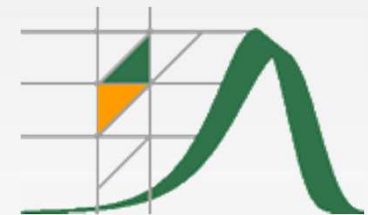


The HFD and HFC in a nutshell

Dmitri Jdanov and Tomas Sobotka
on behalf of the HFD team



The Human Fertility Data Project

Human Fertility Database: Expanding research opportunities

Member-initiated meeting

2015 PAA Annual Conference, San Diego, 29 April 2015



MAX PLANCK INSTITUTE
FOR DEMOGRAPHIC
RESEARCH

MAX-PLANCK-INSTITUT
FÜR DEMOGRAFISCHE
FORSCHUNG





Human Fertility Data Project

The Project is a joint endeavour of the **Max Planck Institute for Demographic Research** (MPIDR) and the **Vienna Institute of Demography** (VID).

Aim: to provide free and user-friendly access to a variety of fertility data

The Project consists of two companion databases, based at the MPIDR in Rostock:

- **Human Fertility Database (HFD)**, www.humanfertility.org
- **Human Fertility Collection (HFC)**, www.fertilitydata.org

Official launch:

- **HFD in September 2009** (first time presented at the XXVI International IUSSP Population Conference, Marrakech, Morocco, September 2009)
- **HFC in August 2013** (first time presented at the XXVII International IUSSP Population Conference, Busan, Korea, August 2013)

Inspiration: Human Mortality Database (HMD)
Human Life Table Database (HLD)



Motivation, aims

- Vastly expanding the availability of comparative data on fertility around the world
- Open access data
- Putting together datasets and data collections that were not easily accessible or comparable
- Providing standardised sets of measures and indicators that can readily be used
- Going beyond the period TFR: giving a boost to aggregate-level & comparative research fertility
- Expanding the range of available fertility measures and indicators: birth order, parity, and cohort dimensions



“Classical” sources of demographic data at population level & their limitations

➤ Publications by national statistical offices / other official institutes

- Limited to one country
- More detailed data or documentation often unavailable in English
- Mostly focused on recent data, without historical series
- No universal format of data, vastly different levels of detail

➤ Collections of official data (Eurostat, WHO mortality database)

- Data published as they are provided
- Mixture of “good” and “bad” data
- Limited historical coverage
- Data not fully comparable across time and countries
- No documentation about data collection, estimations, reliability, data manipulation, etc.

➤ Official reports or data collections which must cover specific range of countries (UN, World Bank, CIA)

- Mixture of real data, sample-based estimates, and modeled data
- Data are not comparable across time and countries
- Relatively short data series
- Fragmentary documentation



HFD: a new type of demographic databases

- **Why to spend resources on production of demographic data in times of easy data availability on the Internet?**

Substantial “added value” compared to other data sources: very detailed, well documented, uniform, high quality data, comparable across time and space.

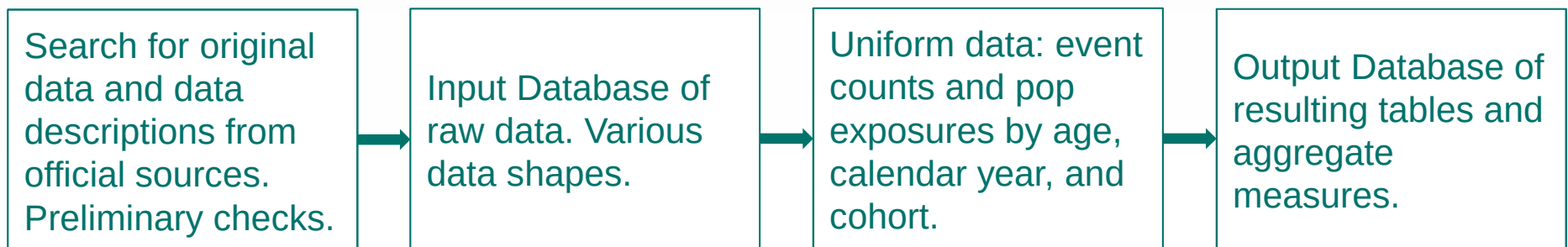
- **HFD main principles:**

free & open access to all raw data, thorough documentation of data sources and computational methods, data in user-friendly format, data comparability (across countries and time), reproducibility, flexibility, *responsibility for the data* => ***data are ready for analysis***

