On the Possibilities of Predicting Cohort Fertility Measures from Period Fertility Measures: Theory and Empirical Evidence

P. C. Roger Cheng Joshua R. Goldstein

Abstract

This paper establishes a formal relationship between period and cohort measures, responds to the literature casting doubts on the usefulness of period measures as cohort estimators, and proposes three tempo-adjusted predictors of cohort quantum which are easy to implement. Empirical evidence from Canada, the U.S., and 23 European countries suggests that our predictors provide satisfactory estimates and outperform some conventional methods in most cases, particularly when the observed cohort experience is truncated at a very young age.

1 Introduction

Period total fertility rates (PTFRs hereafter), adjusted or unadjusted, have long been considered unreliable in estimating the mean family size of associated cohorts. The fluctuant feature in the time series of adjusted period indicators and their obvious discrepancies from the time series of cohort fertility have been presented in many studies, usually illustrated graphically (e.g. Kohler and Ortega, 2002; Ryder, 1990; Schoen, 2004; Smallwood, 2002; Sobotka, 2003; van Imhoff and Keilman, 2000). These previous comparisons of period and cohort measures were, however, based on a *one-to-one* correspondence.[1] Although van Imhoff and Keilman (2000, p.552) indicated that most of fluctuations caused by adjusted PTFRs persist even

[1] Mostly, period measures and cohort measures are related through the mean age at birth of period or cohort.

after *moderate* smoothing, Cheng and Lin (2010) showed that *strong* smoothing can remove fluctuations and then provide a good estimate of the cohort total fertility rate (CTFR hereafter), implying that the correspondence between period and cohort measures can in fact be *many-to-one*, since smoothing a series of values is identical to taking a weighted average of them.

Distinguished from the demographic translation literature (Keilman, 1994, 2000; Ryder, 1964), this paper derives a simple but definite relationship between period and cohort measures as follows. Let $f(a, t)$ be an age- and time-specific fertility rate, with a and t representing age and time, respectively. For convenience of exposition, a is assumed to fall within the range $[0, \beta]$, where β is a large finite number such that $f(a, t) = 0$ for all $a \geq \beta$ and all t. Summing the rates over childbearing ages for cohort c yields the cohort quantum:

$$\text{CTFR}(c) = \int_0^\beta f(a, c + a)\, da.$$

Similarly, we define the period quantum for calendar year t as

$$\text{PTFR}(t) = \int_0^\beta f(a, t)\, da,$$

and further the period age-specific fertility proportion as $p(a, t) = f(a, t)/\text{PTFR}(t)$, where

$$\int_0^\beta p(a, t)\, da = 1, \tag{1}$$

for any t. Replacing $f(a, c + a)$ with $\text{PTFR}(c + a)\, p(a, c + a)$ leads to

$$\text{CTFR}(c) = \int_0^\beta \text{PTFR}(c + a)\, p(a, c + a)\, da, \tag{2}$$

showing that a CTFR can be expressed as a linear combination of PTFRs with $p(a, c + a)$ as the nonnegative coefficient for PTFR at time $c + a$. Alternatively, a

CTFR can be expressed as a linear combination of Bongaarts-Feeney adjusted total fertility rates (BFs hereafter), i.e.

$$\mathrm{CTFR}(c) = \int_0^{\beta} \mathrm{BF}(c+a)\, w(a, c+a)\, da, \tag{3}$$

where $\mathrm{BF}(t) = \mathrm{PTFR}(t)/\big[1 - r(t)\big]$, $r(t)$ represents the change in the mean age of childbearing at time $t = c + a$, and the corresponding coefficient is $w(a,t) = \big[1 - r(t)\big]\, p(a,t)$.[2] In other words,

Proposition 1.

> *The CTFR of cohort born at time c is a linear combination of the BFs from time c through time $c + \beta$.*

Note that this proposition holds without premise. In the same way, the CTFR can also be linked to period measures as proposed in Kohler and Philipov (2001) or in Goldstein and Cassidy (2010).

