Using dynamic microsimulation to assess the effect of changing population structures on long-term trends in the period TFR

Karel Neels, University of Antwerp (Belgium), Karel.Neels@uantwerpen.be

Introduction

The period TFR continues to be the most widely used aggregate-level indicator of period fertility, which feeds directly into aggregate-level models of population dynamics (e.g., cohort-component projection) that are commonly used (European Commission, 2018; United Nations, 2017). A myriad of factors have been identified that affect the tempo and quantum of period fertility such as shift in the parity distribution of fertility (Ryder, 1980), increasing education (Neels et al., 2017; Ni Bhrolchain & Beaujouan, 2012), economic cycles (Neels et al., 2013; Sobotka et al., 2011) or the increasing diversity of the population in terms of migrant background to name but a few (Kulu et al., 2017), but the effects of such factors on the period TFR have been difficult to quantify for several reasons. Not only does a lack of data usually prevent breaking down numerators and denominators of age-specific fertility rates consistently by a larger number of characteristics, also the unidimensional (direct) standardization for age has been shown to yield problematic results when applied to specific subgroups (e.g., migrant groups). The problems encountered when attempting to link the period TFR to potential determinants of period fertility also imply that quantifying the impact of anticipated changes in such factors – e.g., the unfolding diversity by migrant background in many European societies – on future trends in the period TFR is highly problematic. As a result, analysis of the period TFR is incapable of accurately informing projection sets of the potential variation in period fertility over time induced by various determinants of family formation.

The limitations of the period TFR to link period fertility to a large number of determinants, while accounting for multiple time clocks that influence fertility over the life course, suggest that individual-based model may be better suited to model trends in aggregate-level fertility (Billari, 2015). Event-history or hazard models have been widely used to link tempo and quantum of various life course transitions (e.g., having a first child) to potential determinants at the individual, household or contextual level (Putter et al., 2007; Singer & Willett, 2003; Vikat et al., 2007), but few attempts have been made to integrate entry into parenthood and subsequent parity progression into a single model that can subsequently be used for dynamic microsimulation of trajectories of family formation and – through aggregation of individual trajectories by age and year – simulation of conventional aggregate-level fertility indicators. Using population-wide longitudinal microdata for Belgium and a late entry design for the period 2001-2010, this paper develops a compartmental hazard model of entry into parenthood and subsequent parity progression which feeds into a dynamic microsimulation model that generates maternity histories for individual women aged 15-50 between 2011-2070. Finally, by aggregating individual trajectories to aggregate-level fertility indicators, the event logs resulting from microsimulation models incorporating different determinants are used to quantify the impact of anticipated changes in these determinants on long-term trends of aggregate-level fertility measures such as the period TFR.

Data and methods

Population-wide census and register-data

The analysis uses retrospective microdata from the 2011 Belgian Census which was complemented retrospectively (for the period 2000-2010) and prospectively (for the period 2011-2022) with longitudinal microdata from the population registers. The linked census-register data provide population-wide longitudinal microdata covering all legal residents in the population over an extended follow-up period from 1985 up to 2022. Second, the data provide annual information on household composition through a coded identifier of the household which allows to identify all individuals belonging to the same household and which allows to observe household transitions over the life course of all household members involved (e.g. birth of a child in the household). The data drawn from the population register provide individual-level data on first nationality, place of birth and year of immigration which allows to identify migrants of the first generation (immigrated as adults) and the intermediate or 1.5 generation (immigrated as children), further differentiated by country of origin and duration of residence. In addition, the population register provides individual-level data on descent (cf. coded identifier of each parent and grandparent) which allows to identify second and later generations of migrants, but also allows to reconstruct kinship networks. The 2011 Census provides information on the highest level of education obtained in the period 2000-2010, whereas information on education is updated throughout the observation period using data from the registry of educational certificates granted by the regional educational systems in Belgium.

Hazard and microsimulation models

The microsimulation consists of three successive steps (Figure 1). First, a late-entry compartmental hazard model of entry into parenthood and parity progression is estimated using observed fertility histories for women aged 15-50 between 2001 and 2010 (Figure 1.a). Women entering the observation window on January 1st 2001 are heterogenous with respect to both age and parity, and are followed until they reach their 50th birthday, emigrate, decease or reach the end of the observation period on December 31st 2010. Women enter the compartment of the model corresponding to their parity when entering the observation period and progress to other compartments of model when having a (or more) child(ren). In the baseline model, the hazard of having a first child is modelled as a 4th-order polynomial function of age (centred at age 14), whereas progression to second and higher-order is modelled as a function of i) parity, ii) duration since index birth (step function allowing point estimates for the first seven years after the index birth and a linear specification thereafter) and age at index birth (quadratic function). Separate parameter estimates are allowed for second up to seventh-order births, while eighth and higher-order births are pooled into a single compartment of the model. More elaborated models further consider the impact of i) enrolment in education and highest level of education (Wood et al., 2014), ii) variation in economic and policy contexts (Neels et al., 2013), and iii) migrant background (origin, migrant generation and duration of residence) (Kulu et al., 2017):

