Selected Wittgenstein Centre databases on fertility across time and space

Kryštof Zeman^{1,*} and Tomáš Sobotka¹

This contribution provides a brief structured overview of selected open-access databases and data websites developed by the Wittgenstein Centre for Demography and Global Human Capital, in some cases in cooperation with partner institutions or based on the data supplied by partner institutions. The presented databases primarily focus on fertility data or have a broader focus, but provide detailed coverage of fertility data. These databases also showcase the breadth and scope of fertility-related research at the Wittgenstein Centre, as they range from detailed data that focus on one country and even one city (*Birth Barometer* for Austria and Vienna), to data that provide comprehensive coverage of one continent (*European Demographic Data Sheet*), to databases that offer global coverage (*Human Fertility Collection*, *Wittgenstein Centre Human Capital Data Explorer*). These databases also cover very different time periods, with the oldest data in the *Human Fertility Collection* pertaining to 1751, and the *Wittgenstein Centre Data Explorer* featuring projections up to 2100.

We review the following six databases:

- Human Fertility Database (https://www.humanfertility.org)
- Human Fertility Collection (https://www.fertilitydata.org)
- Birth Barometer Monitoring fertility in Austria (http://www.birthbarometer.at)
- European Demographic Data Sheet (http://www.populationeurope.org)
- Wittgenstein Centre Human Capital Data Explorer (http://dataexplorer.wittgensteincentre.org)
- Cohort Fertility and Education Database (http://www.cfe-database.org)

DOI: 10.1553/populationyearbook2020.dat01

¹Vienna Institute of Demography (VID), Austrian Academy of Sciences, Wittgenstein Centre for Demography and Global Human Capital (IIASA, OeAW, University of Vienna), Vienna, Austria

^{*}Correspondence to: Kryštof Zeman, krystof.zeman@oeaw.ac.at

1 Human Fertility Database

https://www.humanfertility.org

The *Human Fertility Database (HFD)* is a joint project of the Max Planck Institute for Demographic Research (MPIDR) in Rostock, Germany, and the Vienna Institute of Demography (VID) in Vienna,



and is based at the MPIDR. The HFD provides free and user-friendly access to detailed and high-quality data on period and cohort fertility, facilitating research on changes and inter-country differences in fertility in the past and in the modern era.

What's special about the HFD database?

The HFD is entirely based on official vital statistics, and places great emphasis on data checking and documentation, and on facilitating data comparability across time and countries by applying a uniform methodology. When available, the HFD also incorporates data on fertility rates by birth order, including parity-specific fertility tables that are not readily accessible from other sources.

What data are covered in the HFD?

Currently, the HFD contains data for 42 countries and regions. For each country, the HFD aims to collect the longest possible continuous time series. The length of the currently available datasets varies across countries. The longest time series of data stem from Sweden (dating to 1891).

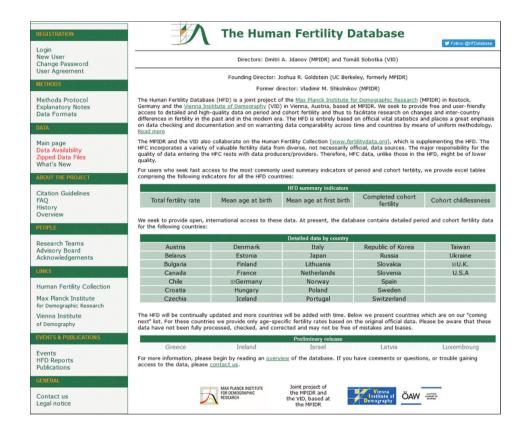
What's new in the Human Fertility Database?

Besides continuously expanding its data coverage to incorporate the latest data, the HFD is evolving to respond to the research needs of different users, and to reflect changing fertility trends.