Also note that Equations (2) and (3) always hold *regardless of* birth order.[3] A few critiques (e.g. Keilman, 2000; van Imhoff, 2001; van Imhoff and Keilman, 2000) on the BF (or other period measures constructed using age-specific fertility rates) emphasized that the use of incidence rates for *non-repeatable* events is *erroneous.* Specifically, since the shift in tempo affects both the numerator and the denominator of such rates, any index derived using the sum of rates for a given year can introduce extra tempo distortions and thus cannot be interpreted as a proper quantum

[2] Bongaarts and Feeney (1998, p.282–284) had proposed to compare the completed fertility of true cohorts with an average of their adjusted total fertility rates over the years during which the true cohorts were in their choldbearing years. However, they did not provide a formal inference or mathematical proof on this comparison, and the weights they provided (see their footnote 7) also differ from w.

[3] When $f(a,t)$ denotes fertility rate combining all parities, the $\mathrm{BF}(t)$ in Equation (3) differs from the formula derived in Bongaarts and Feeney (1998, Equation 4).

indicator. They are right in the sense of such period measures being utilized in the aforementioned one-to-one manner. But if one concludes further that any variant of the period-sum method must cause the same problem and is thus meaningless, one has unfortunately reached the wrong conclusion by ignoring the relationships between CTFR and period measures provided in Equations (2) and (3). An analogy may be useful. If a professor at Rutgers University (New Brunswick, NJ) wants to attend a conference held in Washington D.C., she/he can drive or take a train, heading south, to the conference venue. Or, she/he can go north first to Newark Liberty International Airport (EWR) and then take a plane to get to D.C. The erroneous conclusion is like saying that if one wants to visit a city located to the south of one's own, she/he must not start going northward.

"When data are available on completed childbearing, measuring cohort fertility is straightforward" (Ní Bhrolcháin, 2011, p.850) and Equation (3) adds no value. However, when cohort childbearing is unfinished Equation (3) becomes a useful starting point to develop approaches in predicting the CTFR. In Section 2, we propose three new tempo-adjusted methods which incorporate the cumulated fertility of incomplete cohorts, overcoming the deficiency that the BF indicator "dispenses with the entire past of the fertility process but for each pair of adjacent periods" (Ní Bhrolcháin, 2011, p.853). Along with two convetional methods, the performance of approaches are evaluated with historical data from Canada, the U.S., and 23 European countries. Section 3 introduces the data, the experiment design, and the estimation details. Empirical results presented in Section 4 suggest that our predictors outperform conventional ones in most cases, particularly when the cohort experience is truncated at very young ages. Section 5 concludes.

2 Methods in Predicting CTFR

When data of cohort childbearing is truncated at age A, the CTFR can be divided into two parts — observed and unfinished — and written

$$\text{CTFR}(c) = \int_0^A f(a, c+a)\, da + \int_A^\beta f(a, c+a)\, da.$$

The strategy to complete the CTFR, i.e. to estimate $f(a, c+a)$ for $a \in (A, \beta\,]$ in the unfinished part, is mostly based on particular assumptions regarding future movements in these rates.

2.1 Conventional Methods

For example, one can assume that recent experience would continue into the future and use the age-specific fertility rates in the latest year observed, $T = c + A$, (or the average fertility rates in the latest few years) as an imputation of the future rates for each age, so that the estimated CTFR becomes

$$\text{est. CTFR}(c) = \int_0^A f(a, c+a)\, da + \int_A^\beta f(a, T)\, da. \tag{4}$$

This method is hereafter denoted by the **Freeze-Rate**, since the future rates are "frozen" as those in the latest year observed.

Alternatively, one can utilize fertility rates in the latest few years observed to derive linear extrapolations age by age, so that the estimated CTFR can be

$$\text{est. CTFR}(c) = \int_0^A f(a, c+a)\, da + \int_A^\beta \big[f(a, T) + \Delta_a \cdot (a - A)\big]\, da. \tag{5}$$

where $\Delta_a = \big[f(a, T) - f(a, T-k)\big]/k$ for some chosen k. We denote this procedure hereafter by the **Linear-Extrapolation**.

Other conventional approaches, such as curve fitting models,[4] the Evans method (Evans, 1986), the Li-Wu model (Li and Wu, 2003), and the Willekens-Baydar approach (Willekens and Baydar, 1984), all work in a similar manner. They make their own particular assumptions regarding the structure behind the data, estimate related parameters with information from either the observed part of the target cohort or the experience of previous cohorts, and then produce their predictions.

