Fertility Projection Using Hazard-Based Microsimulation Models, with an Application to Belgian Fertility, 1960-2015.

Karel Neels & Jonas Wood, University of Antwerp

1. Background

Several factors have been identified as potential drivers of fertility postponement and the decline of period fertility in recent decades, including the introduction and diffusion of contraceptive technology; rising enrolment and educational attainment; increasing labour force participation (e.g. incompatibility between childcare and labour force participation, expected increase in earnings and wage penalties, loss of training opportunities and depreciation of job-specific human capital); a shift to an individualistic family model (referring to the changing role of children with children becoming the locus of emotional and financial investment); rising gender equity in educational outcomes and labour force participation under institutional constraints (e.g. lack of childcare and supporting family policies; low benefit levels and/or rigid labour markets); variation in economic context and housing markets; changing partnership patterns (e.g. higher separation/divorce, lack of suitable partners as a result of hypogamy/hypergamy and rising education), as well as declining real wages (Blossfeld, Klijzing, Mills, & Kurz, 2005; Lesthaeghe & Neels, 2002; McDonald, 2006; Mills, Rindfuss, McDonald, Velde, & Force, 2011).

Despite the abundance of candidate causal factors, little quantitative evidence is available on the contribution of specific determinants to aggregate change in tempo and quantum of period fertility by birth order (Neels, Murphy, Ni Bhrolchain, & Beaujouan, 2017; Ni Bhrolchain & Beaujouan, 2012). Using exhaustive longitudinal microdata from the 2001 Belgian census and the national population registers, this paper uses discrete-time hazard models to document the contributions of rising educational enrolment, changing patterns of union formation as well as variation in both economic conditions and family policies, to explaining aggregate change in tempo and quantum of first births as well as progression to second and higher-order births between 1960 and 2000. Subsequently, discrete-time hazard models for entry into parenthood and progression to second and higher-order births are integrated into a microsimulation framework to generate prospective out-of-sample predictions of aggregate fertility trends by birth order, and assess whether and to what extent variation in the aforementioned factors is capable of accounting for trends in Belgian order-specific fertility between 2000 and 2015.

2. Data & Methods

Longitudinal microdata and contextual indicators

The analyses use longitudinal microdata from the 2001 census that provide information on i) age at last attended education and highest level of education, ii) date of first entry into a co-residential union with a partner and first marriage, as well as iii) full fertility histories for all women aged 14 and older (Deboosere & Willaert, 2004). Validation against vital registration has shown that period and cohort indicators estimated retrospectively from the 2001 census agree well with national statistics on period fertility trends between 1960 and 2000, as with independent estimates of cohort fertility patterns for women born after 1930 (Neels & Gadeyne, 2010). Apart from the longitudinal microdata drawn from the census, the analyses use time-series data on harmonized unemployment rates and consumer price index between 1960 and 2000 drawn from OECD.

Discrete-time hazard models of entry into parenthood and parity progression

Discrete-time random-effects hazard models are used to estimate the effects of individual-level covariates and contextual variables on first birth hazards of women aged 15-49 years between 1960 and 2000, as well as parity progression to second and higher-order births. For entry into parenthood, individual level covariates include time-varying age (centered at age 15 in years, quadratic effect), a time-varying dummy-variable indicating enrolment in education, duration since leaving full-time education (in years, quadratic effect), highest level of education, a dummy variable indicating whether women have ever entered a co-residential union and duration since first union (in years, quadratic effect). An interaction between education and time-varying age is included to allow interaction between age and the baseline hazard function. Apart from the longitudinal microdata drawn from the 2001 census, time-series data on harmonized unemployment rates between 1960 and 2000 drawn from OECD are included as an indicator of economic and labour market context. The unemployment rates have been included into the model with lags varying from 1 to 10 years to reflect the direct negative effect of adverse labour market conditions and the positive recuperation effect at longer time lags respectively (Neels, Theunynck, & Wood, 2013). Finally, to assess whether age groups are differentially affected by variation in economic context, cross-level interactions between 5-year age-groups and the aggregate-level unemployment rates are included in the model.