- Adding HFD summary indicators. For users who do not need detailed data and
 appreciate having quick access to key indicators across all of the countries and
 periods covered, we are adding easily accessible Excel files with main HFD
 Summary Indicators, which are available without registration. In 2020, these
 indicators are being expanded to cover parity distribution among women born
 since 1935, and the tempo-adjusted TFR developed by Bongaarts and Feeney
 (1998).
- Adding the Short-Term Fertility Fluctuations (STFF) time series and a
 visualisation toolkit. The COVID-19 pandemic is likely to affect fertility rates.
 To give users initial insight into the unfolding changes, we are adding time
 series of the most recent monthly data on the number of live births, which will

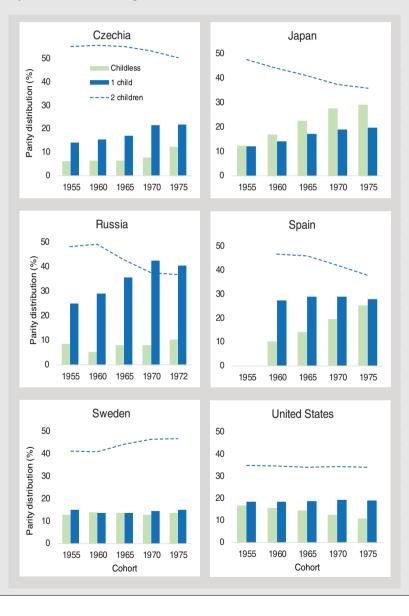
be updated each month. This monitoring should be launched in early 2021, in time to reflect the first effects of COVID-19 on births. This initiative follows a similar data series on Short-Term Mortality Fluctuations (STMF), developed in the Human Mortality Database (https://www.mortality.org).

Data for more countries are gradually being incorporated into the database. At
present, five countries are in the "Preliminary Release" category, with a limited
range of preliminary data provided to the users: Greece, Ireland, Israel, Latvia
and Luxembourg.



Highlight: Changing parity distribution: childlessness vs. having one child

One of the special features of the HFD is its focus on fertility data by birth order, including cohort parity distributions. Here we illustrate their use by plotting the changes in the share of women who are childless, have one child, or have two children by age 42. We look at six countries with different levels of and trends in parity distribution among women born in 1955–1975.



2 Human Fertility Collection

https://www.fertilitydata.org

The *Human Fertility Collection (HFC)* is a joint project of the Max Planck Institute for Demographic Research (MPIDR) in Rostock, Germany, and the Vienna Institute of Demography. The HFC has been designed to supplement the Human Fertility Database (HFD) and to incorporate a variety of international fertility data that are valuable for fertility research, but that do not meet all of the quality standards of the HFD.

What's special about the HFC database?

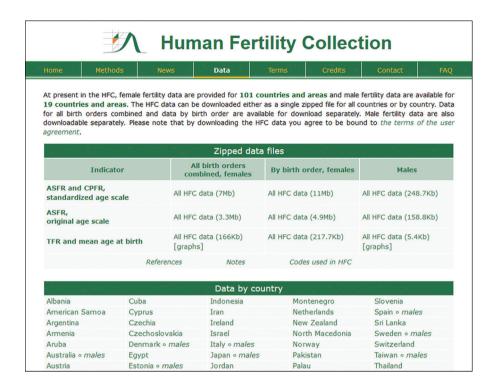
The HFC features estimates from national statistical offices and many other data sources, including survey data, statistical and scientific publications and data reconstructed by individual researchers or research teams (e.g., the European Demographic Observatory – ODE). This flexibility allows the HFC to expand the geographical coverage of the data to less developed parts of the world, and to feature historical data and estimates. At present, the HFC includes the following period fertility data specified for all birth orders combined and by birth order (when available) in a standardised format: age-specific fertility rates (ASFR), cumulative fertility rates, total fertility rates and mean ages at birth.

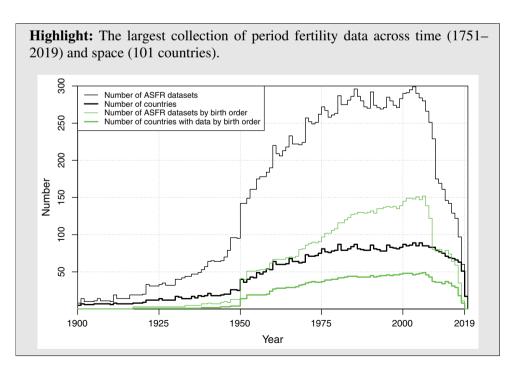
What data are covered in the HFC?

Data for 101 countries and areas are currently provided in the HFC. Data for all birth orders combined can be found from 1751 onwards (Sweden), and the earliest year for which birth order-specific data are available is 1917 (USA). For a selected number of countries, these data are available for males as well as for females.