- **Target users:**

researchers, university faculty members, students.

Principal data flows





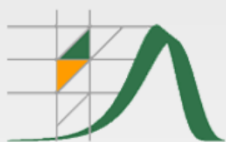
HFD vs. HFC: what's the difference?

➤ Human Fertility Database / HFD

- Attention to data quality & comparability
- All estimates computed using standardised methods
- Carefully checked, often intensive collaboration with providers
- Detailed documentation
- Wide range of indicators including cohort data and detailed parity-specific fertility tables
- Taking Human Mortality Database for population exposure
- Drawbacks: 1) Too perfect? Labour-intensive, quality limits to expansion. 2) Difficulties in keeping it up to date

➤ Human Fertility Collection / HFC

- Simpler framework: less complex, very limited data checks, incorporates multiple sources of data
- Fewer indicators
- Global scale: Wide and expanding data coverage

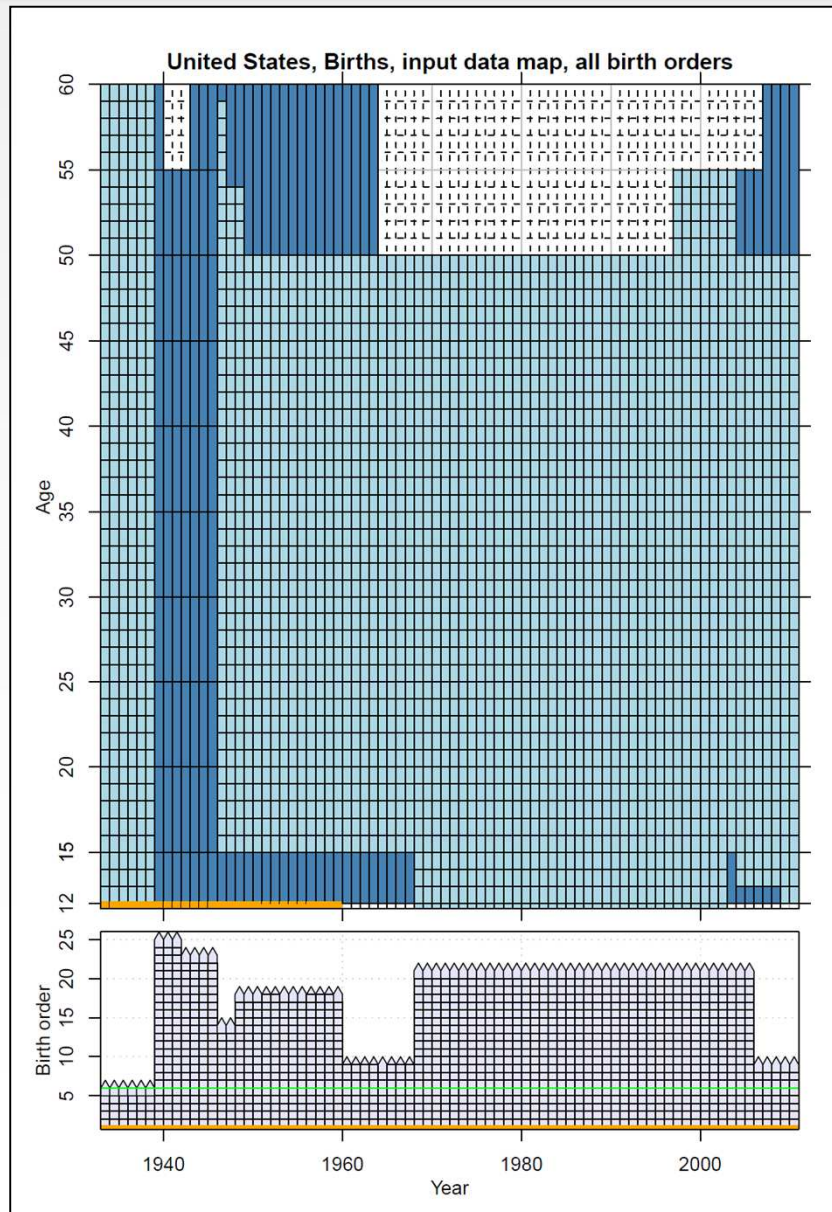
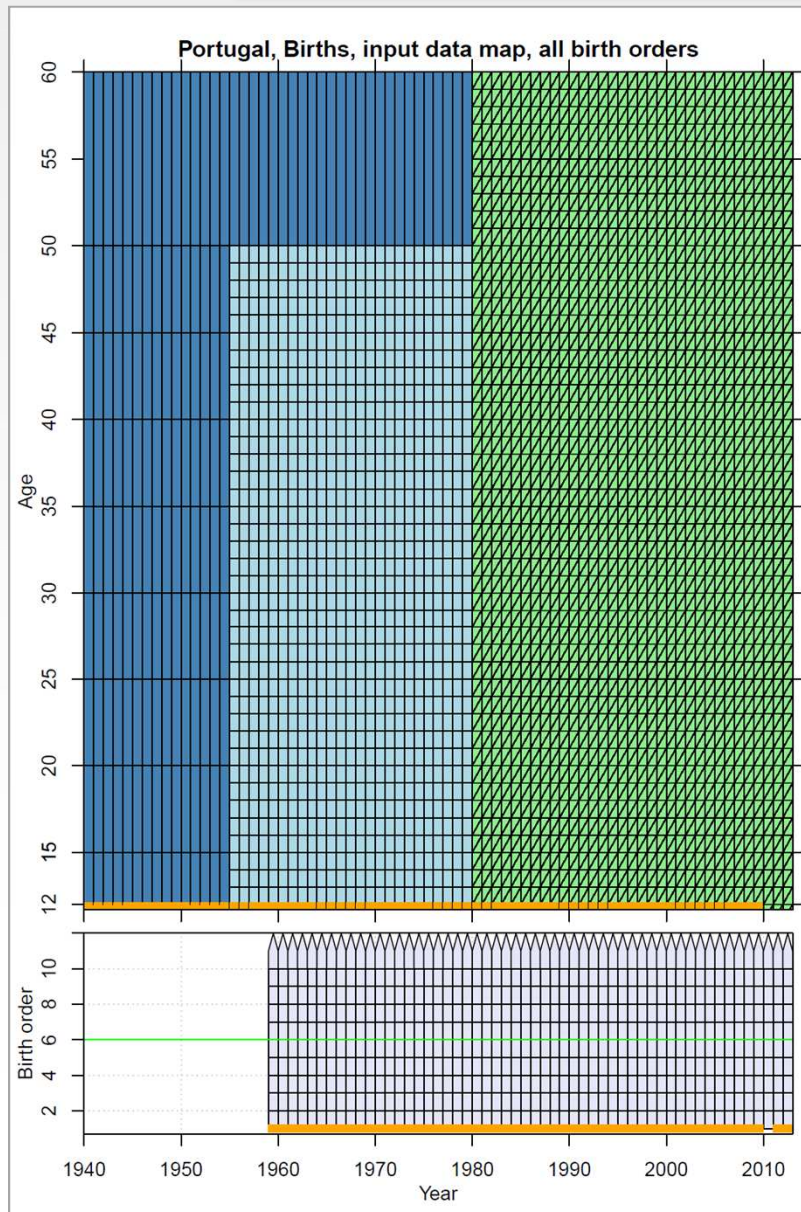