Models for parity progression to second and higher-order births are similar to the model for first births, but duration since index birth is now used as the exposure dimension and women's age at index birth is included to account to for variation in second and higher-order births as a result of postponement of parenthood. As selection into parenthood has been found to occur as a result of reaching of stable employment and/or income positions (Neels & De Wachter, 2010), variation in harmonized unemployment rates does not significantly affect progression to second and higher-order births significantly. In contrast, variation in consumer price index as an indicator of the variation of real income of households has been found to significantly affect progression to second and higher-order births regardless of women's level of education.

Contribution to aggregate change in tempo and quantum of order-specific fertility

To estimate the contribution of various factors on aggregate change in order-specific mean ages at childbearing ($\mathrm{MAC}i$) and synthetic parity progression ratios ($\mathrm{SPPR}i$) over the period considered, fitted or estimated birth hazards for women aged 15-49 are retrieved from the discrete-time hazard models for single calendar years between 1960 and 2000 which were used to construct synthetic life tables of entry into parenthood and parity progression to second and higher-order births. Subsequently, MAC1 and SPPR1 are derived from the period life tables, which amounts to indirect standardization of MAC1 and SPPR1 for the individual-level and contextual variables included in the model (Neels et al. 2014). Finally, zero-order correlations were calculated between observed time-series of MAC1 and SPPR1 and series derived from the models, as well as correlations between both series after first-order differencing. The latter provide more conservative estimates of the correspondence between observed and estimated series than the more conventional zero-order correlations.

Fertility projection using microsimulation models

The parameter estimates of the discrete-time hazard models for entry into parenthood and progression to second and higher-order births are subsequently applied to women under age 50 in the 2001 Census to generate individual-level fertility outcomes on an annual basis for the period 2000-2015. The simulation for the period 2000-2015 uses the observed harmonized unemployment rate and consumer price index as exogenous inputs. The simulated births and women exposed to risk are subsequently aggregated to generate conventional period fertility indicators such as the order-specific period TFRs (and associated mean ages at childbearing), the overall period TFR (and mean age at childbearing), synthetic parity progression ratios (and associated synthetic birth intervals).

Since the focus of the microsimulation is projection of order-specific fertility trends for the period 2000-2015 using the nested hazard models estimated for the period 1960-2000, information on mortality and emigration is taken from the register follow-up of the 2001 Census throughout the simulation period. Similarly, information on increasing enrolment in education, as well as information on age, sex, education and parity of the female immigrant population is taken from the 2011 Census and linked register data for the period 2000-2015.

3. Results

Hazard models for entry into parenthood and parity progression, 1960-2000

Figure 1a compares the observed trend in SPPR1 between 1960 and 2000 with the simulated trend from a model that incorporates the effect of changing educational careers on entry into parenthood through lengthening educational careers and rising enrolment at older ages, rising levels of education and the fact that young adults at a given age have typically left education more recently in the 1980s and 1990s than was the case in the 1960s and 1970s, resulting in a younger social age (Neels et al., 2017; Ni Bhrolchain & Beaujouan, 2012; Skirbekk, Kohler, & Prskawetz, 2004). According to the model, rising education through the various pathways mentioned above entails a decrease in SPPR1 – and an associated rise in synthetic childlessness – that corresponds quite well to the observed trend, but the model fails to account for the period variation around this structural shift. Similarly, rising enrolment in education accounts for a structural increase in the mean age at first birth, but the model fails to account for the acceleration and deceleration of fertility postponement over the period considered.

