Projecting the presence

Estimating incomplete parity- and age-specific fertility rates

Carlo G. Camarda * Harald Wilkoszewski †

Extended abstract submitted
for the PAA Annual Meeting 2008

1 Introduction

Complex demographic analyses often require long-time series of detailed vital rates. In many cases, the data available do not meet these requirements: whereas crude indicators are given for long-time series, detailed rates are only available for short periods of time. In the area of fertility for example Total Fertility Rates (TFR) are easily obtained, while parity- and age-specific fertility rates (PASFR) are often incomplete. This paper provides a new approach to deal with this kind of data structure. We suggest a modified application of the Lee Carter model (LC), which usually is being used in the area of mortality (Lee and Carter, 1992); given the fact that fertility patterns are generally speaking more erratic than mortality trends, the linear assumption in the forecast of the LC does not lead to reasonable results. Therefore, we use the information given by the available TFR to project PASFR.

In our paper we concentrate on Germany; this for two reasons. First, fertility data available for Germany are a typical example of the data structure described above. PASFR for West Germany only exist for the period 1957 to 1985, for East Germany only for 1956 to 1989 (Figure 1). Second, fertility patterns in East and West have differed significantly both before and after reunification. Especially East Germany showed wavering trends. Our aim is to complete the incomplete PASFR for the projected period until 2004 for both parts of Germany taking into account the given TFR. The data we use come from the Fertility Database at the Max Planck Institute for Demographic Research.

*Max Planck Institute for Demographic Research, Statistics Laboratory; email:camarda@demogr.mpg.de

†Max Planck Institute for Demographic Research, Population and Policy Laboratory; email: wilkoszewski@demogr.mpg.de

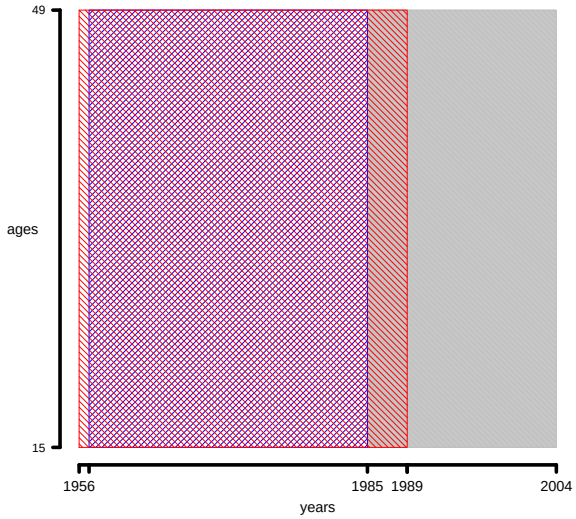

Figure 1: Availability of parity- and age-specific fertility rates in Germany. West Germany is depicted with blue and East Germany in red. The grey area denotes the period we aim to project.

2 Methodological approach

The LC was first applied in the area of fertility by Lee (1993). Further applications can be found in Lee and Tuljapurkar (1994), Lee and Skinner (1999), Li and Wu (2003), Lee (2004) and Goldstein (2004).

In our case the LC decomposes the ASFR of each parity (i) as follows

$$f_{xy}^{(i)} = a_x^{(i)} + b_x^{(i)} k_y^{(i)} \tag{2.1}$$

where $a_x^{(i)}$ gives the average of the fertility rates over the whole period, $b_x^{(i)}$ represents the patterns of deviation from this average over age, and $k_y^{(i)}$ denotes the variance over time at any given age. We use the Singular Value Decomposition to estimate the LC.

The classic approach to forecast vital rates on the basis of the LC is to extrapolate the $k_y^{(i)}$ in (2.1) by applying univariate ARIMA models (Box and Jenkins, 1970). In the case of mortality, the outcome of this method is a random walk with drift, since the estimated $k_y^{(i)}$ generally form a linear downward trend. In the case of fertility, the pattern of the $k_y^{(i)}$-values is not that straight forward. For each parity in West Germany we can observe some increase for the first years of the given period, followed by a rather long decreasing trend and an indication of a levelling off during the last years (see Figure 2). For East Germany the trends are even more erratic (see Figure 3). Without additional assumptions being made, it is obvious that forecasts of these patterns would not lead to reasonable results. Since in our case, the forecast period was actually not the future but rather the past, for

which some more information on the overall fertility level was available, we modified the approach.

