Bayesian Poisson Regression for Reconstructing Fertility Rates

***Durowaa-Boateng Afua, **Yildiz Dilek**

**OeAW, Wittgenstein Centre (IIASA, OeAW, Univ. Vienna)*
***IIASA, OeAW, Wittgenstein Centre (IIASA, OeAW, Univ. Vienna)*

**afua.durowaa-boateng@oeaw.ac.at, **dilek.yildiz@oeaw.ac.at*

Introduction

The impact of education on fertility has been widely studied. Many of these studies have found an association between higher educational attainment levels and decreasing fertility rates (e.g,[1], [2], [3], [4], [5], [6]). However, access to quality historic data on fertility by level of education has been limited, especially in developing countries. The main source of detailed fertility data in many developing countries has been provided by the Demographic and Health Surveys (DHS). However due to the data collection methods employed, the quality of data is sometimes sub-optimal [7]. Certain issues including but not limited to incomplete and missing data sets may arise from collection methods [8]. To deal with missing data, Schoumaker, (2013) proposes a Poisson regression model for reconstructing birth histories from DHS waves [9]. We extend the model by Schoumaker to include education differences in fertility. The number of observations decrease when broken down by age and education. To deal with the small sample size we propose a Bayesian model to estimate the true age and education specific fertility rates and the uncertainty around them. We present preliminary results for Brazil and Turkey.

Methodology

Bayesian Poisson Regression

The Bayesian Poisson model focuses on estimating the expected number of births, μ_{jit} (for age i and time t and education level j). This is given as:

$$\mu^*_{jit} \sim Pois(\mu_{jit}) \tag{1}$$

$$\log(\mu_{jit}) = \log(e_{jit}) + \beta_0 + \sum_k \beta_{ik} \cdot A_{kit} + \sum_h \delta_h \cdot T_{hit} + \sum_j \nu_j \cdot E_{jit} + \sum_l \eta_l \cdot E_{lit} \cdot A_{lit} \tag{2}$$

$$\beta_0 \sim N(h_0, h_1) \tag{3}$$

$$\beta_{ik} \sim G(h_2, h_2) \tag{4}$$

$$\delta_h \sim N(h_3, h_4) \tag{5}$$

$$\nu_j \sim G(h_5, h_5) \tag{6}$$

$$\eta_l \sim G(h_6, h_6) \tag{7}$$

where e_{jit} represents the expected number of women exposed (of childbearing) at time t at age i, with educational level j. β_0 is the intercept of the regression model, β_{ik} is the coefficient of age and δ_h is the coefficient of time whereas ν_j is associated with the levels of education. The levels of education were specified as: No education, Primary education, Secondary education and Higher education. η_l is the coefficient associated with the interaction term of level of education and age, h_0 to h_6 are hyper parameters for the coefficients. Values of the hyper parameters were subjectively selected based on the convergence of the model.

Application

Three DHS survey waves are available in total for Brazil (1986, 1991 and 1996), however, the second wave of 1991 focused only on the North eastern part of Brazil. The data analysis including the second wave for Brazil showed that this particular region in Brazil recorded much higher fertility rates than the national average. As such our analysis did not include the second wave. In the case of Turkey, there were five surveys in the DHS database (1993, 1998, 2003, 2008 and 2013) and all five surveys were used. We present the preliminary results of the Bayesian Poisson model application to Brazilian and Turkish data sets.

Preliminary Results

Observed in Figure 1, is the clear inclination that Brazilian women with high levels of education have fewer children. Throughout the years under observation, the overall fertility slightly declines, especially for women with secondary education. A striking observation is that the age at which women have more children (peak of fertility) increases by level of education. This trend is expected as most women with higher education are well above age 19.