Figure 2 shows the cross-correlations between the annual time-series of the harmonized unemployment rates and the observed parity progression ratios for first to fourth births. The correlations were obtained after first-order differencing of both time-series to eliminate trends which would otherwise typically inflate cross-correlations. As a result of first-order differencing, the cross-correlations thus correlate year-to-year change in the unemployment rate to year-to-year change in the synthetic parity progression ratio. Lags of different years have been implemented to verify whether and at which (distributed) lags unemployment rates should be included as contextual variables in hazard models of entry into parenthood and parity progression. Negative lags imply in this case that change in current synthetic parity progression ratios is correlated to change in future unemployment rates. Apart from the synthetic parity progression to third birth among women with higher secondary education, correlations with negative lags are never significant. Significant correlations at negative lags could arise as a result of autocorrelation in the time series of unemployment – as a result of which current levels to some extent predict future levels – but the autocorrelation has not been eliminated from the time-series in Figure 2. Cross-correlations at a lag of zero correlate change in synthetic parity progression ratios to change in unemployment levels in the same year. The cross-correlations at zero lags are typically not significant, which is most likely due to the fact that conceptions and the decision to have a child occurs roughly a year before the birth. Cross-correlations that link current change in synthetic parity progression ratios to past variation in unemployment levels (positive lags in Figure 2) should therefore be more relevant, which is confirmed by the results in Figure 2. Variation in unemployment rates is significantly correlated with variation in first births, but not second and higher-order births. Variation in SPPR1 is correlated to variation in unemployment rates one to three years earlier, suggesting the unemployment rates should be included at lags of one year, or alternatively distributed lag models could be implemented.

Figure 3 compares the observed trend in the synthetic parity progression ratio to a first birth between 1960 and 2000 to the simulated trend from a model that includes age (cubic) and the unemployment rate with a lag of one year. The latter model introduces the temporal variation that is lacking in Figure 1, but the variation is at the same time excessive relative to the observed series. With respect to the mean age at entry in parenthood, the simulated trends suggest acceleration and deceleration in response to economic conditions, but the simulated trends now lacks the structural shift that was apparent from Figure 1. As unemployment predominantly affect younger adults entering the labour market (Neels et al., 2013), Figure 4 compares the observed trend in SPPR1 with the simulated trend from a model that additionally includes an interaction between five-year age groups and the lagged unemployment rate, allowing the effect of unemployment to taper with increasing age. The period variation in the simulated series from this model now corresponds closely to the observed temporal variation in the synthetic parity progression ratio to first births, but the simulated trend in the period mean age at first birth does not appropriately pick up the postponement of first births after 1975, suggesting that the factors accounting for such a structural shift should be included in the model.

Figure 5 compares the observed trend in the synthetic parity progression ratio to first births to the simulated trend from a more elaborate model including the factors accounting for rising education, the indicators relating to entry into a co-residential union, as well as the unemployment rate and the interaction between the latter and age. Due to the trend in both the observed and the simulated synthetic parity ratio to first births, the zero-order correlation between both time-series equals 0.93 - implying that 87 per cent of the variance in the synthetic parity progression ratio is accounted for by variation un unemployment – whereas the correlation between the first-order differenced time-series, which correlates change in synthetic parity progression to change in unemployment, equals 0.40 and the average absolute deviation between both time-series amount 0.0125, implying that the model predicts the observed series with an average error of 1,25 per cent (a proportional reduction in error of 67 per cent relative to the null model including no covariates). Similarly, the model predicts the period mean age at first birth with an average error of 0.34 (4 months), a proportional reduction in error of 70 per cent relative to a constant only null model.

Previous models have not yet taken into account that recuperation at older ages typically takes place as a result of recession-induced postponement at younger ages (Neels et al., 2013). Figure 6 compares the observed trend in the synthetic parity progression ratio to first births to the simulated trend additionally allowing for fertility recuperation at lags of 10 years. For the synthetic parity progression ratio the average error is further reduced to 0.0095, implying that the model predicts the SPPR1 with an average error of less than 1 per cent over a 40 year observation period. Similarly, the model predict the mean age at first birth with an average error of 0.28 years (3,3 months).

As shown in Figure 2, the harmonized unemployment rate does not significantly affect synthetic parity progression to second, third and fourth-order births. Figure 7 plots the cross-correlations between the consumer price index and synthetic parity progression ratios to first up to fourth births. In contrast to unemployment, consumer price index has no effect on first births, but significantly affects progression to second, third and fourth births at a lag of one year, typically corresponding to the conception of the births considered.