Note that we do not intend to include conventional approaches other than the Freeze-Rate and Linear-Extrapolation methods selected for this paper owing to the following considerations: (1) The Freeze-Rate and Linear-Extrapolation methods are easy to implement and the results shown in this paper can be replicated and verified without difficulty, suggesting that such methods can be regarded as a common reference for researchers to evaluate performances across various approaches. (2) The Freeze-Rate method actually outperforms several conventional approaches in predicting the CTFR; see evidence in Cheng and Lin (2010, Table 1).[5]

2.2 Tempo-Ajusted Approaches

Based on the identity that $f(a, c+a) = \mathrm{BF}(c+a)\, w(a, c+a)$ as in Equation (3), one can take the BF value in the latest year observed, $T = c + A$, as an imputation of the future BF values and obtain

$$\text{est. CTFR}(c) = \int_0^A f(a, c+a)\, da + \mathrm{BF}(T) \int_A^\beta w(a, c+a)\, da.$$

This BF-freezing equation, however, is not an estimator of the CTFR until one specifies how to impute the integration of future coefficients $w(a, c+a)$ for $a \in$

[4] For example, a modified version of the Coale-McNeil's (1972) double exponential model (e.g., Bloom 1982; Chen and Morgan 1991), the Hadwiger function (e.g., Chandola, Coleman, and Hiorns 1999), and the linearized Gompertz model (Myrskylä and Goldstein 2010).

[5] They used a different term, the Naive, to name the Freeze-Rate method.

$(A, \beta\,]$.[6] According to:

Proposition 2.

Assume that the period pattern $p(a,t)$ is constant for time $t \in [\,c, c+\beta\,]$.[7] *Then*

$$\int_0^{\beta} w(a, c+a)\,da = \int_0^{A} w(a, c+a)\,da + \int_A^{\beta} w(a, c+a)\,da = 1$$

holds for cohort c.

and

Proposition 3.

Assume that the period pattern $p(a,t)$ is constant for time $t \in (T, c+\beta\,]$, where $T = c + A$. Then

$$\int_A^{\beta} w(a, c+a)\,da = \int_A^{\beta} p(a, T)\,da,$$

holds for cohort c.

We propose the following two esitmators:

$$\text{est. CTFR}(c) = \int_0^{A} f(a, c+a)\,da + \text{BF}(T)\left[1 - \int_0^{A} w(a, c+a)\,da\right] \tag{6}$$

and

$$\text{est. CTFR}(c) = \int_0^{A} f(a, c+a)\,da + \text{BF}(T) \int_A^{\beta} p(a, T)\,da, \tag{7}$$

respectively. These two innovative methods are denoted hereafter by the **Freeze-BF1** and the **Freeze-BF2** since the future BF values are "frozen" as in the latest year observed, T. The proofs of Propositions 2 and 3 are provided in Appendix.

[6] Note that a seemingly straightforward imputation of $w(a, a+c)$ using $w(a,T)$ for $a \in (A, \beta\,]$ is in fact identical to the Freeze-Rate method.

[7] That is, only the first moment of period fertility proportions (i.e., the mean age at birth) is allowed to change over time, while the second and higher order moments of this pattern remain constant.

Table 1: Mean and Standard Deviation for Distribution of Coefficient Sums by Parity

birth order	all	1	2	3+
mean	1.0064	1.0082	1.0046	1.0020
std	0.0273	0.0218	0.0218	0.0253
number of completed cohorts	863	272	272	272

Note: For details regarding our data set refer to Section 3.

Note that the assumption in Proposition 2 is the same as in Bongaarts and Feeney (1998),[8] and that the Freeze-BF1 and the Freeze-BF2 are theoretically identical when the assumption is met. van Imhoff and Keilman (2000) and van Imhoff (2001) showed that this constant shape assumption is violated by the data in countries such as Italy, Netherlands, and Norway. Table 1, however, presents that the distribution of $\int_0^\beta w(a, c+a)\,da$ across all completed cohorts from our data set is highly concentrated around one.