Basic facts about HFD and HFC

	HFD	HFC
Launched	September 2009	August 2013
Initial number of countries	6	73
Current number of countries	26 (31 incl. subpopulations)	89
Coming soon	Italy, Poland, Spain, Ireland, Denmark, Latvia	Bangladesh, China, Fiji, Mexico, Cuba, Iran; data for regions
Foreseeable expansion	45-50 countries (limited by data quality)	110-140 countries Data for subpopulations, regions
Period (min – max)	1891-2012, continuous data series	1751-2013
Update of country series	Annual/biannual; update depends on publishing of official data and update of the population exposure (HMD)	Intensive data search
Available indicators	Period and cohort ASFR, fertility tables and summary indicators, birth counts, pop exposures. Data by single age groups and birth order	Period ASFR, TFR, MAB Data by single age groups and birth order
Number of users	~ 4.000	n/a

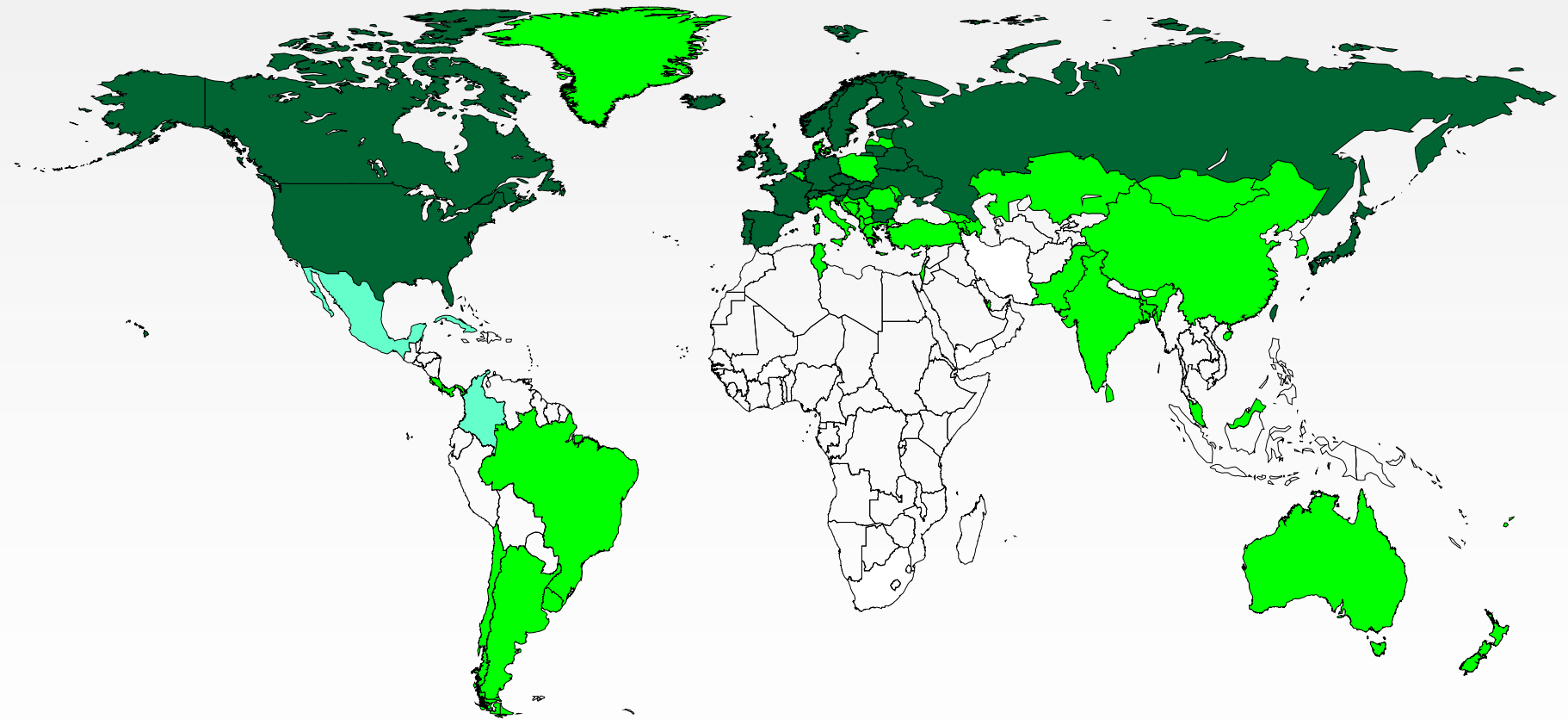


HFD building blocs: Lexis elements





The HFD/HFC Data Map



HFC and HFD countries



HFC countries

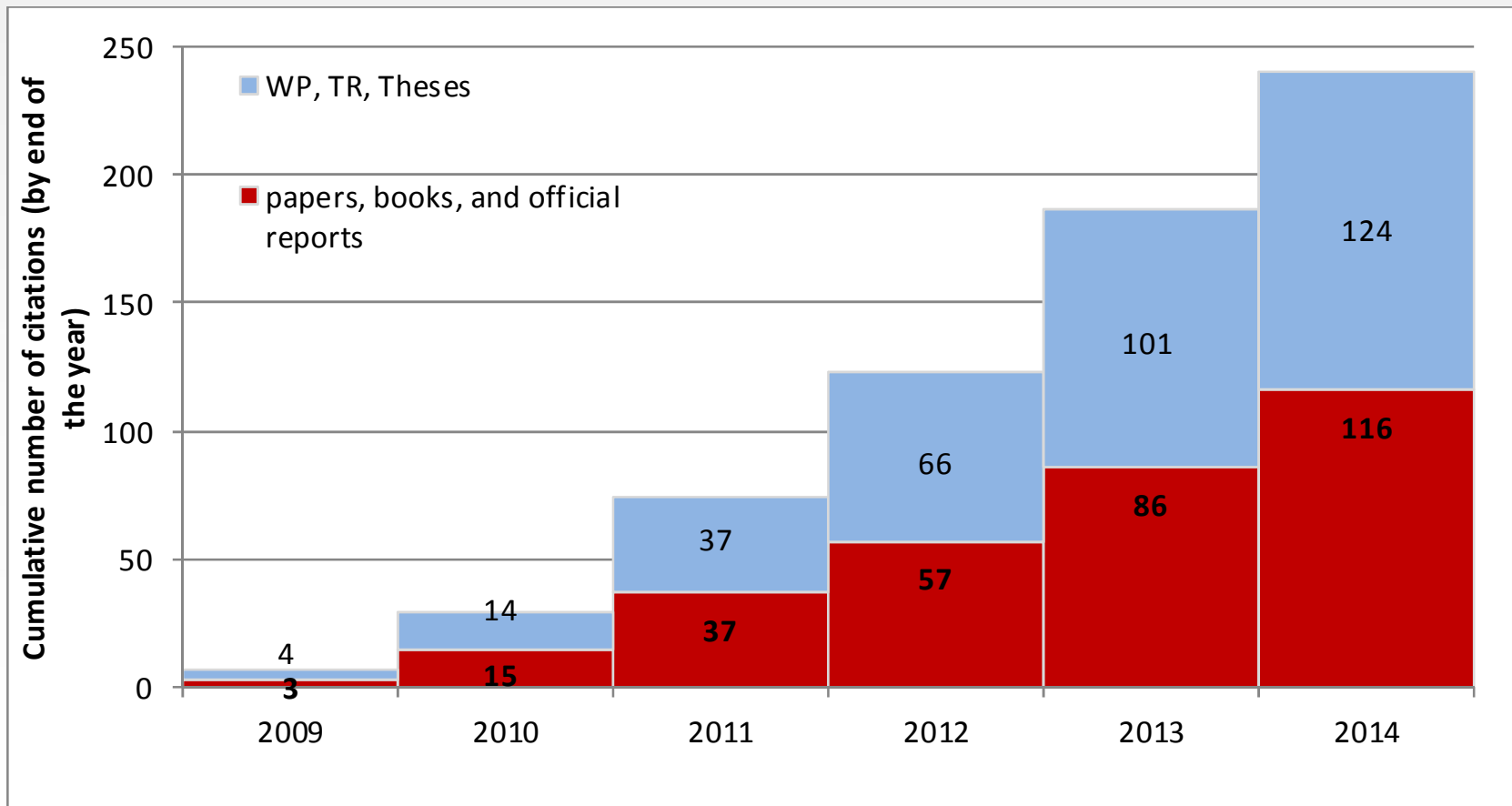


Coming soon to HFC





HFD citations



The full list of publications using the HFD/HFC data is available at HFD web site (HFD publications). *If you don't find your publication on the list please send it to us!*



