

ADJUSTMENTS, VARIABLE CODES AND DATA SOURCES

Of the 29 countries taking part in the project, 23 have population censuses and seven have population registers. Other sources were identified by some countries but the census or the population register is generally the most suitable source (see attached table on data sources).

For the great majority of the countries involved, data by country of birth are available. For some countries the situation is, however, more problematic. In the case of Japan, for example, the data by country of origin and level of education were not published or processed at the time of the drafting of this document even if they appear in the census. In the case of the Netherlands, the data on education are not available from the population register and it was thus necessary to use the labour force survey averaged over several years (2000-2002), in order to estimate the foreign-born by level of education and country of birth (for those countries of birth for which there were samples large enough to support reliable estimates).

Korea and Japan do not identify the foreign-born in their censuses. For these countries, because naturalisations are rare, nationality can serve as a reasonable proxy for country of birth. This approximation was not possible, however, in the case of Germany where the only data available, from the annual Microcensus (1999-2002), does not record the place of birth, although it does record the nationality and whether or not a person was born in Germany. In this case to compile data on expatriates the following assumptions and adjustments were made: (i) for non-German citizens born abroad, it was assumed that their place of birth was the same as their nationality; (ii) for "unknown" place of birth or nationality in the Microcensus, a response was attributed according to the distribution observed when a response was available; and (iii) for German citizens born abroad, the German Socioeconomic Panel (GSOEP), which does identify the place of birth, was used for those countries for which the sample was large enough to produce reliable estimates. The data included in the publicly available OECD data file, however, does not include the adjustments which were made through the GSOEP.

With regard to the structure of the levels of qualification retained, it was decided to take into consideration five levels compatible with the International Standard Classification of Education (ISCED): ISCED 0/1/2: Less than upper secondary; ISCED 3/4: Upper secondary and post-secondary non-tertiary; ISCED 5A: "Academic" tertiary; ISCED 5B: "Vocational" tertiary; ISCED 6: Advanced research programmes. The detail at the higher levels, however, was available only for a subset of countries. For France, Switzerland, Luxembourg and Austria 5A and 6 are not distinguishable; for the United States, Turkey, Mexico and Spain 5A and 5B are not distinguishable; for the Slovak Republic, Korea, Netherlands and Hungary 5A, 5B and 6 are not distinguishable.

The objective was to minimize residual (i.e. "other") categories, with regard to the coding of countries of birth. An attempt was made to preserve the maximum information available while distinguishing between continental/regional residual categories whenever this was possible (ie. "other Africa", "other Europe", "other Asia", "other South and Central America and Caribbean", "other Oceania", "other North America").

With regard to split, recomposed or newly constituted countries, there was little choice but to respect the coding in the national data collection, which varies from one country to another. In the United States, for example, people born in Korea have the choice of three ways to indicate their country of birth: Korea, North Korea or South Korea. More than 80% of them (80% of the nationals

and 85% of foreigners) indicated having been born in Korea¹, without further specification. In the censuses of many OECD countries the Czech Republic and the Slovak Republic are aggregated under the name of the former Czechoslovakia. The same applies to the former USSR and the former Yugoslavia and Yemen.

To produce a consistent list of countries of birth across receiving countries, some minor adjustments had to be made, especially with respect to small islands and overseas territories. This recoding explains the small differences that might exist with national estimates for foreign-born and native-born populations. The following recodings were carried out:

AUS	DNK	FRA	GBR	PRT	USA²
<ul style="list-style-type: none"> • Heard & McDonald Islands 	<ul style="list-style-type: none"> • Faeroe Islands • Greenland 	<ul style="list-style-type: none"> • French southern territories • Tromelin Island • Guadeloupe • Martinique • Reunion • Juan De Nova Island • Guyane • Mayotte • Saint-Pierre-et-Miquelon 	<ul style="list-style-type: none"> • Channel Islands • Isle of Sark • Isle of Man 	<ul style="list-style-type: none"> • Madeira Islands • Azores Islands 	<ul style="list-style-type: none"> • US minor island • Christmas isle • Wake Island • Palmyra Atoll • Navassa Island • Midway Islands • Johnston Atoll • Howland Island • Baker Island

¹ It is not possible to distinguish between Koreans who emigrated to the United States before and after 1953.

² People born in Puerto Rico are considered as foreign born in the United States.

Data sources

<u>Country</u>	<u>Data year(s)</u>	<u>ISO code</u>	<u>Type of source</u>	<u>Source description</u>
Australia	2001	AUS	CEN	Australian Census of Population & Housing
Austria	2001	AUT	CEN	Census of Population
Belgium	2001	BEL	GSS	General Socio-Economic Survey
Canada	2001	CAN	CEN	Census of Population
Czech Rep	2001	CZE	CEN	Census of population
Denmark	2002	DNK	REG	Register-based population and labour force statistics
Finland	2000	FIN	REG	Population statistics
France	1999	FRA	CEN	Census of Population
Germany	1999-2002	DEU	LFS	Microcensus
Greece	2001	GRC	CEN	Census of population
Hungary	2001	HUN	CEN	Census of Population
Ireland	2002	IRL	CEN	Census of Population
Italy	2001	ITA	CEN	Census of Population
Japan	2000	JPN	CEN	Census of Population
Korea	2000	KOR	CEN	Census of population
Luxembourg	2001	LUX	CEN	Census of Population
Mexico	2000	MEX	CEN	Census of population
Netherlands	1995-2000	NDL	REG	Matched data from the Population Registers, the Tax Department and the Ministry of Justice
Netherlands	1999-2001	NDL	LFS	Labour Force Survey
New Zealand	2001	NZL	CEN	Census of Population and Dwellings
Norway	2003	NOR	REG	Various administrative and statistical registers
Poland	2001	POL	CEN	Census of population
Portugal	2001	PRT	CEN	Census of population
Slovak Rep	2001	SVK	CEN	Census of population
Spain	2001	ESP	CEN	Census of Population
Sweden	2003	SWE	REG	Total Population Register TPR
Sweden	2003	SWE	EDU	Education register
Switzerland	2000	CHE	CEN	Census of Population
Turkey	2000	TUR	CEN	Census of Population
United Kingdom	2001	GBR	CEN	Census of Population
United States	2000	USA	CEN	Census 5% Public Use Microdata Sample