What's new in the Human Fertility Collection?

- The latest data are being updated faster than in the Human Fertility Database. The fertility rates for the most recent available period are added soon after they are published. Currently, data for 18 countries are available up to 2019.
- Data for more countries are being gradually incorporated. Data for non-European countries are being added as well, including for Costa Rica, Indonesia, Tunisia and Sri Lanka.
- The database also aims to cover the fertility of men. Data on male fertility rates by age for 19 countries were added in November 2019.
- Other data dimensions are planned in the future, including fertility of migrants and fertility rates for sub-national regions.





3 Birth Barometer - Monitoring fertility in Austria

http://www.birthbarometer.at

The *Birth Barometer (Geburtenbarometer* in German) features upto-date data, graphs and analyses on fertility trends in Austria. Six individual presentations provide in-depth coverage of period and cohort fertility, fertility timing, migrant fertility, male fertility and fertility in Vienna. The underlying data are available for download in a single Excel file.

What's special about the Birth Barometer?

The Birth Barometer, which started in 2005, has been completely redesigned and expanded. The website now covers a wider range of topics, including male fertility and migrant fertility. In addition to the German version of the website, the Birth Barometer is now available in English in order to provide information and data to an international audience. The main indicators are computed in both an annual and a monthly (or quarterly) format. The fertility trends are covered in topical presentations that are based on simple graphs, short accompanying texts and publicly accessible datasets. All indicators are computed from the data provided by Statistics Austria.

What data are covered in the Birth Barometer?

The Birth Barometer focuses on fertility in Austria and Vienna. It features long time series (starting in 1951) of the main fertility indicators, as well as more detailed fertility indicators from 1984 onwards. Monthly data series are available from 2002.

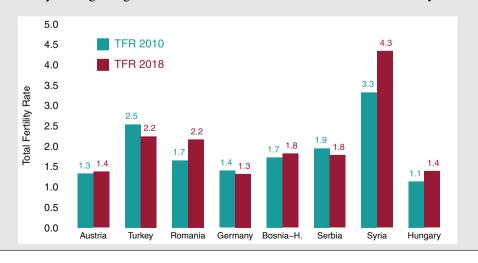
What's new in the Birth Barometer?

In 2019, the format of the Birth Barometer has been changed to make it more accessible to the wider public and to journalists. The presentations now offer a short overview of data, indicators and trends; with an emphasis on data visualisation and short explanatory texts. The topical presentations cover a range of issues, including recent fertility trends, cohort fertility rates, male fertility, migrant fertility, fertility timing and fertility in Vienna. The core indicators will be updated every year. In the future, the Birth Barometer will provide an expanded coverage of migrant fertility by country of birth. Data that reflect the impact of the COVID-19 pandemic on fertility and fertility intentions will be added as well.



Highlight: Total fertility rate by country of birth of the mother, 2010 and 2018

Almost 29,000 children born in Austria in 2018 had a mother born in another country, while 57,000 had an Austrian-born mother. There are large differences in the fertility rates of migrant women coming from different countries. Women born in Germany and Hungary have fertility rates that are almost identical to those of Austrian-born women. Women born in south-eastern Europe and Turkey have higher fertility rates, with their TFR reaching around two births per woman. Women born in Syria had very high and increasing fertility rates between 2010 and 2018. This sharp rise in fertility can be attributed to high fertility among refugee women soon after their arrival in the host country.



4 European Demographic Data Sheet

http://www.populationeurope.org

The European Demographic Data Sheet (EDS) has been published biannually since 2006. It reviews, explores and visualises recent population trends in Europe. The Data Sheet also provides a snapshot of the current research of the Wittgenstein Centre for Demography and Global Human Capital and collaborating researchers, including the key results of population projection scenarios. The online version provides an expanded selection of maps and ranking tables, as well as theme-focused texts and figures.

What's special about the Data Sheet?