Why prefer using HFD over other data sources?

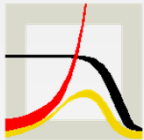
Things you might not know about the official fertility data

- Population exposures may suddenly change after each population census: discontinuity in fertility data series
 - Underreported emigration may deflate period fertility
 - Problems for countries with intensive migration
 - Central & Eastern Europe after 1989
- Mismatch between births and population (exposure) data
 - East Asia, Central & Eastern Europe
 - Can inflate period fertility
- Problems with definitions and reporting of birth order
 - Birth order partly estimated, inconsistencies in time series (Italy, Russia, Spain)
 - Birth order defined within marriage only (historical data for Germany, Switzerland, UK...)
 - Cases of unknown birth order

HFD deals with these issues by exploring & documenting them, adjusting or excluding the data & warning the users



HFD Project Team



MAX PLANCK INSTITUTE
FOR DEMOGRAPHIC
RESEARCH



MPIDR:

Vladimir M. Shkolnikov (Co-Director of HFD)
Dmitri A. Jdanov (Co-Director of HFC)
Aiva Jasilioniene (Project coordinator)
Olga Grigorieva
Dora Kostova
Pavel Grigoriev
Karolin Kubisch
Sigrid Gellers-Barkmann

VID:

Tomáš Sobotka (Co-Director of HFD & HFC)