$$
\begin{cases}
\text{P0:} \quad \int_t^{t+1} h(a)da = \begin{matrix} e^{\hat{\alpha}}.\, e^{\hat{F}(DL_{ti})}.\, e^{\hat{F}(AL_i)}.\, e^{\hat{F}(DL_{ti},AL_i)}.\, e^{\hat{F}(EF_i)+\hat{F}(EF_i,DL_{ti},AL_i)}. \\ e^{\hat{F}(M_i)+\hat{F}(M_i,EF_i,DL_{ti},AL_i)}.\, e^{\hat{F}(P_i)+\hat{F}(P_i,EF_i,DL_{ti},AL_i)}.\, e^{\hat{F}(M_i,P_i,EF_i,DL_{ti},AL_i)}. \\ e^{\hat{F}(EC_{ti})+\hat{F}(EC_{ti},EF_i,DL_{ti},AL_i,M_i,P_i)}.\, e^{\hat{F}(PC_{ti})+\hat{F}(PC_{ti},EF_i,DL_{ti},AL_i,M_i,P_i,EC_{ti})}. \end{matrix} + \alpha_i \\
\text{P1:} \quad \int_t^{t+1} h(a)da = \begin{matrix} e^{\hat{\alpha}}.\, e^{\hat{F}(DL_{ti})}.\, e^{\hat{F}(AL_i)}.\, e^{\hat{F}(DL_{ti},AL_i)}.\, e^{\hat{F}(EF_i)+\hat{F}(EF_i,DL_{ti},AL_i)}. \\ e^{\hat{F}(M_i)+\hat{F}(M_i,EF_i,DL_{ti},AL_i)}.\, e^{\hat{F}(P_i)+\hat{F}(P_i,EF_i,DL_{ti},AL_i)}.\, e^{\hat{F}(M_i,P_i,EF_i,DL_{ti},AL_i)}. \\ e^{\hat{F}(EC_{ti})+\hat{F}(EC_{ti},EF_i,DL_{ti},AL_i,M_i,P_i)}.\, e^{\hat{F}(PC_{ti})+\hat{F}(PC_{ti},EF_i,DL_{ti},AL_i,M_i,P_i,EC_{ti})}. \end{matrix} + \alpha_i \\
\qquad\qquad\qquad\qquad \vdots \\
\text{P8+:} \quad \int_t^{t+1} h(a)da = \begin{matrix} e^{\hat{\alpha}}.\, e^{\hat{F}(DL_{ti})}.\, e^{\hat{F}(AL_i)}.\, e^{\hat{F}(DL_{ti},AL_i)}.\, e^{\hat{F}(EF_i)+\hat{F}(EF_i,DL_{ti},AL_i)}. \\ e^{\hat{F}(M_i)+\hat{F}(M_i,EF_i,DL_{ti},AL_i)}.\, e^{\hat{F}(P_i)+\hat{F}(P_i,EF_i,DL_{ti},AL_i)}.\, e^{\hat{F}(M_i,P_i,EF_i,DL_{ti},AL_i)}. \\ e^{\hat{F}(EC_{ti})+\hat{F}(EC_{ti},EF_i,DL_{ti},AL_i,M_i,P_i)}.\, e^{\hat{F}(PC_{ti})+\hat{F}(PC_{ti},EF_i,DL_{ti},AL_i,M_i,P_i,EC_{ti})}. \end{matrix} + \alpha_i
\end{cases}
$$

where DL_{ti} reflects duration since leaving education, AL_i the age at leaving education, EF_i the level (and potentially field) of education, M_i a typology of migrant generation and origin group, P_i parental background, and EC_{ti} and PC_{ti} reflect distributed lags of economic context and policy context respectively. The model compartments include relevant 2-way and higher-order interactions between individual characteristics, parental background, economic and policy contexts. Model specifications are calibrated using observed fertility histories for the period 2011-2020.

In a second stage, the compartmental model of entry into parenthood and parity progression is used to simulate individual fertility histories of women age 15-50 on a year-by-year basis between the first of January 2021 and December 31st 2070 (Figure 1.b). As children enter the risk set throughout the simulation period when turning 14 years of age, the starting population for the simulation includes all females in the population between ages 0 and 50 on January 1st 2021. Women are no longer considered at risk of entering parenthood or having a child when they turn 50.

Finally, the third stage uses the event logs that where generated in the second stage using dynamic microsimulation (Figure 1.c). By aggregating births by age of the mother (and birth-order) as well as women at risk by age on an annual basis, age-(order-)specific fertility rates are generated on an annual basis from which the conventional period TFR and order-specific TFR_i can be calculated. Using event logs from microsimulation models incorporating different determinants of fertility, we quantify the impact of anticipated changes in these determinants on aggregate-level fertility indicators such as the period TFR over the period 2021-2070. In line with previous research, we consider variation in the period TFR associated with economic prospects, ongoing educational expansion and unfolding population heterogeneity by migration background.

