# An APC Approach in Transforming Cohort Fertility Levels into Schedules: Evidence from European Countries

P. C. Roger Cheng, Ph.D. (Corresponding Author)

Email address: paochih@mgt.ncu.edu.tw

Department of Economics, National Central University, Taiwan.

No.300, Jhongda Rd., Jhongli City, Taoyuan County 32001, Taiwan.

Tel: 886-3-422-7151 ext. 66310; Fax: 886-3-422-2876;

Eric S. Lin, Ph.D.

Email address: slin@mx.nthu.edu.tw

Department of Economics, National Tsing Hua University, Taiwan.

# An APC Approach in Transforming Cohort Fertility Levels into Schedules: Evidence from European Countries

P. C. Roger Cheng and Eric S. Lin

Preliminary Draft

## ABSTRACT

In our previous work (Cheng and Lin 2010) we developed a simple age-period-cohort framework in forecasting cohort fertility schedules, and made full use of 1917–2005 U.S. data to obtain robust outcomes. Our new approach 1) is easy to implement, 2) can estimate the entire cohort fertility schedule with relatively mild data requirement, and 3) has been shown to outperform other popular methods in the literature in forecasting sense. In sum, the procedure is divided into two steps: the first is to estimate cohort fertility levels and the second is to transform a level into a schedule. In this paper, we extend our previous work by incorporating fertility data from European countries, keep the focus on the second step, mark any period of 25 consecutive years as a sample, and carry out validation experiments to all cohorts covered in each sample period. Performances of our approach and some representative ones proposed in the literature will be evaluated and compared, according to the root mean square error (RMSE) criterion.

# 1 Introduction

Forecasting uncompleted cohort fertility schedule is a quite important and challenging task for demographers. It is important because a cohort fertility schedule provides much useful information (such as the proportion of women who will have a child, the level we human beings reproduce ourselves, the distribution of the age gap between mothers and their children, and so forth) which forms a sound basis for policy makers to design their population, social, and health policies. On the other hand, it is challenging because a lot approaches have been proposed in the literature (refer to Cheng 2010, Cheng and Lin 2010, and Myrskylä and Goldstein 2010) while few perform very well (refer to Cheng and Lin 2010, Table 1).

According to the World Factbook 2010 published by Central Intelligence Agency (CIA),[1] global fertility rates in general have a declining tendency. This trend is most pronounced in developed countries, such as some European ones, and the estimated period total fertility rates (PTFR) in 2010 are only 1.29, 1.32, 1.36, 1.37 and 1.39 for Poland, Spain, Slovakia, Greece and Austria, respectively. It is, however, quite challenging for policy makers to reconcile these low period fertility rates with the aforementioned population, social and health policies, given the well recognized fact that PTFRs are distorted by the tempo and/or spread effects. Thus, it would be much more informative to derive the complete fertility schedules for European countries in a cohort perspective.

Recently, we develop a new approach in forecasting cohort fertility schedules, which is easy to implement and is able to estimate the entire cohort fertility schedule

---

[1] `https://www.cia.gov/library/publications/the-world-factbook/rankorder/2127rank.html`

with relatively mild data requirement. In particular, we make full use of 1917–2005 U.S. data to obtain robust outcomes showing that the new approach outperforms other competing methods. Our approach is distinguished from other methods in the fundamental strategy. Briefly, previous studies assume a particular structured mechanism behind the data, estimate related parameters of the structure, and then produce age-specific fertility forecasts constituting the uncompleted schedule. The cohort total fertility rate (CTFR) is just one estimate among others and they are generated from the projected schedule at the same time. On the contrary, we allow the fertility level and the fertility schedule to be derived separately but sequentially. More specifically, our method consists of two steps: the first is to estimate cohort fertility levels (i.e., CTFRs), and the second is to transform a level into a schedule with a simple age-period-cohort (APC) framework.

In this paper, we extend our approach to incorporate fertility data from 17 European countries, keep the focus on the second step, and carry out validation experiments. Performances of our approach along with the linear extrapolation and the frozen rate methods will be evaluated and compared based on the root mean square error (RMSE) criterion. The remainder of this paper is organized as follows: Section 2 briefly describes the fertility data from European countries. In Section 3, we provide an overview of the frozen rate and the linear extrapolation methods along with ours, and then introduce how we evaluate their performances in forecasting and how the comparison is carried out. Experimental results will be presented and discussed in Section 4, and Section 5 investigates the extent to which the curve approximation will be affected when the input level is biased. Section 6 summarizes and concludes.

## 2 Data

The data employed are age-specific fertility rates (ASFR; all births combined) by year and age which are derived mainly from the Human Fertility Database[2] and the Eurostat Database.[3] In addition, we thank Dr. Michaela Kreyenfeld for providing us with the W. Germany data. Note that not all countries listed in the Eurostat Database are included in this study because their data ranges are not lengthy enough to construct complete fertility schedules for more than 10 cohorts. Table 1 presents the date ranges and related information for 18 countries; besides 17 European ones the U.S. is also included for purposes of comparison.

## 3 Comparison among approaches

To evaluate the performance of the new APC framework developed in Cheng and Lin (2010), it is appropriate to adopt some rivals to compare with. In this paper, we adopt two approaches along with ours to forecast the uncompleted parts of cohort fertility schedules with various truncation ages and then to assess the extents to which they approximate to the observed ones.