Kryštof Zeman

Founding Director: Josh Goldstein

HFD Advisory Board:

John Wilmoth (UN)
Wolfgang Lutz (VID)
Laurent Toulemon (INED)
Jean-Paul Sardon (INED, ODE)

Other researchers involved:

Dimiter Philipov (VID)
Evgueni L. Soroko (HSE, Moscow)
Tatiana Kharkova (HSE, Moscow)
Evgueni M. Andreev (NES, Moscow)
Edward Nash (Rostock)
Tim Riffe (Berkeley)
E. E. Campos de Lima (Sao Paulo)



Collaborators

HFD + HFC rely on voluntary contributions by dozens of collaborating researchers, institutions and statistical offices

- The databases would not materialise without their involvement
- Helping with collecting, correcting, standardising and documenting the data

Selected recognitions & examples

- Jean-Paul Sardon (France): providing the records of the ODE database for Europe
- Everton E. Campos de Lima, Mathias Nathan, Ruben Castro: coordinating HFC+HFD data collection for Latin America
- Patrick Gerland, Thomas Spoorenberg (UN Pop. Division): providing data for HFC
- ISTAT, Antonella Guarneri, Sabrina Prati (Italy): collecting, adjusting and documenting Italian HFD+HFC data
- Daniel Devolder (Spain): adjusting and documenting Spanish data
- Ward Kingkade (US Census Bureau): collecting and documenting US data by birth order