References

Billari, F. C. (2015). Integrating macro- and micro-level approaches in the explanation of population change [Article]. *Population Studies-a Journal of Demography*, *69*, S11-S20. https://doi.org/10.1080/00324728.2015.1009712

European Commission. (2018). *The 2018 Ageing Report: Economic and Budgetary Projections for the EU Member States (2016-2070)*. European Commission, Institutional Paper 079.

Kulu, H., Hannemann, T., Pailhe, A., Neels, K., Krapf, S., Gonzalez-Ferrer, A., & Andersson, G. (2017). Fertility by Birth Order among the Descendants of Immigrants in Selected European Countries [Article]. *Population and Development Review*, *43*(1), 31-60. <Go to ISI>://WOS:000397855600002

Neels, K., Murphy, M., Ni Bhrolchain, M., & Beaujouan, E. (2017). Rising Educational Participation and the Trend to Later Childbearing. *Population and Development Review*, *43*(4), 667-693.

Neels, K., Theunynck, Z., & Wood, J. (2013). Economic recession and first births in Europe: recession-induced postponement and recuperation of fertility in 14 European countries between 1970 and 2005. *International Journal of Public Health*, *58*, 43-55. https://doi.org/http://dx.doi.org/doi:10.1007/s00038-012-0390-9

Ni Bhrolchain, M., & Beaujouan, E. (2012). Fertility postponement is largely due to rising educational enrolment. *Population Studies*, *66*(3), 311-327.

Putter, H., Fiocco, M., & Geskus, R. B. (2007). Tutorial in biostatistics: Competing risks and multi-state models [Article]. *Statistics in Medicine*, *26*(11), 2389-2430. https://doi.org/10.1002/sim.2712

Ryder, N. (1980). Components of Temporal Variations in American Fertility. In R. W. Hiorns (Ed.), *Demographic Patterns in Developed Societies*. Taylor & Francis.

Singer, J., & Willett, J. (2003). *Applied longitudinal data analysis. Modeling change and event occurrence.* Oxford University Press.

Sobotka, T., Skirbekk, V., & Philipov, D. (2011). Economic Recession and Fertility in the Developed World [Article]. *Population and Development Review*, *37*(2), 267-306. https://doi.org/10.1111/j.1728-4457.2011.00411.x

United Nations, Department of Economic and Social Affairs, Population Division. (2017). *World Population Prospects: The 2017 Revision, Methodology of the United Nations Population Estimates and Projections* (U. Nations, Ed. Vol. Working Paper No. ESA/P/WP.250). United Nations.

Vikat, A., Speder, Z., Beets, G., Billari, F. C., Bühler, C., Désesquelles, A., Fokkema, T., Hoem, J. M., MacDonald, A., Neyer, G., Pailhé, A., Pinnelli, A., & Solaz, A. (2007). Generations and Gender Survey (GGS): Towards a Better Understanding of Relationships and Processes in the Life Course [Research Material]. *Demographic Research*, *17*, 51.

Wood, J., Neels, K., & Kil, T. (2014). The Educational Gradient of Childlessness and Cohort Parity Progression in 14 Low Fertility Countries. *Demograhpic Research*, *31*(46), 1365-1416.

Figure 1. Analytical steps in dynamic microsimulation of the period TFR.

Figure 1.a Model estimation (2001-2010) and calibration (2011-2020) using late entry design

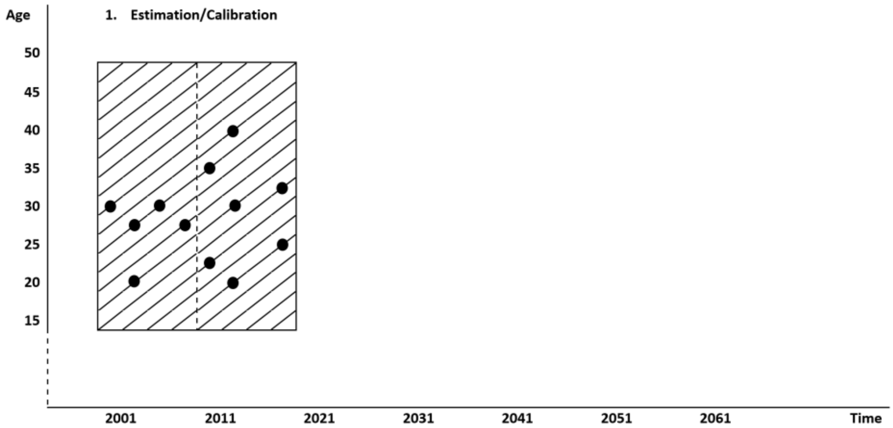

Figure 1.b Prospective microsimulation of fertility histories for women between age 15-50