### 3.1 Frozen Rate method

The first and most simplified method is to freeze ASFRs at ages of the uncompleted part in the last observed year, as if the incomplete cohort schedule will follow this particular pattern. We refer to this method (denoted as the Naive approach in our previous work) as the Frozen Rate approach hereafter. According to the experimen-

---

[2] Human Fertility Database. Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria). Available at `http://www.humanfertility.org` (data downloaded on Feb. 11, 2010).

[3] Available at `http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search\_database` (data downloaded on May 3, 2010).

Table 1. Data ranges and related information for 18 countries

| Country | Data range<br>Year | <br>Age | Birth cohorts in data range | Complete cohorts in data range |
|---|---|---|---|---|
| Austria | 1951–2008 | 14–49 | 1902–1994 | 1937–1959 |
| Bulgaria | 1960–2008 | 15–49 | 1911–1993 | 1945–1959 |
| Czech | 1950–2008 | 14–49 | 1901–1994 | 1936–1959 |
| Denmark | 1960–2008 | 15–49 | 1911–1993 | 1945–1959 |
| Finland | 1960–2008 | 15–49 | 1911–1993 | 1945–1959 |
| W. Germany | 1952–2008 | 15–49 | 1903–1993 | 1937–1959 |
| Greece | 1961–2008 | 15–49 | 1912–1993 | 1946–1959 |
| Hungary | 1960–2008 | 15–49 | 1911–1993 | 1945–1959 |
| Iceland | 1963–2008 | 15–49 | 1914–1993 | 1948–1959 |
| Italy | 1960–2007 | 15–49 | 1911–1992 | 1945–1958 |
| Netherlands | 1950–2008 | 14–49 | 1901–1994 | 1936–1959 |
| Norway | 1961–2008 | 15–49 | 1912–1993 | 1946–1959 |
| Portugal | 1960–2008 | 15–49 | 1911–1993 | 1945–1959 |
| Russia | 1959–2008 | 14–49 | 1910–1994 | 1945–1959 |
| Slovakia | 1950–2008 | 14–49 | 1901–1994 | 1936–1959 |
| Sweden | 1891–2007 | 14–49 | 1842–1993 | 1877–1958 |
| Switzerland | 1944–2007 | 14–49 | 1895–1993 | 1930–1958 |
| U.S. | 1917–2006 | 14–49 | 1868–1992 | 1903–1957 |

Note: Data come from the Human Fertility Database include Austria, Czech, Netherlands, Russia, Slovakia, Sweden, Switzerland, and the U.S. The West Germany data are provided by Dr. Michaela Kreyenfeld. Others are from the Eurostat Database.

tal results on the U.S. fertility data of first birth by Cheng and Lin (2010, Table 1), this method outperforms many prior approaches such as curve-fitting models, the Lee-Carter method (Lee and Carter 1992), and the Willekens-Baydar method (Willekens and Baydar 1984). Although its performance is not as good as the Evans method (Evans 1986), the latter requires data periods longer than 30 years and is thus excluded from the comparison in this paper.

### 3.2 Linear Extrapolation method

The other method takes the observed ASFRs in the last five years to fit linear trends age by age, and then extrapolates the ASFRs in the following years. The predicted ASFRs however could be negative, in which case we replace negative values with zero.

### 3.3 APC-TRUE approach

The framework developed in Cheng and Lin (2010) can be summarized as follows:

1. Apply a simple age-period-cohort (APC) model to estimate age-, cohort-, and period-specific effects from the sample data.
2. Estimate CTFRs for incomplete cohorts from the sample data.
3. Fit out-of-sample period effects in a model that uses estimated CTFRs as control totals.
4. Fill in missing data by adding the in-sample age and cohort effects to the out-of-sample period effects.

We have shown, in our previous study, that the model for out-of-sample period effects in step 3 is the key to transforming a fertility level into a fertility schedule with the U.S. first birth data. In this paper, the locus function of out-of-sample period effects

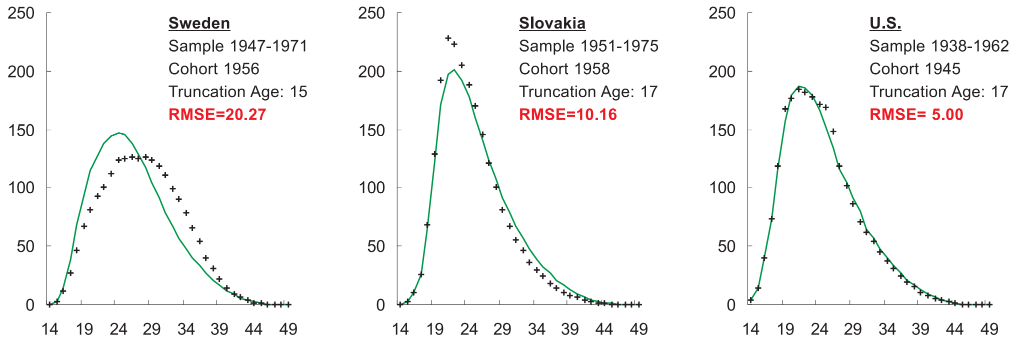

Figure 1. Examples for various extents of approximation to observed schedules

Note: Points (+) are observed (true) schedules and lines are forecast ones.