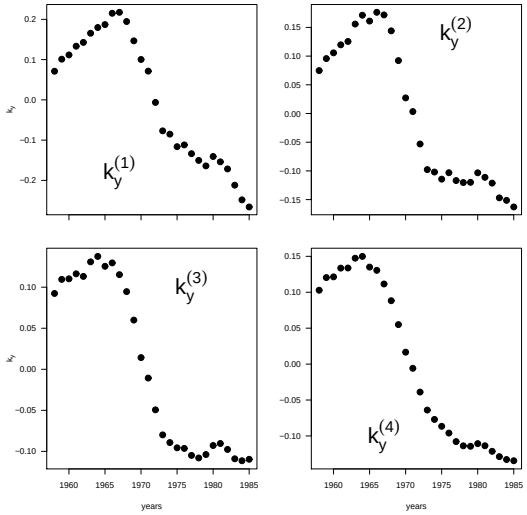

Figure 2: Patterns of $k_y^{(i)}$ $i = 1, 2, 3, 4$, Germany West, 1957-1985.

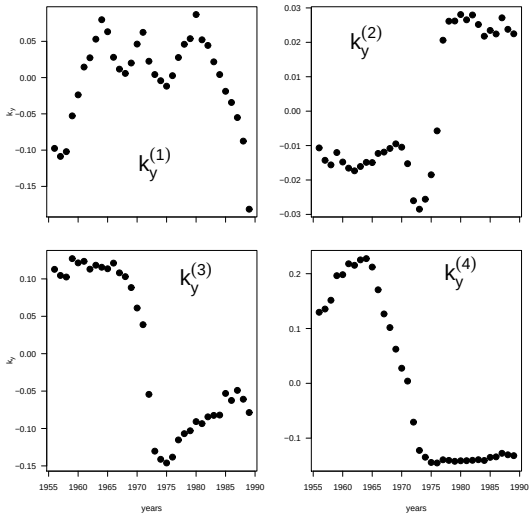

Figure 3: Patterns of $k_y^{(i)}$ $i = 1, 2, 3, 4$, Germany East, 1956-1989.

In order to account for significant changes in the German fertility regimes, we used the TFR (see Figure 4) to adjust the projected PASFR for the period 1986/1990 to 2004. We applied an iterative process, where we fixed the k_y-values for the fourth parity (the one with the least variance to be expected) to the last value given, set the mean of the second and third parity for the projected period to a certain level, and estimated the distribution of the first parity by taking the TFR into account; we repeated this process with the newly given distributions until convergence was reached. From the estimated $k_y^{(i)}$ we obtained the final PASFR.

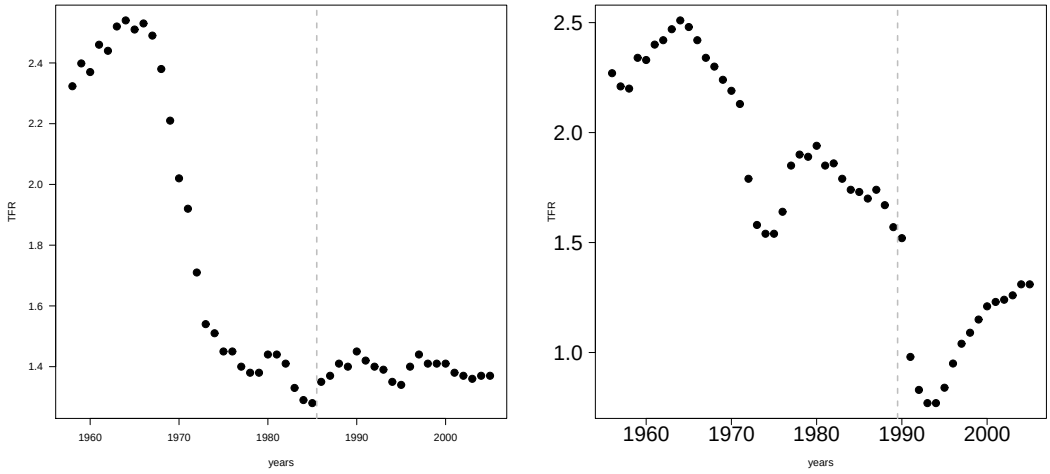

Figure 4: TFR Germany West and East.

3 Preliminary results

The results of our estimation are displayed in Figure 5 and 6. In the West German case, our approach produces sound outcomes, as parities 1 to 3 capture the trend of the TFR over the projected period. In the East German case, the values for parity 1 go well below the level of parity 2. We assume that this is due to the specifics of the LC; since the pattern of deviations over age from the average of the fertility rates for the whole period ($b_x^{(i)}$) is fixed, the model cannot capture future changes of these deviations such as postponement of births. However, we argue that our estimations for East Germany very likely show a realistic picture of the dramatic fertility decline's composition after the fall of the Berlin Wall: the consequences of this strong period effect might have been much bigger on those people who were in the decision making process of whether or not to have a first child than on those who already had founded a family. In the paper we will further investigate into this question in order to test the plausibility of our results.