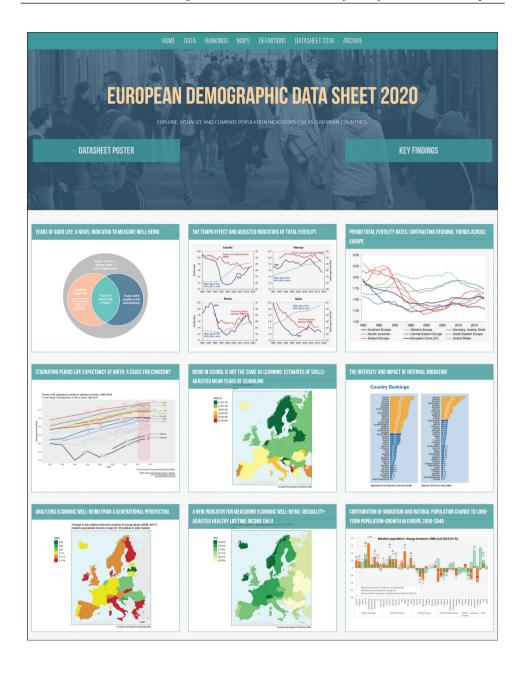
The Data Sheet covers data for all countries in Europe with a population of more than 100,000 people. It features a range of fertility indicators, including several indicators that are not available for all European countries from other sources, such as tempo-adjusted TFRs, completed cohort fertility rates and childlessness.

What data are covered in the EDS?

The European Demographic Data Sheet (EDS) covers 45 European countries, Japan and the United States. The core indicators pertain to the most recent period; most of the data in the latest (2020) Data Sheet cover the year 2018. Other indicators vary every two years, and most of them showcase recent research at the Wittgenstein Centre for Demography and Global Human Capital.

What's new in the European Demographic Data Sheet 2020?

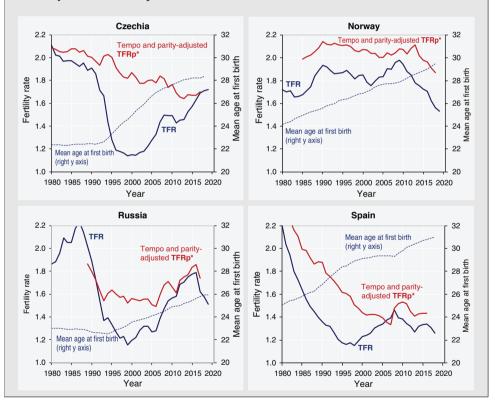
The 2020 issue puts a spotlight on health and well-being, including the indicators measuring years of life spent in good health. It also looks at the recent stagnation of life expectancy in selected countries in Europe and in the United States, the measurement of internal migration, and adjusted estimates of schooling. The online version offers users an expanded selection of maps, ranking tables and additional data. The 2022 issue of the European Demographic Data Sheet will provide initial findings on the impact of the COVID-19 pandemic on mortality and birth trends in Europe.



Highlight: Tempo- and parity-adjusted TFR

Alternative indicators to the period TFR have been developed to provide a more accurate measure of the mean number of children per woman in a calendar year, which is not affected by changes in the timing of births. The European Demographic Data Sheet features the tempo- and parity-adjusted total fertility indicator (TFRp*; Bongaarts and Sobotka 2012), which is based on age- and parity-specific fertility rates, as well as on changes in mean age at birth. When available, the TFRp* is shown for 2016. For countries lacking the required data, the Data Sheet displays the tempo-adjusted TFR (TFR-BF) proposed by Bongaarts and Feeney (1998).

Illustration: Trends in period TFR, TFRp* and mean age at first birth in Czechia, Norway, Russia and Spain



5 Wittgenstein Centre Human Capital Data Explorer

http://dataexplorer.wittgensteincentre.org

The Wittgenstein Centre Human Capital Data Explorer allows users to explore, select and download data on past reconstructions and future projections of the global population by age, sex and education, published in Lutz et al. (2018). This dataset is based on the work of a large team of researchers at the Wittgenstein Centre



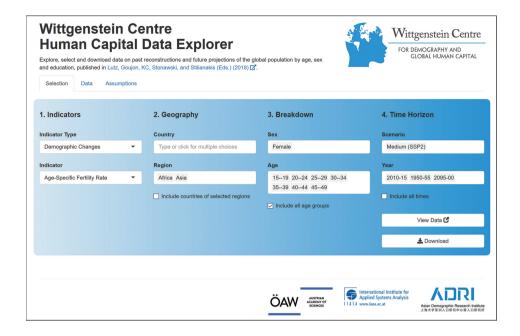
for Demography and Global Human Capital and at other institutions, and was initially developed in Lutz et al. (2014).