(as defined in our previous study) is restricted to the linear one for simplicity. In addition, since the focus is on investigating how well the transformation process performs when it is applied to European data, our experiments will be based on a premise that the actual CTFRs are known. We therefore denote the approach APC-TRUE hereafter.

### 3.4 Performance measurement and comparison design

To measure the extent to which an array of forecasted ASFRs ($\hat{f}_i,\ i = 1, 2, \ldots, n$) approximates to that of the pairwise matched observations ($f_i,\ i = 1, 2, ..., n$), we adopt the root mean square error (RMSE)

$$\text{RMSE} = \left[ n^{-1} \sum_{i=1}^{n} e_i^2 \right]^{1/2}$$

as the criterion, where $e_i = \hat{f}_i - f_i$ (in live births per thousand women). Figure 1 presents three examples with RMSE being about 20, 10, and 5 to give the reader an idea as to how well these curves fit the observed ones. In this paper, we adopt 5 as a standard of ‘excellent’ and 10 of ‘fair’ approximation.

Instead of depleting all of the observations once only, we follow the process

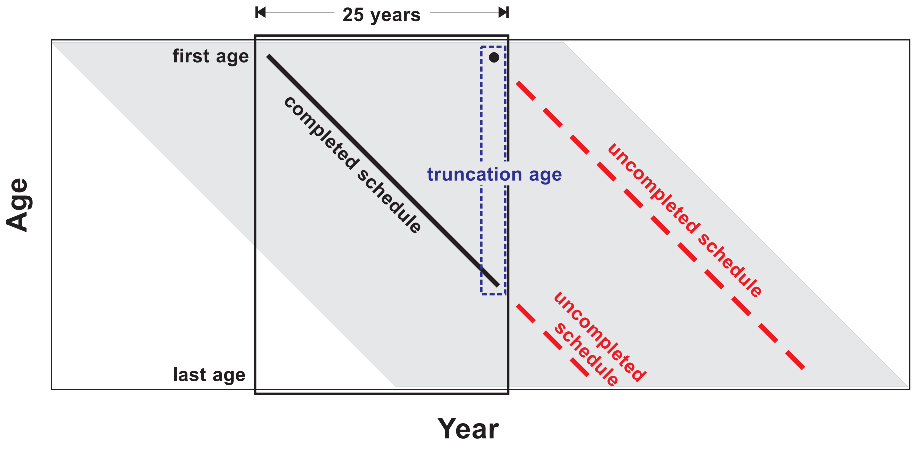

Figure 2. An illustration of how the data is utilized

Note: Given a 25-year sample, there are 25 cohorts who lack their future fertility schedules only and whose ages in the last year of the sample are denoted by *truncation ages*. The shaded area indicates the full information regarding cohort fertility schedules that can be derived from the entire data range.

adopted in our previous study to yield more robust results. Specifically, we mark any period of 25 consecutive years during the data range as a sample, and there will be several samples for each country, depending on the length of the data range. Within each 25-year sample, all cohort fertility schedules are incomplete; some lack their past records, some lack their future experiences, and some lack both. The paper focuses on comparing the performances of three approaches in forecasting the uncompleted schedules for cohorts who have not finished their future fertility.

Figure 2 illustrates how we utilize the entire data set (represented by the largest year-age rectangle). As can be seen, within the 25-year rectangle there are 25 cohorts whose early fertility experiences (represented by diagonal solid lines) are available, while their future schedules (represented by diagonal dashed lines) are not. We denote the age in the last year of the sample for a particular cohort as *truncation age*; a younger truncation age represents a larger percentage of uncompleted fertility. The shaded area indicates the full information regarding cohort fertility schedules that can be derived from the entire data range. In this illustration, one can examine how well the forecast schedule by a particular approach approximates to the actual one for all 25 cohorts based on the RMSE criterion. As the 25-year sample moves along the year dimension, one can collect all RMSE values and obtain a corresponding distribution for each truncation age. Note that the number of RMSE values available differs across truncation ages, and it could even be zero for some truncation ages, depending on how long of the data range is.

## 4 Experimental results

Obtaining all experimental results for each method, we next investigate not only the extent to which our APC-TRUE approach outperforms the other two but also the limit it can reach under a reasonable RMSE standard.

Figure 3 depicts the distributions of forecast errors (measured by RMSE) by truncation age for 18 countries, with stock charts marking the maximum, the third quartile, the median, the first quartile, and the minimum. As can be expected, the younger the truncation age, the larger the forecast errors. We also indicate the 50% truncation (in a little excess of the peak of a schedule) in each panel for the reader's reference.[4] For some countries, the data range is not lengthy enough so that all cohorts investigated have finished more than half of total fertility, we mark the average truncation percentage at the minimum truncation age. A visual inspection of Figure 3 concludes that:

1. With few exceptions, the APC-TRUE approach outperforms the other two in forecasting incomplete cohort fertility schedules. Even in the case of Greece

---

[4] Note that each cohort reaches the 50% truncation at a particular age which may differ across cohorts. We therefore use the mean truncation age as a representative.