Figure 8a compares the observed trend in the synthetic parity progression ratio to second births to the simulated trend from a model including duration since index birth, women's age at first birth, educational variables (time-varying enrolment, educational level and duration since leaving education), partnership variables (ever having entered a co-residential union and duration since first entry into a co-residential union) and temporal variation in consumer price index (allowing a lag of one year). Similar to the models for first births, the simulated trend provides a close approximation of the observed trend in synthetic parity progression to second births. For third births, Figure 8b compares the observed trend in synthetic parity progression ratio to third births to the simulated trend from a model including duration since index birth, women's age at first birth, educational variables (time-varying enrolment, educational level and duration since leaving education), partnership variables (ever having entered a co-residential union and duration since first entry into a co-residential union) and temporal variation in consumer price index (allowing a lag of one year). Provided that dummy variables are additionally included to account for the unprecedented drop in SPPR3 between 1965 and 1975 (Neels, 2006), the model provides a close approximation of the observed trend. Models with a similar specification have been set up for fourth and higher order births, despite the dwindling importance of these higher-order births in period fertility after 1975 (Neels, 2006).

Out of sample simulation for the period 2000-2015

Previous results indicate the discrete-time hazard models developed for entry into parenthood and progression to second and higher order births provide allow accurate predictions of the observed trends between 1960 and 2000. The out of sample microsimulations for the period 2000-2015 based on the hazard models for entry into parenthood and parity progression will be presented at the European Population Conference 2020.

4. References

Blossfeld, H. P., Klijzing, E., Mills, M., & Kurz, K. (2005). *Globalization, Uncertainty and Youth in Society*. London: Routledge.

Deboosere, P., & Willaert, D. (2004). *Codeboek algemene socio-economische enquête 2001*. Retrieved from

Lesthaeghe, R., & Neels, K. (2002). From the first to the second demographic transition: An Interpretation of the Spatial Continuity of Demographic Innovation in France, Belgium and Switzerland. *European Journal of Population, 18*, 325-360.

McDonald, P. (2006). Low fertility and the state: The efficacy of policy. *Population and Development Review, 32*(3), 485-+. doi:10.1111/j.1728-4457.2006.00134.x

Mills, M., Rindfuss, R. R., McDonald, P., Velde, E. T., & Force, E. R. S. T. (2011). Why do people postpone parenthood? Reasons and social policy incentives. *Human Reproduction Update, 17*(6), 848-860. doi:10.1093/humupd/dmr026

Neels, K. (2006). *Reproductive strategies in Belgian fertility: 1930-1990.* (Vol. 38). Brussels and The Hague: NIDI-CBGS Publications.

Neels, K., & De Wachter, D. (2010). Postponement and recuperation of Belgian fertility: how are they related to rising female educational attainment? *Vienna Yearbook of Population Research, Vol. 8*, 77-106.

Neels, K., & Gadeyne, S. (2010). Het effect van opleidingsniveau op tempo en kwantum van rangspecifieke vruchtbaarheid in België: 1960-2000. In K. Matthijs, B. Van de Putte, J. Kok, & H. Bras (Eds.), *Leven in de lage landen. Historisch-demografisch onderzoek in Vlaanderen en Nederland.* Leuven/Den Haag: Acco.

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.

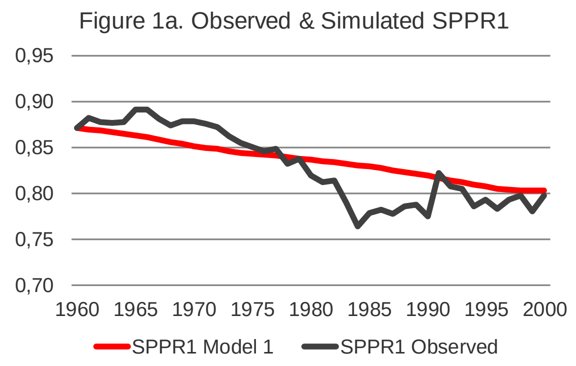
Neels, K., Theunynck, Z., & Wood, J. (2013). Economic recession and first births in Europe: recession-induced postponement and recuperation of fertility in 14 European countries between 1970 and 2005. *International Journal of Public Health, 58*, 43-55. doi:http://dx.doi.org/doi:10.1007/s00038-012-0390-9