What's special about the Wittgenstein Centre Human Capital Data Explorer?

The Wittgenstein Centre Human Capital Data Explorer gives users easy access to a wide array of standardised data on population and human capital stocks, and demographic indicators by country, region, sex and age. The Data Explorer provides data for past periods, as well as projection scenarios and their underlying assumptions for all countries around the world. The Data Explorer also incorporates two new indicators of population ageing that go beyond the traditional conceptualisation of ageing based on fixed age boundaries: the age at which a person's remaining life expectancy is below 15 years, and the proportion of the population with a remaining life expectancy below 15 years. The Graphic Explorer allows users to visualise population pyramids and population size by education for each country and world region, and to create maps featuring the available indicators.

What data are covered in the Wittgenstein Centre Human Capital Data Explorer?

The Wittgenstein Centre Human Capital Data Explorer covers 201 countries and 29 regions for the 1950–2015 period. Data by education are shown for the population aged 15+, and are stratified by up to eight education categories. The website presents a set of five scenarios of future population and human capital trends for the projection period of 2015–2100. On top of exploring assumptions about future trends in fertility, mortality and education, the projections study the effects of several migration assumptions applied to the context of the set of Shared Socioeconomic Pathways (SSP) scenarios related to the Intergovernmental Panel on Climate Change (IPCC).

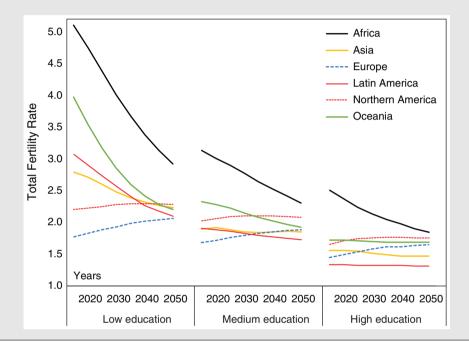


What's new in the Wittgenstein Centre Human Capital Data Explorer?

The present version (2.0), developed in 2018, benefited from the partnership with the Centre of Expertise on Population and Migration (CEPAM) at the European Commission Joint Research Centre in Italy, and with the Asian Demographic Research Institute (ADRI) in China (Lutz et al. 2018).

Highlight: Projected fertility decline by education in global regions

The figure shows the projected total fertility rate in broad world regions by level of education, based on the medium scenario for 2015–2050 developed by the Wittgenstein Centre. The *low* education category includes women with no education or primary education only; the *medium* education category covers women with lower and upper secondary education; and the *high* education category includes women with postsecondary and tertiary education.



6 Cohort Fertility and Education Database

http://www.cfe-database.org

The *Cohort Fertility and Education (CFE) database* provides high-quality data on completed cohort fertility and parity distribution by level of education (ISCED-97). The database focuses on women and



men who have completed or almost completed their family building. The data come from censuses and large sample surveys and are freely accessible to all interested users.

What's special about the CFE database?

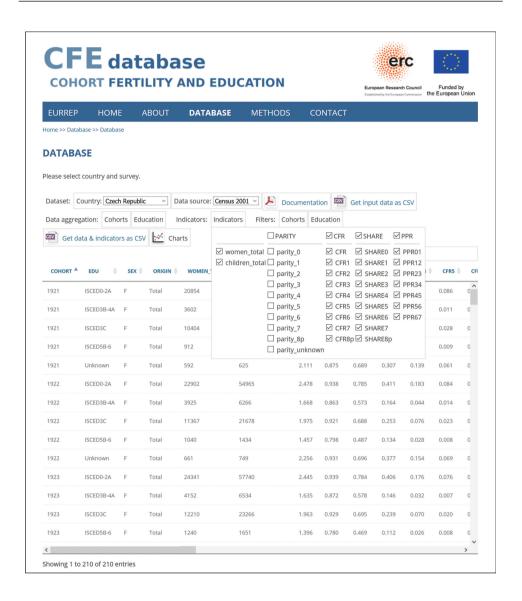
The database offers a wide range of standardised cohort fertility indicators by level of education: completed fertility rate (CFR) by birth order, share of women (and men, when available) by number of children ever born (including those who are childless), and parity progression ratios (PPR). In addition, the database includes source data on the absolute number of women (and men, when available) by birth cohort, education and number of children ever born. For some countries, data are also available by country of birth, distinguishing between two broad categories (born in the country vs. born abroad).