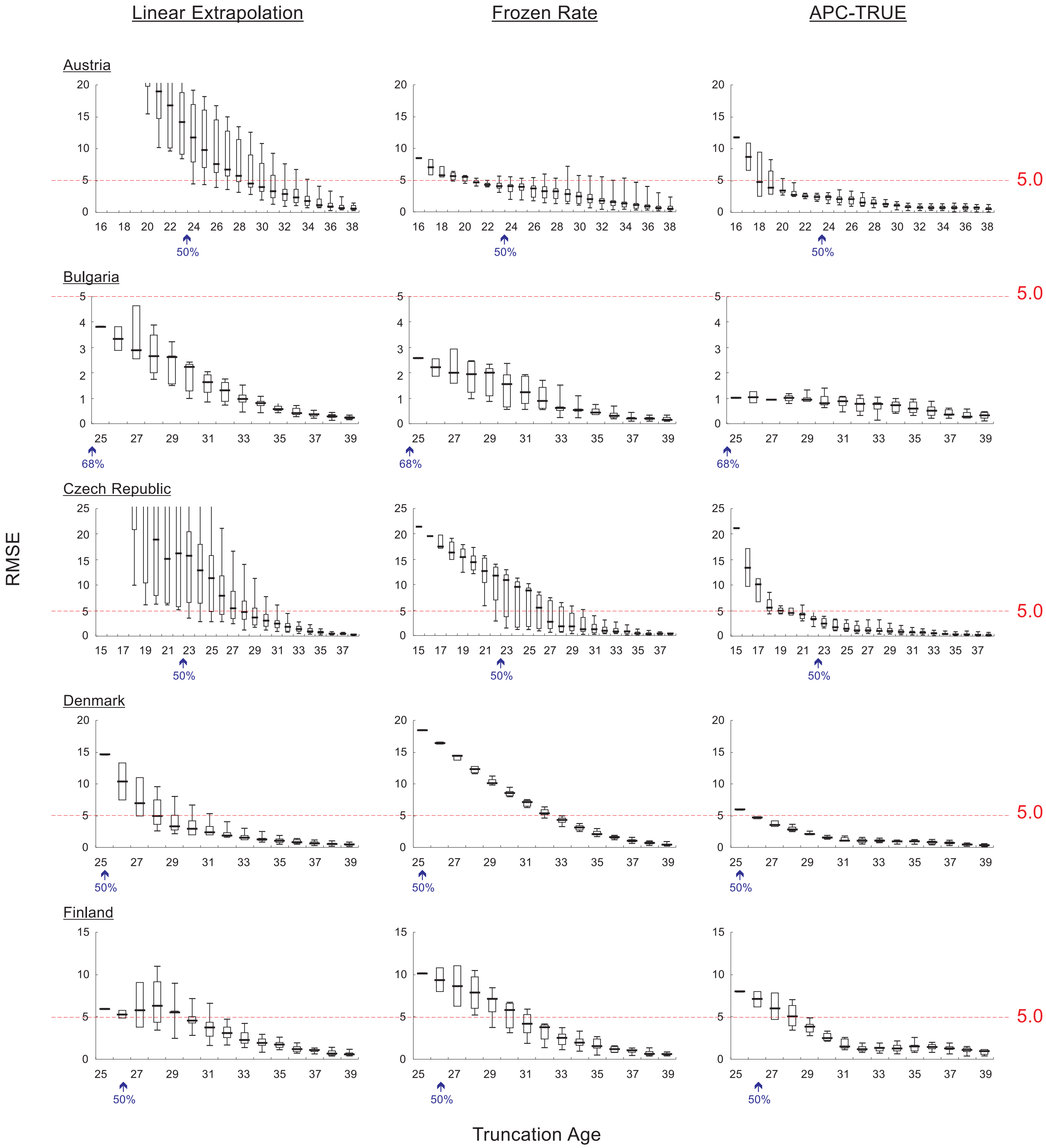

Figure 3. Forecast errors by truncation age and country: all births combined

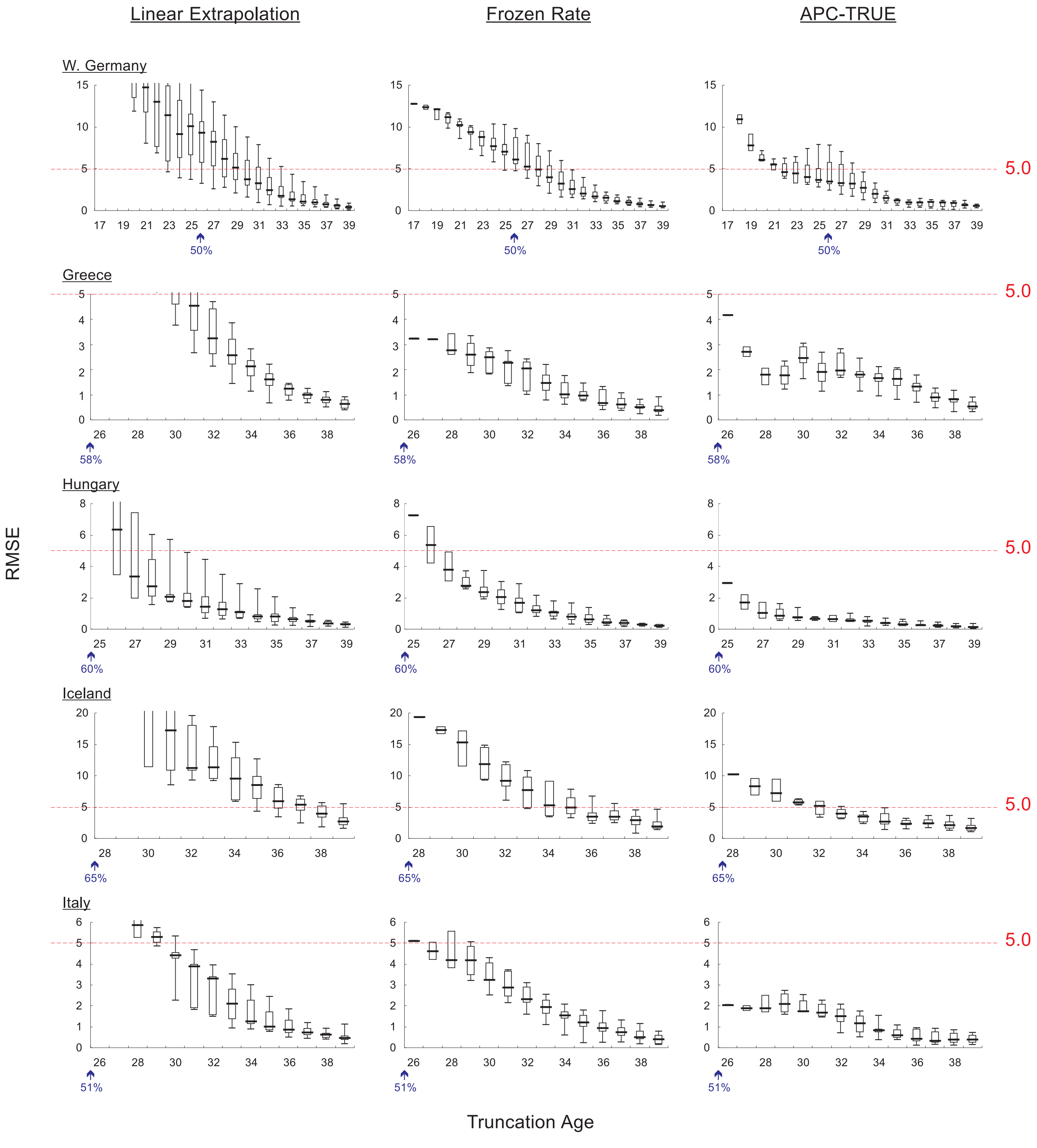

Figure 3: (Continued)