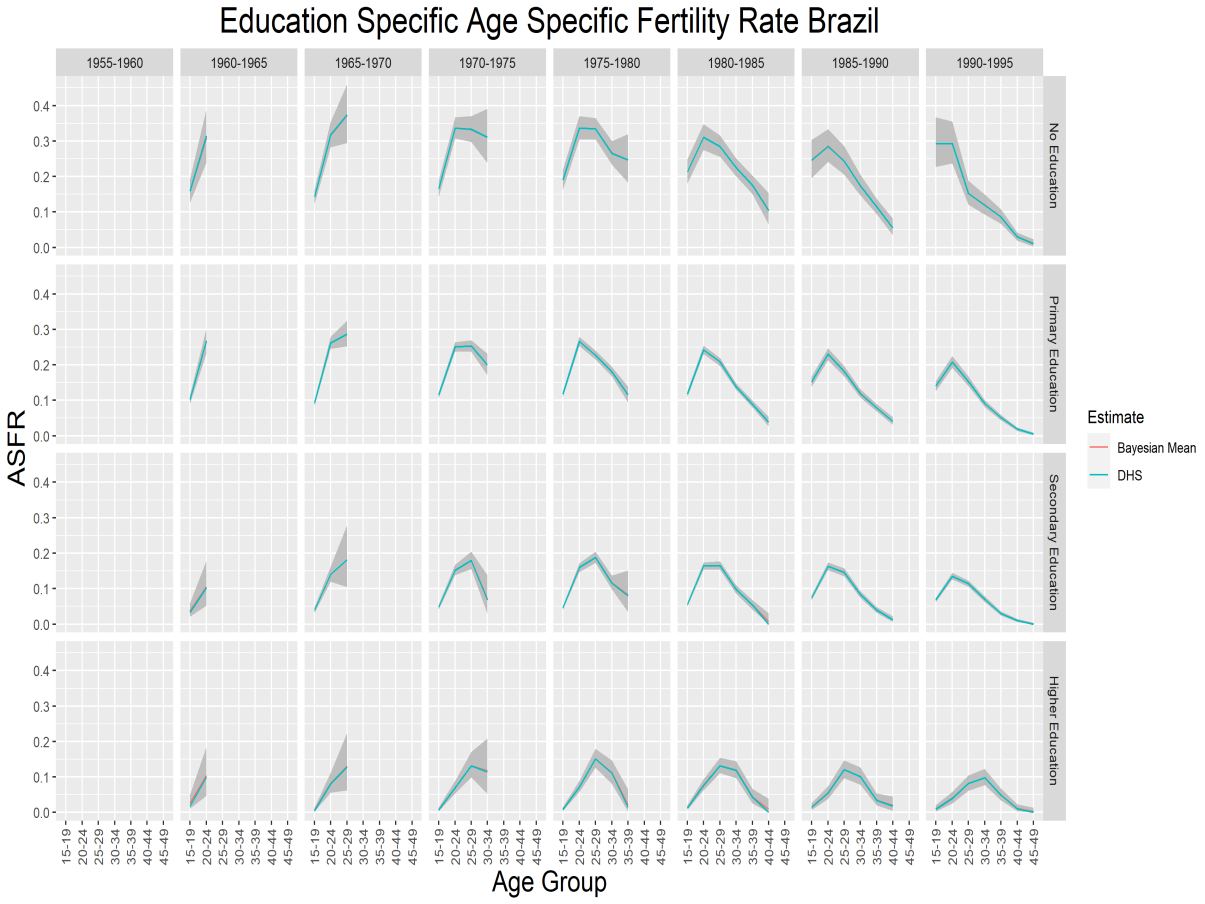

Figure 1: Bayesian Education Specific Age Specific Fertility Rate of Brazil

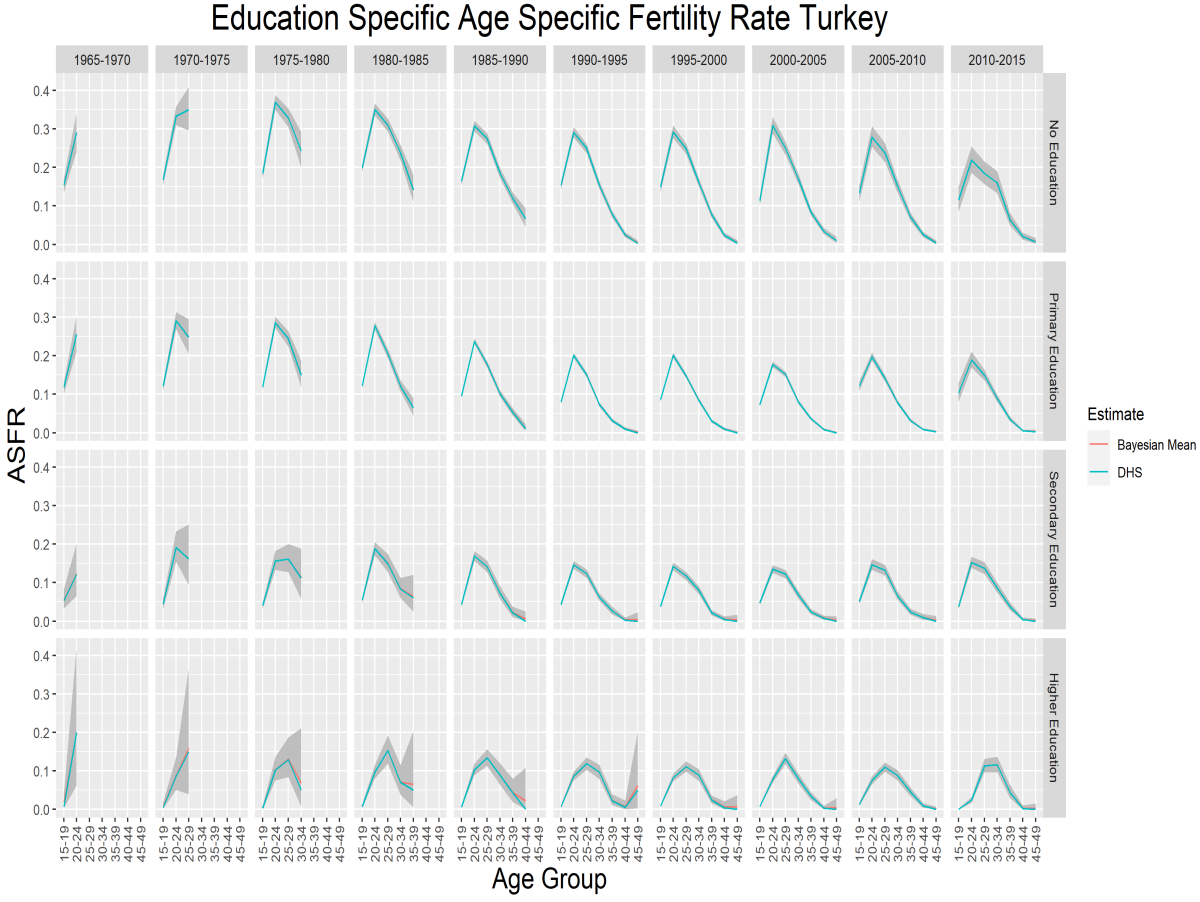

Figure 2: Bayesian Education Specific Age Specific Fertility Rate of Turkey

Similarly in Figure 2, the fertility rates decrease over the years for each educational level. Comparable to Figure 1, the peak of fertility increases in age as education progresses for Turkish women. A captivating observation is that for each educational group, the overall curve rarely changes for many years, with the exception of Higher education.

Conclusion

We focus on reconstructing fertility data using Bayesian Poisson regression by level of education. We took into account education, age, period factors as well as the interaction of age and education. The preliminary results supports existing literature that education has a strong influence on fertility. It may be explained that women with higher education are exposed to factors that may delay or reduce number of children born. These factors include but are not limited to: knowledge and access to contraceptives, focus on career and time spent in education [11]. At the moment only one data source, the DHS data base, was used to test the performance of the model. Future work on the model would focus on combining other data sources from countries with similar fertility rates. Thus enabling the model to further learn from the other countries and to complete age schedules.

Bibliography

[1] Ester Lazzari, Ryohei Mogi, and Vladimir Canudas-Romo. Educational composition and parity contribution to completed cohort fertility change in low-fertility settings. *Population Studies*, pages 1–15, 2021.