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

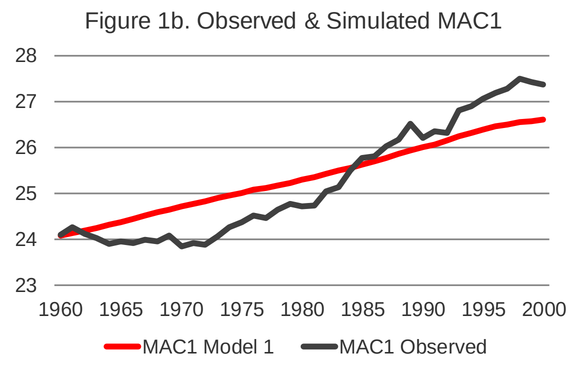
Skirbekk, V., Kohler, H. P., & Prskawetz, A. (2004). Birth month, school graduation, and the timing of births and marriages. *Demography, 41*(3), 547-568. doi:10.1353/dem.2004.0028

Appendix: Tables & Figures

Figure 1. Estimated synthetic parity progression ratio to first births and period mean age at first births from the hazard model including centred age (cubic), time-varying enrolment, educational level and duration since leaving education (quadratic), Belgium, 1960-2000.



R = 0,8694
R^2 = 0,7559
Rdif = 0,1039
|e| = 0,0211 (0,0380 under Constant Schedule)

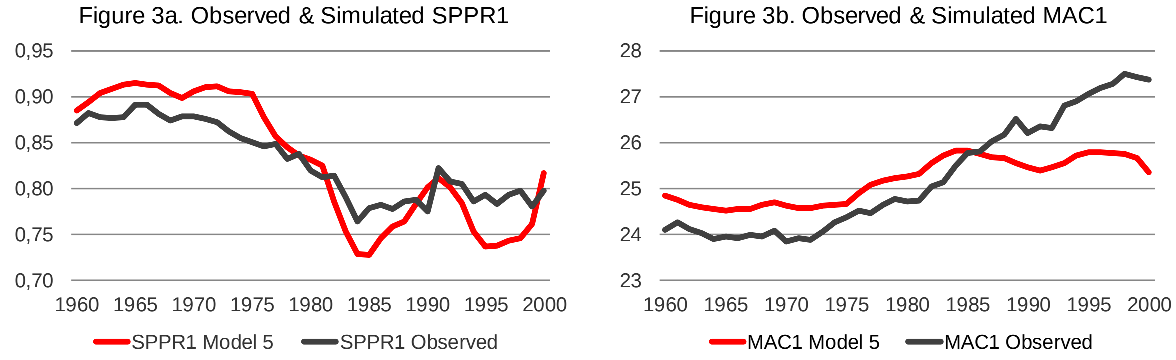


R = 0,9567
R^2 = 0,9152
Rdif = 0,3094
|e| = 0,5013 (1,1247 under Constant Schedule)

Figure 2 Cross-correlations between annual time-series of harmonized unemployment rates and observed synthetic parity progression ratios (first to fourth births) after first-order differencing and considering negative lags (births correlated to future unemployment rates), a zero lag (births correlated to unemployment rates in the same year) and positive lags (births related to past unemployment rates), Belgium, 1960-2000.