What data are covered in the CFE database?

Currently, the CFE database contains data from 80 censuses and surveys from 45 countries. Most of the data come from the 2001 and 2011 rounds of population censuses, with each dataset covering women (and men) aged 40 to 80 at the time of the survey or census. Regionally, the database covers both highly developed countries with relatively high levels of education and relatively low fertility (mostly in Europe), as well as selected middle-income countries, especially in Latin America. The oldest data come from the census in 1966 (South Korea), and, together with some other datasets, cover the cohorts born since the late 19th century.

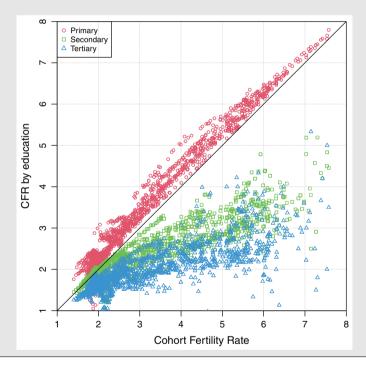
What's new in the CFE database?

In recent years, the CFE database expanded its geographical coverage and incorporated population censuses from Latin America and South-Eastern Europe, with the latter region generally being under-represented in demographic databases and in the literature. In the future, the CFE database aims to incorporate the data from the new wave of population censuses due around 2021.



Highlight: Cohort fertility by education: the link between completed fertility for all women and by education group

What is the link between the education-specific fertility rates and the overall fertility levels in a country? When plotting the data for all cohorts in the censuses and surveys covered by the CFE database, the largest contrast appears at high fertility levels. In societies with very high levels of fertility either today or in the past, (including historical data on women born in the late 19th and early 20th centuries), the fertility of women with secondary and tertiary education is typically much lower than the fertility of women with primary education only. When we look at cohorts with low fertility (typically, women in the highly developed countries born since the late 1940s), the absolute differences in fertility levels diminish.



Acknowledgements

Many thanks to Barbara Simunics for her assistance with the preparation of screenshots and illustrations.

References

- Bongaarts, J. and T. Sobotka 2012. A demographic explanation for the recent rise in European fertility. *Population and Development Review* 38(1): 83–120. https://doi.org/10.1111/j. 1728-4457.2012.00473.x
- Bongaarts, J. and G. Feeney 1998. On the quantum and tempo of fertility. *Population and Development Review* 24(2): 271–291. https://doi.org/10.2307/2807974
- European Demographic Datasheet 2020. Wittgenstein Centre, Vienna Institute of Demography and International Institute for Applied Systems Analysis, http://www.populationeurope.org
- Human Fertility Database 2020. Max Planck Institute for Demographic Research and Vienna Institute of Demography, http://www.humanfertility.org
- Human Fertility Collection 2020. Max Planck Institute for Demographic Research and Vienna Institute of Demography, http://www.fertilitydata.org
- Lutz, W., A. Goujon, S. KC, M. Stonawski and N. Stilianakis (eds) 2018. *Demographic and human capital scenarios for the 21st century: 2018 assessment for 201 countries*. Publications Office of the European Union.
- Lutz, W., W. P. Butz and S. KC (eds) 2014. *World population and human capital in the twenty-first century*. Oxford: Oxford University Press.
- Wittgenstein Centre for Demography and Global Human Capital 2018. Wittgenstein Centre Data Explorer Version 2.0 (Beta), http://www.wittgensteincentre.org/dataexplorer
- Zeman, K., Z. Brzozowska, T. Sobotka, E. Beaujouan and A. Matysiak 2017. Cohort Fertility and Education Database, http://www.cfe-database.org
- Zeman, K., T. Sobotka, R. Gisser and M. Winkler-Dworak 2019. Birth Barometer: Monitoring Fertility in Austria. Vienna Institute of Demography, http://www.birthbarometer.at

Open Access This article is published under the terms of the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/) that allows the sharing, use and adaptation in any medium, provided that the user gives appropriate credit, provides a link to the license, and indicates if changes were made.