

Max Planck Institute for Demographic Research<sup>1</sup>  
Rostock  
Germany

# The Human Fertility Database

A.Jasilioniene<sup>1</sup>, T. Sobotka<sup>2</sup>, D. Jdanov<sup>1</sup>, E.M. Andreev<sup>1</sup>, K. Zeman<sup>2</sup>, V.M. Shkolnikov<sup>1</sup>, and J. Goldstein<sup>1</sup>

Vienna Institute of Demography<sup>2</sup>  
Vienna  
Austria



## 1. INTRODUCTION

The Human Fertility Database Project is a joint project of the Max Planck Institute for Demographic Research (MPIDR) and the Vienna Institute of Demography (VID), which aims

- to provide open and user-friendly access to detailed, well-documented and high-quality data on period and cohort fertility,
- to facilitate research on changes and inter-country differences in fertility in the past and in the modern era.

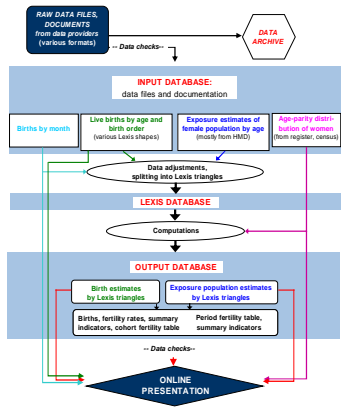
Following the example of the Human Mortality Database (HMD, <http://www.mortality.org>), our guiding principles are *comparability, flexibility, accessibility, and reproducibility*.

The HFD Project includes two principal parts:

**Human Fertility Database (HFD)** will be entirely based on one and the same type of initial birth data, which will be further processed using uniform methodology. The HFD will feature order-specific data on births, two types of order-specific fertility rates – unconditional and conditional rates, cohort and period fertility tables as well as selected aggregate fertility indicators.

**Human Fertility Collection (HFC)** will include sets of fertility rates and fertility tables constructed by other researchers, research organizations, and statistical agencies.

## 2. DATA PROCESSING IN THE HFD



## 2.1 Raw data. Data archive

The data archive stores raw data files and all accompanying documents, supplied by the data providers.

## 2.2 Input database

The Input Database stores the original raw data converted into standard input format and the documents describing the data.

Data files (ASCII format):

1. (Live) births by calendar year, mother's age and/or year of birth, and birth order.
2. Age-parity distribution of female population (from population register, population census or large scale survey).
3. Births by calendar month and year.
4. Female population size by age.

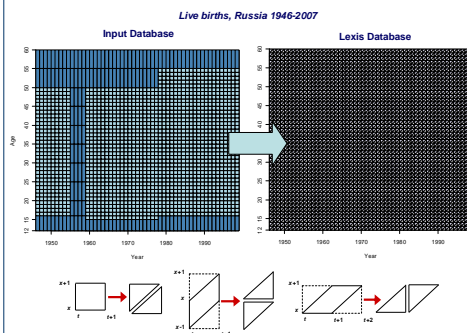
Text files, i.e. documents describing the data (PDF):

1. Background and documentation providing information about data quality, definitions, coverage, etc.
2. Notes providing details about specific data points in the data files.
3. References to the data sources.

## 2.3 Adjustments to raw data

Raw data on births are processed in a standardized adjustment procedure, which involves the following:

- > redistribution of births of unknown age and of unknown birth order,
  - > estimation of the proportion of live births if only information on total births (i.e. including stillbirths) exist,
  - > splitting five-year age groups into one-year age groups,
  - > splitting open age intervals into single age groups,
  - > making ranges of age and birth order 'uniform' (age -12 to 50+, birth order 1, 2, 3, 4, and 5+),
  - > splitting larger Lexis elements into Lexis triangles,
- etc.



## 2.4 Lexis database

All output data are produced on the basis of two 'Lexis' files: 'Lexis file' for births and 'Lexis file' for exposure population. These two files are displayed on the website as separate output files as well.

## 2.5 Output database

The major blocks of output data (order- and non-order specific) are as follows:

- Births, population exposure, and unconditional rates by different Lexis elements.
- Summary indicators of period and cohort fertility.
- Cohort and period fertility tables, conditional fertility rates, and table summary indicators.