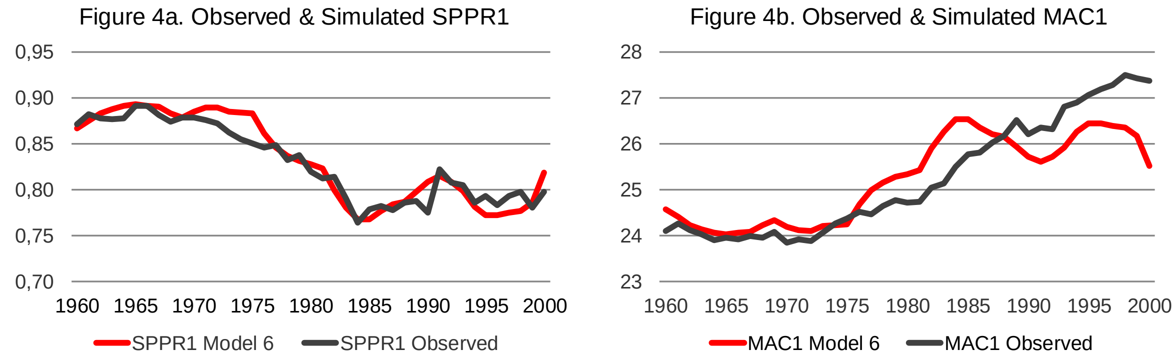
Source: 2001 Belgian Census, calculations by authors. Note: Figures with significant cross-correlations have a gray background.

Figure 3. Estimated synthetic parity progression ratio to first births and period mean age at first birth from the hazard model including centred age (cubic) and the harmonized unemployment rate with a lag of one year, Belgium, 1960-2000.



R	= 0,9561	R	= 0,8644
R^2	= 0,9251	R^2	= 0,7471
Rdif	= 0,3843	Rdif	= 0,3689
$\|e\|$	= 0,0266	$\|e\|$	= 0,7457

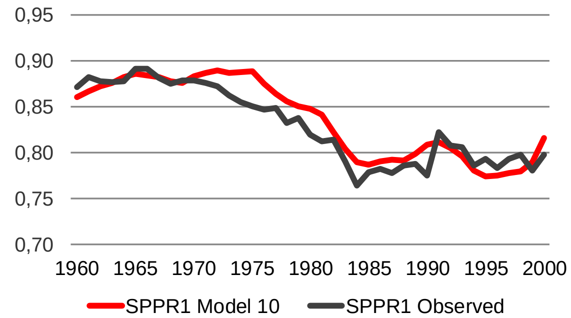
Figure 4. Estimated synthetic parity progression ratio to first births and period mean age at first birth from the hazard model including centred age (cubic), the harmonized unemployment rate with a lag of one year and interaction between unemployment rate and 5-year age groups, Belgium, 1960-2000.



R	= 0,9618	R	= 0,8638
R^2	= 0,9142	R^2	= 0,7462
Rdif	= 0,3961	Rdif	= 0,3698
$\|e\|$	= 0,0095	$\|e\|$	= 0,5090

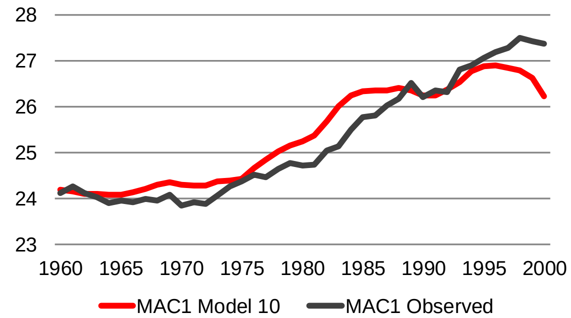
Figure 5. Estimated synthetic parity progression ratio to first births and period mean age at first birth from the hazard model including centred age (cubic), educational variables, partnership variables and the harmonized unemployment rate with a lag of one year (incl. the interaction between unemployment rate and 5-year age groups), Belgium, 1960-2000.