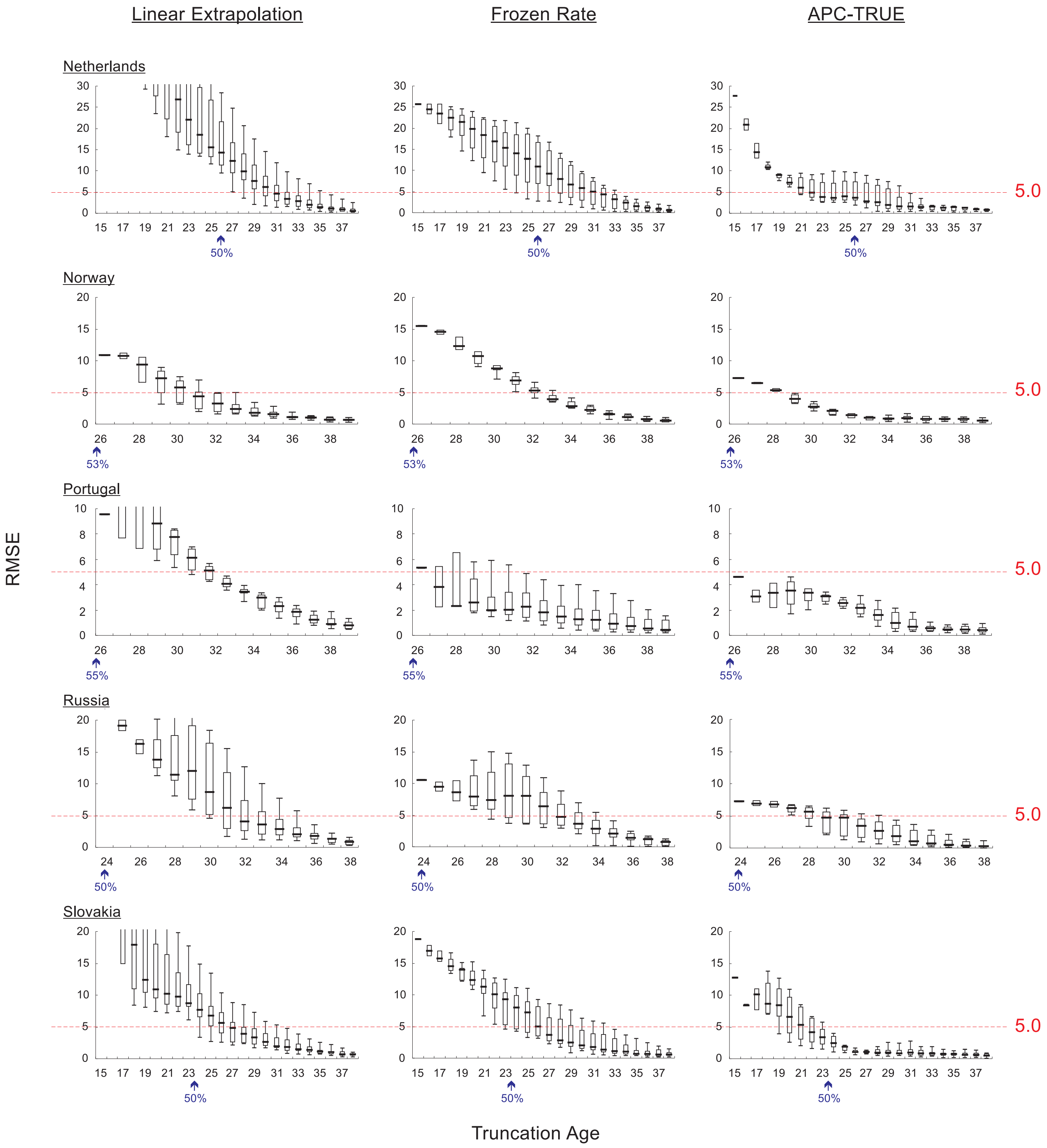

Figure 3: (Continued)

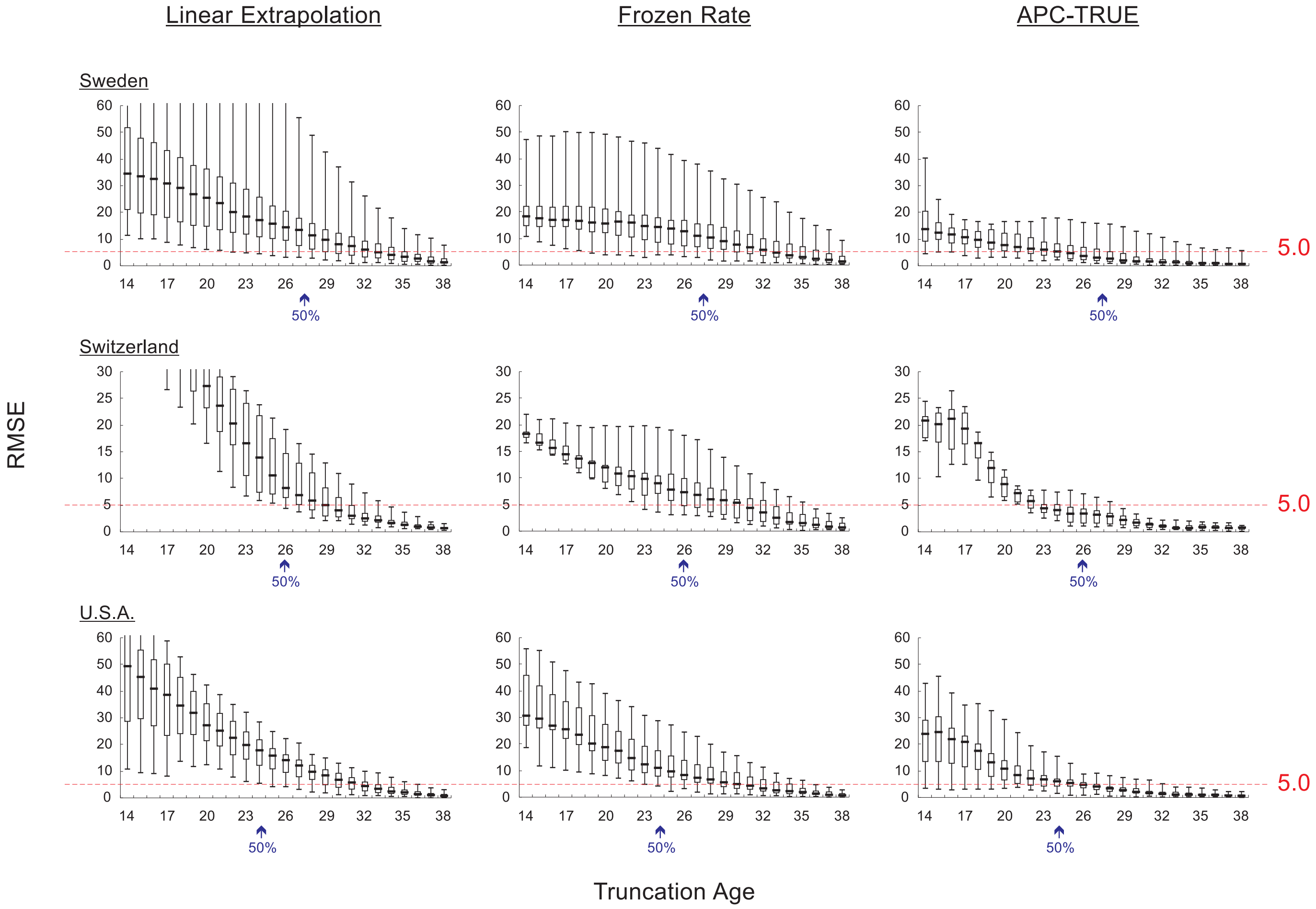

Figure 3: (Continued)

that the performances are close for the APC-TRUE and the Frozen Rate, they are both 'excellently' below the 5% RMSE standard.

