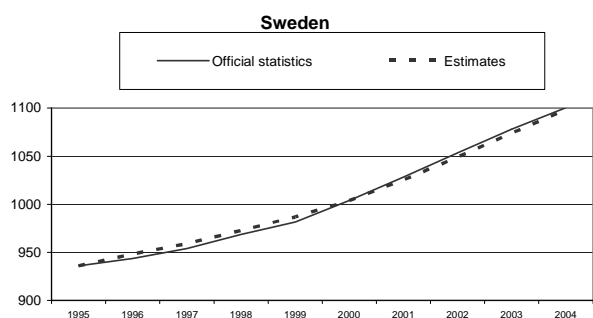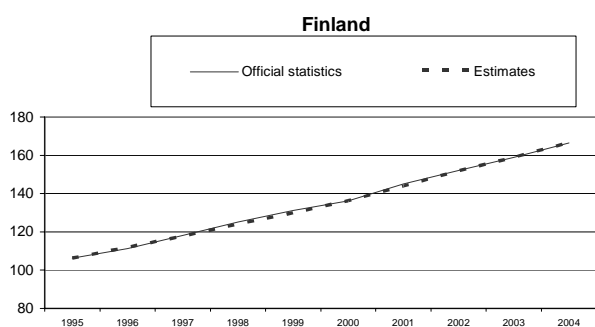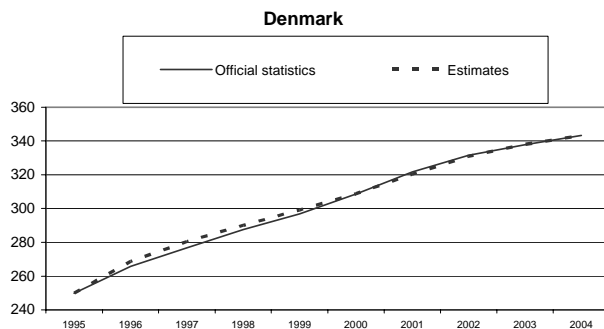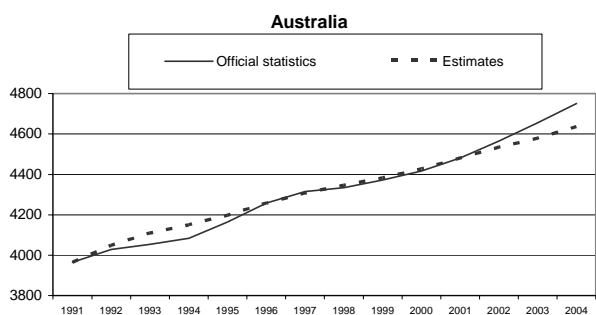**(Estimates produced according to the parametric method)**

Australia

Denmark

Finland

Sweden

Table 3. **Stocks of foreign-born population in selected OECD countries**
Thousands