Figure 5a. Observed & Simulated SPPR1



R = 0,9329
R^2 = 0,8703
Rdif = 0,4009
|e| = 0,0125

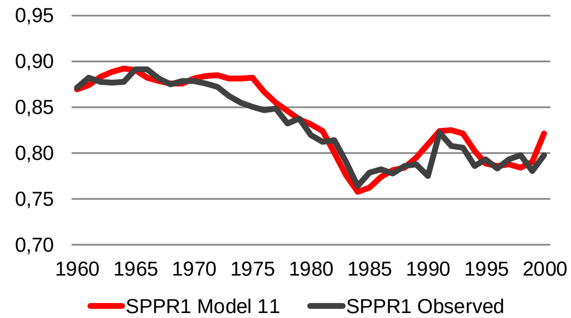
Figure 5b. Observed & Simulated MAC1



R = 0,9515
R^2 = 0,9053
Rdif = 0,4260
|e| = 0,3422

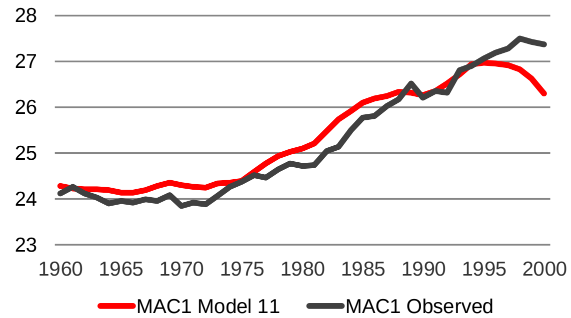
Figure 6. Estimated synthetic parity progression ratio to first births and period mean age at first birth from the hazard model including centred age (cubic), educational variables, partnership variables and the harmonized unemployment rate with a lags of one and 10 years (incl. the interaction between unemployment rate and 5-year age groups), Belgium, 1960-2000.

Figure 6a Observed & Simulated SPPR1



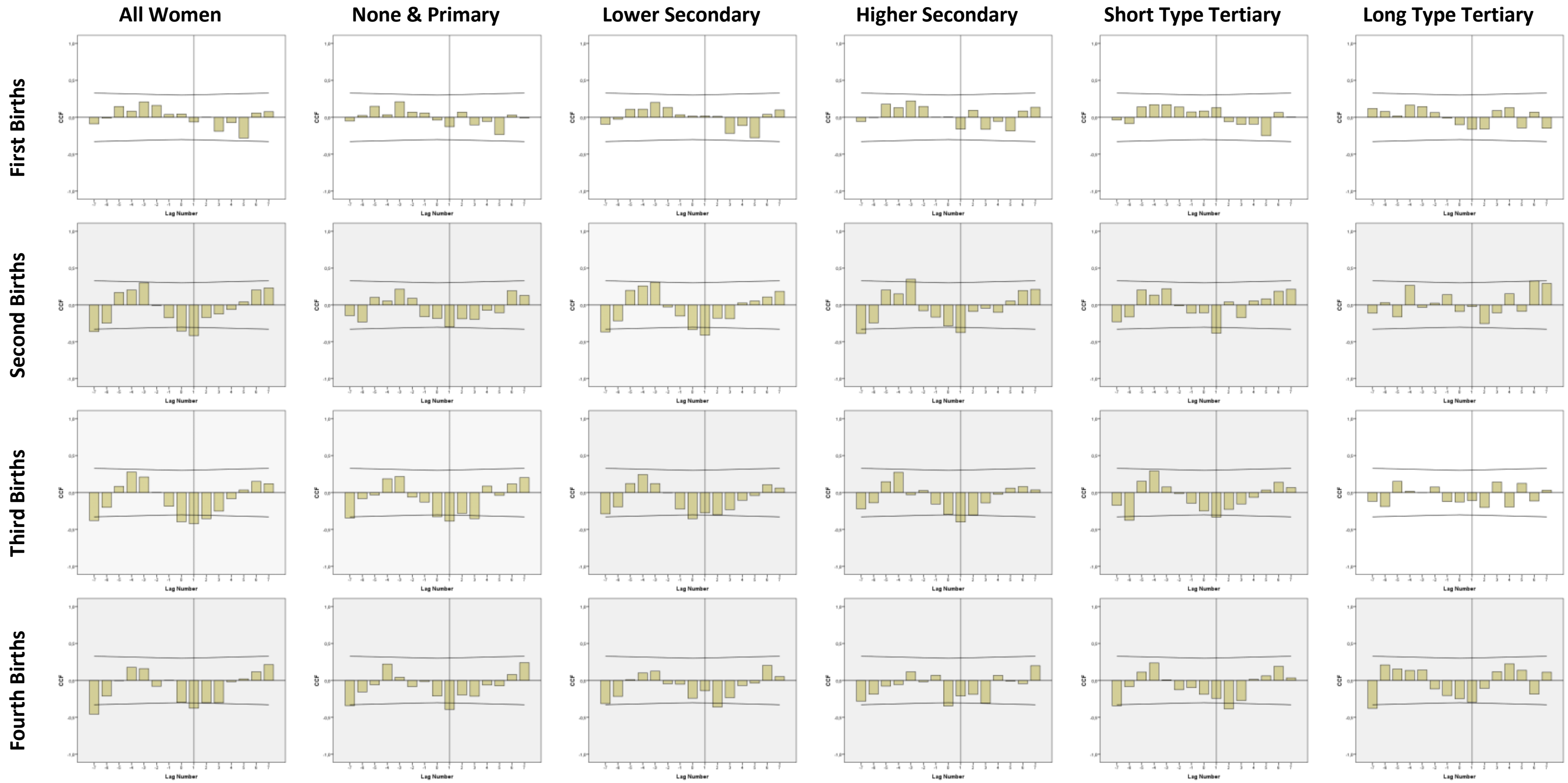
R = 0,9586
R^2 = 0,9189
Rdif = 0,4715
|e| = 0,0095