[2] Wolfgang Lutz and KC Samir. Global human capital: Integrating education and population. *Science*, 333(6042):587–592, 2011.

[3] Alaka Malwade Basu. Why does education lead to lower fertility? a critical review of some of the possibilities. *World Development*, 30(10):1779–1790, 2002.

[4] Endale Kebede, Anne Goujon, and Wolfgang Lutz. Stalls in africa's fertility decline partly result from disruptions in female education. *Proceedings of the National Academy of Sciences*, 116(8):2891–2896, 2019.

[5] John Bongaarts. The causes of educational differences in fertility in sub-saharan africa. *Vienna yearbook of population research*, pages 31–50, 2010.

[6] Øystein Kravdal. Education and fertility in sub-saharan africa: Individual and community effects. *Demography*, 39(2):233–250, 2002.

[7] Bruno Schoumaker. *Quality and consistency of DHS fertility estimates, 1990 to 2012*. ICF International Rockville, Maryland, 2014.

[8] Bruno Schoumaker. Reconstructing fertility trends in sub-saharan africa by combining multiple surveys affected by data quality problems. In *Proceedings of the 2010 Annual Meeting of the Population Association of America*, 2010.

[9] Bruno Schoumaker. Reconstructing long term fertility trends with pooled birth histories. In *XXVII International Population Conference*, 2013.

[10] Jakub Bijak and John Bryant. Bayesian demography 250 years after bayes. *Population studies*, 70(1):1–19, 2016.

[11] Margaret Frye and Lauren Bachan. The demography of words: The global decline in non-numeric fertility preferences, 1993–2011. *Population studies*, 71(2):187–209, 2017.