Distinct from previous approaches, this paper proposes another innovative strategy to complete the CTFR without explicit assumptions regarding future fertility rates. We estimate the cumulated proportion of fertility by truncation age A, i.e.

$$\alpha(A, c) = \int_0^A f(a, c+a)\,da \Big/ \mathrm{CTFR}(c),$$

and then use $\int_0^A w(a, c+a)\,da$ as the estimate of $\alpha(A, c)$ to inflate the observed part:

$$\text{est. CTFR}(c) = \frac{\displaystyle\int_0^A f(a, c+a)\,da}{\displaystyle\int_0^A w(a, c+a)\,da}. \tag{8}$$

Hereafter we denote this method by the **Proportion-Inflation**.

[8] The assumption in Proposition 3 is slightly weaker, for the time period is from T rather than from c.

Before examining the empirical performance of methods in predicting the CTFR, it is worthwhile to incorporate (and interpret) the Freeze-BF1, the Freeze-BF2, and the Proportion-Inflation in a common context implied by Proposition 2 that the CTFR of cohort c can be regarded as not only a linear combination but also a weighted average of the BF values throughout the whole fertility profile (from time c through time $c + \beta$).[9] The Freeze-BF1 and Freeze-BF2 methods in effect impute the weighted average of BF values in the unfinished part (from time T through time $c + \beta$) by BF(T), while the Proportion-Inflation method directly assumes that the weighted average of observed BF values (and the weighted average of unfinished BF values as well) equals the CTFR(c). [10] One can therefore expect a poor performance from these tempo-adjusted methods when the weighted average of unfinished BFs severely deviates from BF(T) or CTFR(c). The deviation is usually caused by a strong quantum effect.

3 Data, Experiment Design, and Estimation Details

The data employed in this study are age-specific fertility rates (ASFRs) by one-year period and by single-year age group, taken from the Human Fertility Database[11] and the Eurostat Database, downloaded in August, 2011. We restrict our analysis to the

[9]A linear combination casts no restrictions on coefficients, while a weighted average requires the sum of all coefficients to equal one.

[10]An analogy may be useful. Suppose that there are a number of people intending to use an elevator with a weight limit and that some of them have entered the elevator. If one wants to predict the total average weight but has only observed the weight of each individual who has entered, one can assume that the average weight of the people waiting outside equals the weight of the last person having entered, or one can use the average weight of the people inside as an estimate directly. The logic of the former is the same as that of the Freeze-BF methods, while the logic of the latter is identical to that of the Proportion-Inflation method. One can regard the only difference between the Freeze-BF1 and Freeze-BF2 methods as how they estimate the number of those who wait outside.

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

set of single-year age groups 15-44 and exclude countries whose data contain less than 10 completed cohorts, so not all countries listed in the databases are included. In addition, we use the separate schedules for East and West Germany, as well as for England/Wales and Scotland in the U.K., rather than their combined data. Thus there are 27 schedules in total, as listed in Table 2. Since non-parity specific data are more readily available than parity-specific data, there are 863 completed cohorts from the former and 272 from the latter, respectively.

For each particular cohort, 25 CTFR predictions with partial fertility information from age 15 through a chosen truncation age (varying between 19 and 43) are produced accordingly. To measure the extent to which an estimated CTFR approximates to the actual CTFR, the prediction error (PE) index is designed to measure how much of the proportion of unfinished fertility has not been correctly estimated, i.e.

$$\text{PE} = \frac{\text{est. CTFR} - \text{CTFR}}{\text{CTFR} - \text{obs. CTFR}} \times 100\%,$$

so that comparisons can be made across various degrees of truncation. For example, suppose that the actual CTFR is 2.0 and the prediction is 1.8, the PE is -16.67% when the observed fertility (obs. CTFR) is only 0.8, indicating that 83.33% of the unfinished part has been estimated. In contrast, if the actual and estimated CTFR remain unchanged while the observed fertility increases to 1.2, then the PE becomes -25.00%, indicating that only 75.00% of the unfinished part has been estimated. The positive or negative sign of PE represents whether the CTFR is over- or under-estimated. Refering to Sobotka (2003, Table 12), we evaluate the performance of an estimate by the following classification considering an absolute PE as very good ($< 7.5\%$), good (7.5–12.4%), average (12.5–19.9%), poor (20.0–37.5%), and very