Related databases

Like HMD or IPUMS, HFD inspires creation of related databases

- Russian Fertility and Mortality database (RusFMD)
 - Based at Center for Demographic Research at the New Economic School, Moscow
 - Strong regional focus
 - http://demogr.nes.ru/index.php/en/demogr_indicat/data
- Latin American Fertility Database
 - Still at the planning stage, aiming for a higher standard, better documentation, and wider coverage than the HFC
 - Based on a network of collaborators, coordinated by Everton E. Campos de Lima
- Cohort Fertility and Education database (CFE)
 - Census and large-scale survey data on completed fertility and parity distribution by level of education
 - Developed by Vienna Institute of Demography / EURREP project
 - www.cfe-database.org



Acknowledgments

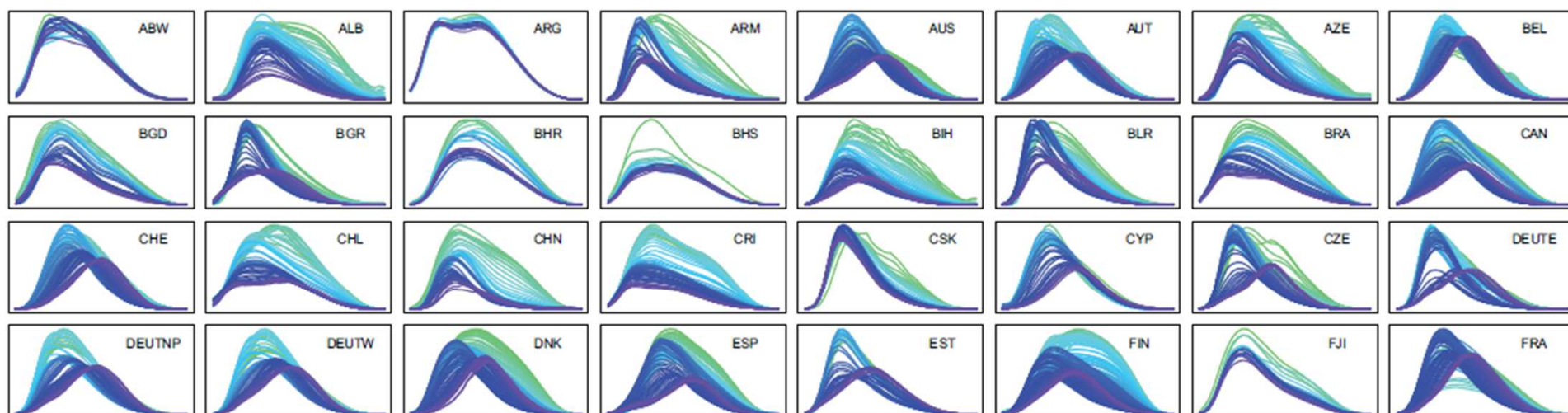
This project would not have been possible without the encouragement and financial support of the Max Planck Institute for Demographic Research (MPIDR) in Rostock, Germany and the Vienna Institute of Demography (VID), Vienna, Austria. The European Commission funded the initial work by the VID as part of the project "Reproductive decision-making in a macro-micro perspective" (REPRO) in the Seventh Framework Programme (Grant Agreement: SSH-CT-2008-217173). Since February 2012 the work of the VID team (Kryštof Zeman and Tomáš Sobotka) was funded by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement n° 284238 (EURREP project).



We acknowledge the key contribution of the HFD Founding Director Joshua R. Goldstein.

Our special gratitude for their valuable comments on the HFD methodology and general advice to the project is due to the members of the HFD Advisory Board.

We are very grateful to our data providers and country experts for their indispensable support. The full list of individual contributors is available at the HFD and HFC web sites.



HUMAN FERTILITY DATA PROJECT

A joint venture of Max Planck Institute for Demographic Research and Vienna Institute of Demography