	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
Australia	4 164.1	4 258.6	4 315.8	4 334.8	4 373.3	4 417.5	4 482.0	4 565.8	4 655.3	4 751.1
% of total population	23.0	23.3	23.3	23.2	23.1	23.0	23.1	23.2	22.8	23.6
Austria	..	..	..	895.7	872.0	843.0	893.9	873.3	923.4	1 059.1
% of total population	..	..	..	11.2	10.9	10.5	11.1	10.8	11.4	13.0
Belgium	*983.4*	*999.2*	*1 011.0*	*1 023.4*	*1 042.3*	1 058.8	1 112.2	1 151.8	1 185.5	..
% of total population	*9.7*	*9.8*	*9.9*	*10.0*	*10.2*	10.3	10.8	11.1	11.4	..
Canada	*4 867.4*	*4 971.1*	*5 082.5*	*5 165.6*	*5 233.8*	*5 327.0*	*5 448.5*	*5 568.2*	*5 670.6*	*5 781.3*
% of total population	*17.2*	*17.4*	*17.7*	*17.8*	*18.0*	*18.1*	*18.4*	*18.6*	*18.7*	*18.9*
Czech Republic	..	..	..	440.1	455.5	434.0	448.5	471.9	482.2	499.0
% of total population	..	..	..	4.3	4.4	4.2	4.4	4.6	4.7	4.9
Denmark	249.9	265.8	276.8	287.7	296.9	308.7	321.8	331.5	337.8	343.4
% of total population	4.8	5.1	5.2	5.4	5.6	5.8	6.0	6.2	6.3	6.3
Finland	106.3	111.1	118.1	125.1	131.1	136.2	145.1	152.1	158.9	166.4
% of total population	2.0	2.1	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.2
France	..	..	..	..	5 868.2	..	..	..	..	..
% of total population	..	..	..	..	10.0	..	..	..	..	..
Germany	*9 377.9*	*9 708.5*	*9 918.7*	*10 002.3*	*10 172.7*	*10 256.1*	*10 404.9*	*10 527.7*	*10 620.8*	..
% of total population	*11.5*	*11.9*	*12.1*	*12.2*	*12.4*	*12.5*	*12.6*	*12.8*	*12.9*	..
Greece	..	..	..	..	..	..	1 122.9	..	..	..
% of total population	..	..	..	..	..	..	10.3	..	..	..
Hungary	283.7	283.9	284.2	286.2	289.3	294.6	300.1	302.8	307.8	319.0
% of total population	2.8	2.8	2.8	2.8	2.9	2.9	3.0	3.0	3.0	3.2
Ireland	..	*251.6*	*271.2*	*288.4*	*305.9*	*328.7*	*356.0*	*390.0*	*416.6*	*443.0*
% of total population	..	*6.9*	*7.4*	*7.8*	*8.2*	*8.7*	*9.3*	*10.0*	*10.5*	*11.0*
Italy	..	..	..	..	..	..	1 446.7	..	..	..
% of total population	..	..	..	..	..	..	2.5	..	..	..
Luxembourg	*127.7*	*130.9*	*134.1*	*137.5*	*141.9*	*145.0*	*144.8*	*147.0*	*148.5*	*149.6*
% of total population	*30.9*	*31.5*	*31.9*	*32.2*	*32.8*	*33.2*	*32.8*	*32.9*	*33.0*	*33.1*
Mexico	385.2	..	..	..	..	406.0	..	..	..	..
% of total population	0.4	..	..	..	..	0.5	..	..	..	..
Netherlands	1 407.1	1 433.6	1 469.0	1 513.9	1 556.3	1 615.4	1 674.6	1 714.2	1 731.8	1 736.1
% of total population	9.1	9.2	9.4	9.6	9.8	10.1	10.4	10.6	10.7	10.6
New Zealand	..	*605.0*	*620.8*	*630.5*	*643.6*	*663.0*	*698.6*	*726.3*	*748.6*	*763.6*
% of total population	..	*16.2*	*16.4*	*16.5*	*16.8*	*17.2*	*18.0*	*18.4*	*18.7*	*18.8*
Norway	240.3	246.9	257.7	273.2	292.4	305.0	315.2	333.9	347.3	361.1
% of total population	5.5	5.6	5.8	6.1	6.5	6.8	6.9	7.3	7.6	7.8
Poland	..	..	..	..	..	..	..	776.2	..	..
% of total population	..	..	..	..	..	..	..	1.6	..	..
Portugal	*533.6*	*529.2*	*523.4*	*516.5*	*518.8*	*522.6*	*651.5*	*699.0*	*704.6*	*704.4*
% of total population	*5.4*	*5.4*	*5.3*	*5.1*	*5.1*	*5.1*	*6.3*	*6.7*	*6.7*	*6.7*
Slovak Republic	..	..	..	..	..	..	119.1	..	..	207.6
% of total population	..	..	..	..	..	..	2.5	..	..	3.9
Spain	..	..	..	..	..	..	2 172.2	..	..	..
% of total population	..	..	..	..	..	..	5.3	..	..	..
Sweden	936.0	943.8	954.2	968.7	981.6	1 003.8	1 028.0	1 053.5	1 078.1	1 100.3
% of total population	10.5	10.7	10.8	11.0	11.8	11.3	11.5	11.8	12.0	12.2
Switzerland	*1 503.2*	*1 509.5*	*1 512.8*	*1 522.8*	*1 544.8*	1 570.8	*1 613.8*	*1 658.7*	*1 697.8*	*1 737.7*
% of total population	*21.4*	*21.3*	*21.3*	*21.4*	*21.6*	21.9	*22.3*	*22.8*	*23.1*	*23.5*
Turkey	..	..	..	..	..	1 278.7	..	..	..	..
% of total population	..	..	..	..	..	1.9	..	..	..	..
United Kingdom	4 030.7	4 131.9	4 222.4	4 335.1	4 486.9	4 666.9	4 865.6	5 075.6	5 290.2	5 552.7
% of total population	6.9	7.1	7.2	7.4	7.6	7.9	8.2	8.6	8.9	9.3
United States	*26 255.4*	*27 269.1*	*28 329.4*	*29 266.8*	*30 135.3*	31 107.9	*32 296.6*	*33 601.2*	*34 736.2*	*35 820.9*
% of total population	*9.9*	*10.1*	*10.4*	*10.6*	*10.8*	11.0	*11.3*	*11.7*	*11.9*	*12.2*

Note: Estimated figures are in italic. Data for Canada, Ireland, New Zealand, the United Kingdom and the United States are estimated with the parametric method. Data for Belgium (1995-1999), Czech Republic, Germany, Luxembourg, Portugal and Switzerland are estimated with the component method.