Table 2: Data from the Human Fertility Database and the Eurostat

	all-birth-combined		parity-specific	
country	periods	completed cohorts	periods	completed cohorts
	from the Human Fertility Database			
Austria	1951–2008	1936–1964 (29)	1984–2008	NA
Canada	1921–2007	1906–1963 (58)	1944–2007	1929–1963 (35)
Czech Republic	1950–2009	1935–1965 (31)	1950–2009	1935–1965 (31)
Estonia	1959–2009	1944–1965 (22)	1959–2009	1944–1965 (22)
Finland	1939–2009	1924–1965 (42)	1982–2009	NA
France	1946–2009	1931–1965 (35)	NA	NA
Germany				
East	1956–2009	1941–1965 (25)	1956–1989	NI
West	1956–2009	1941–1965 (25)	NA	NA
Hungary	1950–2009	1935–1965 (31)	1952–2009	1937–1965 (29)
Netherlands	1950–2009	1935–1965 (31)	1950–2009	1935–1965 (31)
Russia	1959–2009	1944–1965 (22)	1959–2009	1944–1965 (22)
Slovakia	1950–2009	1935–1965 (31)	1950–2009	1935–1965 (31)
Sweden	1891–2008	1876–1964 (89)	1970–2008	1955–1964 (10)
Switzerland	1944–2007	1929–1963 (35)	NA	NA
U.K.				
England/Wales	1938–2009	1923–1965 (43)	NA	NA
Scotland	1945–2009	1930–1965 (36)	NA	NA
U.S.	1917–2006	1902–1962 (61)	1917–2006	1902–1962 (61)
	from the Eurostat			
Belgium*	1954–2009	1939–1965 (27)	NA	NA
Bulgaria	1960–2009	1945–1965 (21)	NA	NA
Denmark	1950–2009	1935–1965 (31)	NA	NA
Greece	1961–2009	1946–1965 (20)	NA	NA
Iceland	1963–2009	1948–1965 (18)	NA	NA
Italy	1952–2008	1937–1964 (28)	NA	NA
Lithuania	1970–2009	1955–1965 (11)	NA	NA
Norway	1961–2009	1946–1965 (20)	NA	NA
Portugal	1950–2009	1935–1965 (31)	NA	NA
Spain	1971–2009	1956–1965 (10)	NA	NA

Note:
1. When a country is included in both databases, we prioritize data from the Human Fertility Database.
2. Countries whose data are insufficient to construct at least 10 completed cohorts (covering age groups 15–44) will be excluded from the analysis.
3. In parentheses are numbers of completed cohorts.
4. NA denotes that data are not available, while NI denotes that data are not included in the analysis because the number of completed cohorts is less than 10.
5. The 1917–1932 U.S. data are taken from Heuser (1976).

* Data in 2001 and 2002 are missing for Belgium. We use spline smoothing to interpolate these values by age.

poor ($> 37.5\%$).[12] In addition, due to the fact that a truncation age A corresponds to a truncation percentile $\alpha(A, c)$ which varies across cohorts and birth orders, we construct PE distributions (or alternatively distributions of the absolute PE) by truncation percentile rather than by truncation age to evaluate the performance of an approach.

This paper investigates and compares the performance of five aforementioned methods, two of which are conventional and three tempo-adjusted. Specific estimation details:

1. The Freeze-Rate method, as described in Equation (4), uses fertility rates in the latest year observed.[13]

2. The Linear-Extrapolation method, as specified in Equation (5), sets $k = 4$ and forces the estimated fertility rates to level after 5 steps of extrapolation. A negative extrapolation is replaced with zero.

3. To apply the Freeze-BF1 and Propotion-Inflation methods, we need to determine the value of $w(a, c + a)$, $a \in [0, A]$, which requires a calculation of $r(t)$, the change in the mean age of childbearing (MAC) at time t, $t \in [c, c+A]$. The calculation basically follows the algorithm specified in Bongaarts and Feeney (1998, footnote 6), $[r(t+1)-r(t-1)]/2$. But for the first year c and the last year T of the observed part, we compute $r(c)$ and $r(T)$ with $\text{MAC}(c+1) - \text{MAC}(c)$ and $\text{MAC}(T) - \text{MAC}(T-1)$, respectively.