2. For countries such as Austria, Czech, Italy, and Slovakia, the whole forecast error distribution (represented by the maximum) of the APC-TRUE falls below 5, the 'excellent' RMSE criterion, at the 50% (and above) truncation. For most of the other countries, at least half of forecast errors (represented by the median) of the APC-TRUE fall below 5 and all the errors fall below 10, the 'fair' RMSE criterion, at the 50% (above) truncation.

3. For countries with a very long data span such as Sweden (1891–2007) and the U.S. (1917–2006), an significant gap between the maximum and the third quartile of forecast errors which does not appear for other countries can be found, indicating that there are some outlying values that may be obtained under certain circumstances. We next investigate this finding in more detail.

Examining cohort fertility schedules in the U.S. one by one enables us to find two significant spikes persisting across cohorts 1915–1923: the first corresponds to year 1942, just one year after the Japanese's attack on Pearl Harbor, dragging America into the world conflict, and the second one is in years 1946–1947, the first and second years after the end of World War II, upon the return of the servicemen from overseas. Similarly, there is a salient spike corresponding to year 1920 (two years after the end of WWI) which persists across cohorts 1880–1900 and another to years 1941–1943 which persists across cohorts 1904–1914 in Sweden.[5] Such spikes in fertility

---

[5] During WWII, Sweden remained neutral but the British attempted to interdict the German-Swedish iron ore trade and sent a fleet into the Baltic Sea to stop shipping reaching Germany until 1940.

schedules can make a forecast curve deviate from the actual one and thus lead to a high RMSE value when the macro shocks occur at the moment right after the end of the sample period. After removing samples for which the aforementioned macro shocks caused by wars cannot be expected during the sample period, the gap between the maximum and the third quartile of forecast errors vanishes.

## 5 Sensitivity examination

So far we have presented the performance of our APC-TRUE approach when applied to ASFR data (all brths combined) from 18 countries. To the results above, the reader might object as follows: "However excellent, the APC-TRUE is after all a *hypothetical* measure since no actual CTFR for an incomplete schedule is available due to the data-range limitations". That is correct, and we have proposed a possible substitute for the actual CTFR in our previous study. In this section, a sensitivity check is carried out to see the performance of our APC framework in curve approximation when there is a bias of the CTFR levels.

Suppose that the CTFR information is biased, upward or downward, to a certain extent, then how will the distribution of forecast errors respond? We investigate situations when the uncompleted fertility levels are ±5% and ±10% biased. For example, when a cohort whose fertility schedule is 50% (60%) truncated, a 5% bias corresponds to a 2.5% (2%) absolute percentage error. There are two types of criteria for the examination: absolute and relative. One can evaluate the performance under the 5 (excellent) or 10 (fair) RMSE criterion, or compared to that of a competing approach, say the Frozen Rate method.

For presentation reasons, Figure 4 illustrates the third quartile of an error dis-