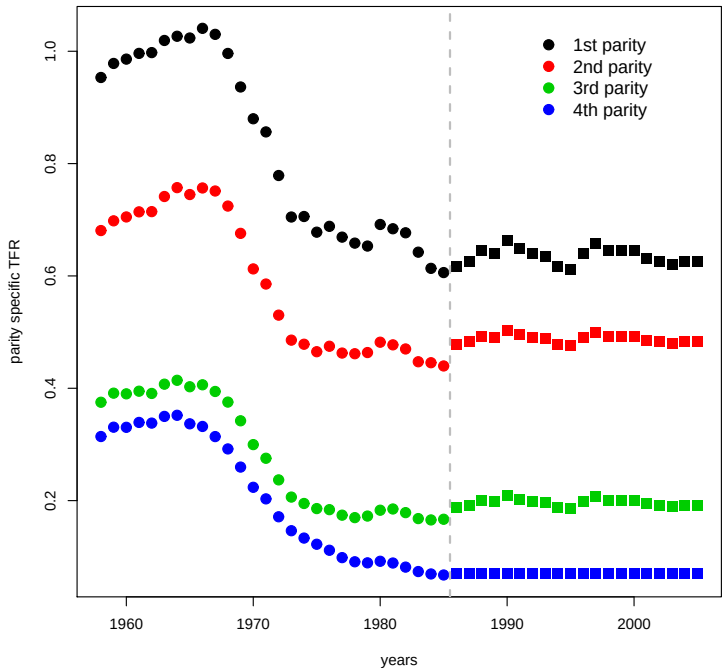

Figure 5: Estimated parity-specific TFR, Germany West.

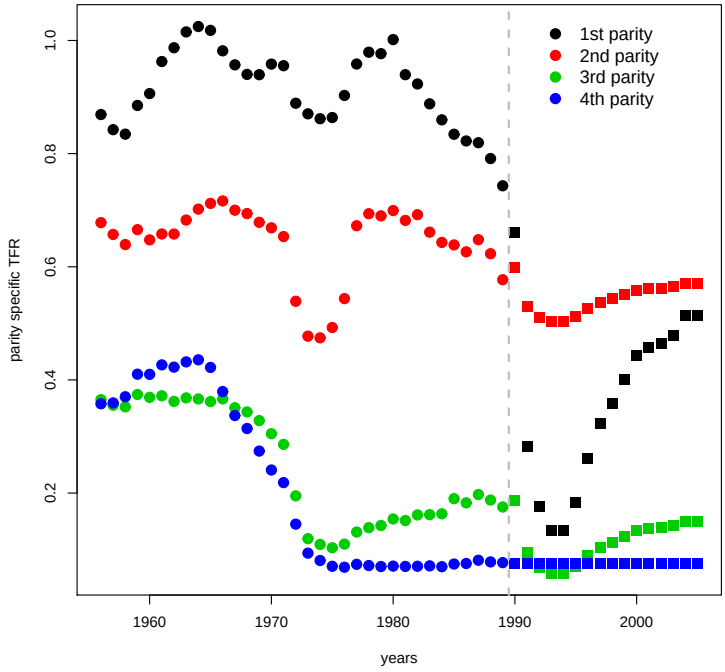

Figure 6: Estimated parity-specific TFR, Germany East.

References

Box, G. and G. Jenkins (1970). *Time series analysis: Forecasting and control*. Holden–Day.

Goldstein, J. R. (2004). Simpler Probabilistic Population Forecasts: Making Scenarios Work. *International Statistical Review 72*, 93–106.

Lee, R. D. (1993). Modeling and forecasting the time series of US fertility: Age distribution, range, and ultimate level. *International Journal of Forecasting 9*, 187–202.

Lee, R. D. (2004). Quantifying Our Ignorance: Stochastic Forecasts of Population and Public Budgets. *Population and Development Review 30*, 153–175.

Lee, R. D. and L. R. Carter (1992). Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association 87*(419), 659–671.

Lee, R. D. and J. Skinner (1999). Will Aging Baby Boomers Bust the Federal Budget? *Journal of Economic Perspective 13*, 117–140.

Lee, R. D. and S. Tuljapurkar (1994). Stochastic Population Forecasts for the United–States - beyond High, Medium and Low. *Journal of the Statistical American Association 89*, 1175–1189.

Li, N. and Z. Wu (2003). Forecasting cohort incomplete fertility: A method and an application. *Population Studies 57*, 303–320.