Figure 6b Observed & Simulated MAC1



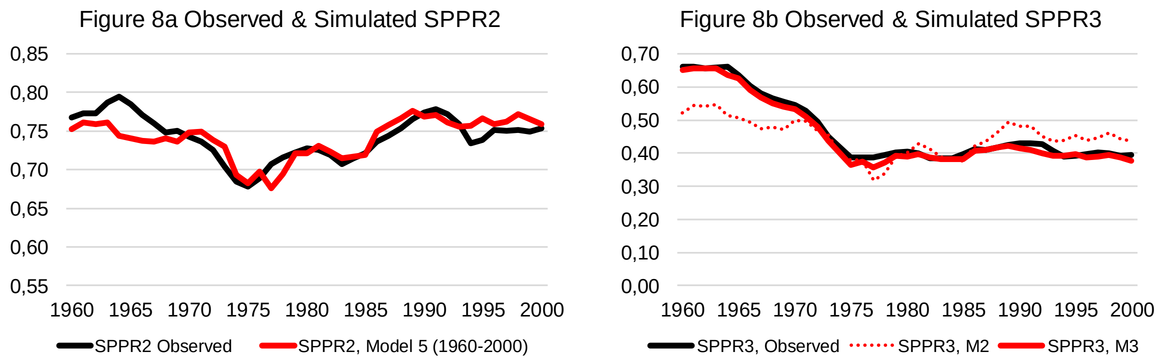
R = 0,9712
R^2 = 0,9432
Rdif = 0,4449
|e| = 0,2832

Figure 7 Cross-correlations between annual time-series of consumer price index (CPI) and observed synthetic parity progression ratios (first to fourth births) after first-order differencing and considering negative lags (births correlated to future CPI), a zero lag (births correlated to CPI in the same year) and positive lags (births related to past CPI), Belgium, 1960-2000.



Source: 2001 Belgian Census, calculations by authors. Note: Figures with significant cross-correlations have a gray background.

*Figure 8. Estimated synthetic parity progression ratio to second and third births and associated mean ages at childbearing from hazard models including duration since index birth (quadratic), mothers' age at index birth (quadratic), educational variables (enrollment, level and duration since leaving education), partnership variables and CPI lagged with one year, Belgium, 1960-2000.**



* The model for third birth births includes a dummy variable to model the strong drop of third births between 1965 and 1975 which cannot be accounted for by rising education and postponement of lower order births, variation in economic conditions (CPI) or changes in the age at which women entered a coresidential union.