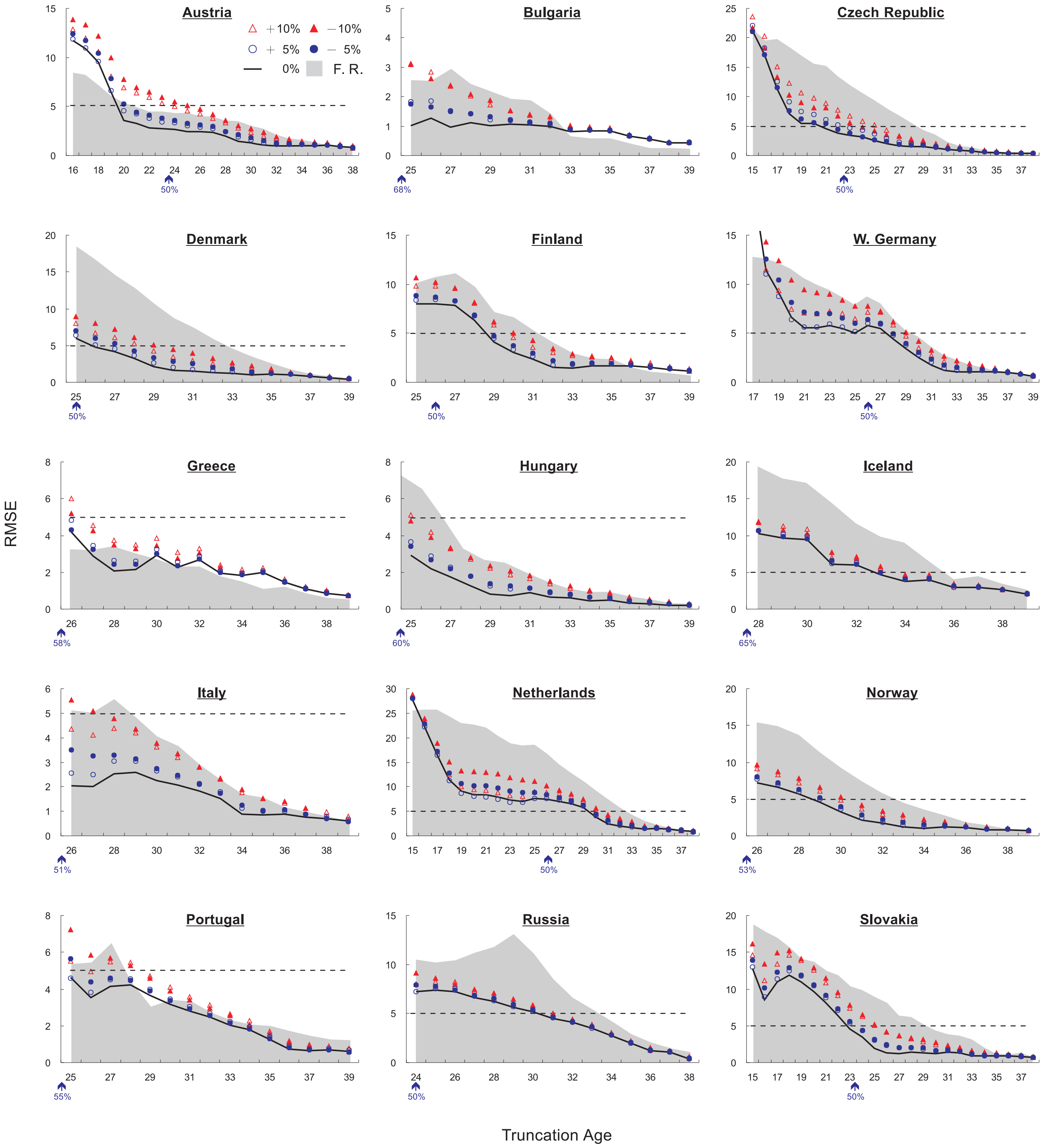

Figure 4. Sensitivity check on third quartile of forecast errors with biased CTFRs

Note: In the legend provided in the upper-left panel, the number denotes the bias percentage of an uncompleted fertility level.

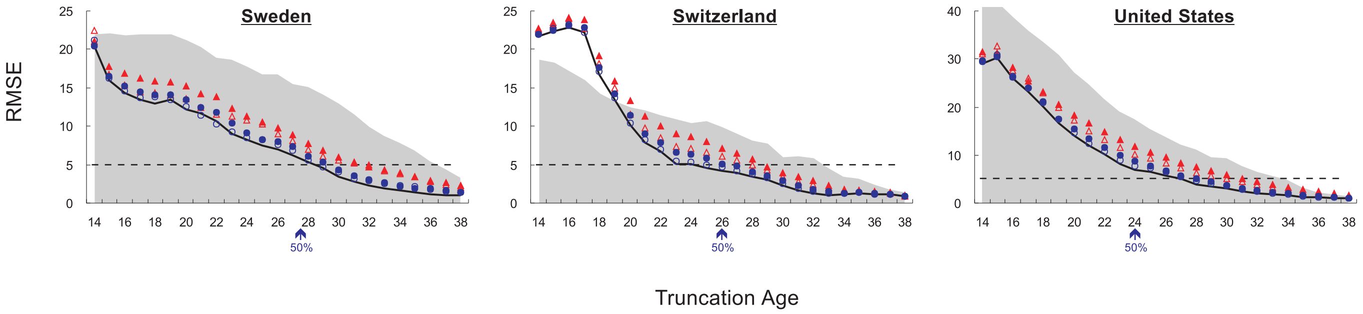

Figure 4: (Continued)

tribution under different situations only. As can be expected, a 10% bias (positive or negative) lead to a poorer performance in curve approximation than a 5% bias unproportionately, indicating that it is worthwhile to keep searching a better CTFR estimate. In addition, at the 50% truncation the APC framework still performs well either under an absolute criterion or a relative one for most countries.

## 6 Summary and conclusions

In this paper, we extend our approach recently developed to incorporate fertility data from 17 European countries, keep the focus on the transformation from a level to a schedule, and carry out validation experiments. Performances of our approach along with the linear extrapolation and the frozen rate methods are evaluated and compared based on the root mean square error (RMSE) criterion. We also investigate the sensitivity of the transformation with different degrees of bias in the level. With few exceptions, this approach performs pretty well under either an absolute or a relative criterion.

# REFERENCES

Cheng, Y. A. (2010). The future of completed cohort fertility in low fertility countries: A comparison of forecasting methods. 2010 Population Association of America meeting, April 14-17, Dallas, U.S.

Cheng, P. C. Roger and Eric S. Lin (2010). Completing incomplete cohort fertility schedules. *Demographic Research* 23(9): 223-256.

Evans, M.D.R. (1986). American fertility patterns: A comparison of white and nonwhite cohorts born 1903–56. *Population and Development Review* 12(2): 267-293.

Lee, R.D. and Carter, L.R. (1992). Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association* 87(419): 659-672.

Myrskylä, M. and J. R. Goldstein (2010). Prediction and error propagation in innovation diffusion models, with applications to demographic processes. MPIDR working paper WP-2010-013.

Willekens, F. and Baydar, N. (1984). Age-period-cohort models for forecasting fertility. Working Paper No. 45, NIDI, The Hague.