For a complete list of the output files see the copy of HFD country page for Austria (at the right side).

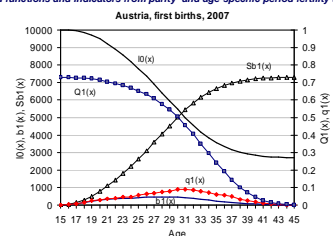
Examples of the output files for period data on births, birth order 1 to 5+ (Austria)

Year	Age	Subpart	Total	1	2	3	4	5p
2007	12	1995	0.00	0.00	0.00	0.00	0.00	0.00
2007	12	1994	0.00	0.00	0.00	0.00	0.00	0.00
2007	13	1994	0.00	0.00	0.00	0.00	0.00	0.00
2007	13	1993	1.00	1.00	0.00	0.00	0.00	0.00
2007	21	1986	1056.00	800.00	226.00	0.00	0.00	0.00
2007	21	1985	1080.00	769.00	287.00	0.00	0.00	0.00
2007	22	1985	1205.00	846.00	359.00	0.00	0.00	0.00
2007	22	1984	1374.00	900.00	415.00	0.00	0.00	0.00
2007	23	1984	1551.00	995.00	459.00	0.00	0.00	0.00

Example of the output file for period fertility tables (Austria)

x	Age at childbearing	Year	w0(x)	w1(x)	q0(x)	l0(x)	b1(x)	l0(x)	l1(x)	w1(x)
1	Both order of child. relates to parity i-1	1991	12	1.00000	0.00000	0.00000	10000	0	10000	0
1	Table number of births of order i	1991	12	1.00000	0.00000	0.00000	10000	0	10000	0
1	Table population of parity i-1	1991	13	0.99999	0.00005	0.00005	10000	0	10000	0
1	Probability of having an i-th birth	1991	14	0.99940	0.00047	0.00047	10000	5	9997	0
1	Conditional age-specific fertility rates	1991	15	0.99850	0.00210	0.00210	9995	21	9984	24
1	Cumulative births of order i	1991	16	0.99617	0.00563	0.00568	9974	65	9942	93
1	Average number of children at highest parity i-1	1991	17	0.98062	0.01737	0.01722	9909	171	9824	261
1	"Likewise" probability of having an i-th birth for women at parity i-1	1991	18	0.95776	0.03030	0.02885	9739	291	9593	552
1		1991	19	0.92325	0.04563	0.04461	9448	422	9237	974

Selected functions and indicators from parity- and age-specific period fertility tables (Austria)



## 2.6 Data checks

Rigorous and standardized data checking will be performed once the raw data are received and later again when the output is obtained.

## 3. HFD TEST WEBSITE: <http://www-t.humanfertility.org>

## 4. FUTURE PLANS

At present, we are in contact with experts or institutions who have received or expect to receive soon the data from the following countries: Austria, Bulgaria, Canada, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Italy, Japan, Lithuania, the Netherlands, Norway, Portugal, Romania, Russia, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Kingdom, and USA.

We trust that the HFD methods protocol will be fully completed and data for 6-7 countries will be prepared for the launch of the database in autumn 2009 (IUSSP conference, Marrakech). The official HFD website will be <http://www.humanfertility.org>.

In the future, the HFD will also include data on births and fertility rates by duration since the last previous birth (birth interval indicators) for countries, for which this kind of data exist.

We believe that in the future the HFD (and HFC) will become a widely used database and will be recognized as the best resource of reliable and detailed fertility data.

**Acknowledgements**  
Creating the Human Fertility Database is a very challenging endeavor, and we are very grateful to our colleagues from various research institutions for their great help. We would like to express our special gratitude to Laurence Toulemon (INED, Paris) and John Wilmoth (University of California, Berkeley) for their valuable feedback on the HFD methodology and general advice in the project. We are also thankful to Jean-Pierre Sillion (INED, Paris) who so kindly shared with us some data that had been collected by European Demographic Observatory (EDO) and also gave us useful methodological suggestions.

Contact: Aina Jasilioniene (ajasilione@demography.mpg.de)