[12]Because Sobotka's error index takes the completed CTFR rather than the unfinished CTFR as the denominator, a conversion of cutting points is called for. Such a conversion is made by assuming the truncation percentile at mean age of childbearing to be 60%.

[13]The authors have tried using average fertility rates in the five latest years observed but did not get better results when predicting the CTFR.

4 Results

4.1 Average Performance of Absolute Prediction Error

Focusing on the absolute value of prediction errors, Table 3 presents the average performance (i.e. the mean absolute PE) of each method across various ranges of truncation percentile and birth orders. Possibly the most striking result observed in the table is that: when the truncation percentile is in the range of $[10\%, 30\%)$ and the birth order is 1, the Proportion-Inflation method yields a mean absolute prediction error as low as 4.99%, outperforming all the other methods investigated in this paper, especially the two conventional approaches. This result is so remarkably impressive because the truncation percentile of the mean age at birth, which is often adopted as a reference age, generally falls between 50% (the median) and 75% (the third quartile), indicating that such "very good" prediction performance is obtained with relatively little information. In addition, the Proportion-Inflation method extends its advantage by providing "very good" or "good" predictions in other ranges of truncation percentile and in the case of birth order 2 when the truncation percentile is no more than 75%. Only when the truncation percentile is in $[75\%, 85\%)$ where a large proportion of fertility has been observed, the leading place for predicting the CTFR of birth orders 1 and 2 is earned by the Freeze-BF2 method. Moreover, as for the all-birth-combined fertility, the Freeze-BF2 method performs the best or the second best while its performance in all ranges of truncation percentile is consistently classified as "average". Finally, the Linear-Extrapolation method always performs the best if the birth order is 3 and above, but its performance never falls into a category better than the "average".

Table 3: Average Performance by Method, Truncation Percentile, and Birth Order

birth order	N	Freeze-BF1	Freeze-BF2	Proportion Inflation	Freeze-Rate	Linear Extrapolation
			truncation percentile $\in [10\%, 30\%)$			
all	2,627	13.57	*13.46*	**13.32**	17.40	18.66
1	585	6.17	*6.04*	**4.99**	10.19	11.74
2	700	9.31	*9.15*	**7.82**	11.91	15.94
3+	829	*27.35*	27.36	31.20	27.62	**27.22**
			truncation percentile $\in [30\%, 50\%)$			
all	2,344	*13.98*	**13.71**	14.05	17.28	16.79
1	507	*6.74*	*6.74*	**5.52**	11.74	11.03
2	581	*10.28*	10.38	**8.69**	13.70	15.71
3+	664	23.53	23.69	29.90	*23.38*	**22.62**
			truncation percentile $\in [50\%, 65\%)$			
all	1,909	*14.35*	**14.05**	15.40	17.44	16.41
1	436	7.79	*7.65*	**6.49**	12.74	11.07
2	503	*10.83*	11.15	**9.61**	14.18	14.62
3+	530	22.09	22.20	31.48	*21.64*	**21.15**
			truncation percentile $\in [65\%, 75\%)$			
all	1,536	*14.78*	**13.55**	17.14	17.38	15.63
1	370	9.47	*8.60*	**8.43**	13.20	10.91
2	376	11.77	*11.53*	**11.26**	15.18	15.71
3+	419	21.18	20.73	32.27	*19.59*	**19.09**
			truncation percentile $\in [75\%, 85\%)$			
all	2,024	16.83	**13.94**	20.92	17.76	*15.40*
1	527	12.16	**9.13**	11.65	13.55	*10.87*
2	518	13.83	**11.80**	*13.72*	15.16	15.66
3+	530	20.92	18.63	35.20	*18.12*	**16.86**

Note: Numbers in the table are mean absolute prediction errors (in %). For each pair of birth order and truncation percentile, the best performance is noted in bold and the second best in italic. We classify an absolute prediction error as very good (< 7.5%), good (7.5–12.4%), average (12.5–19.9%), poor (20.0–37.5%), and very poor (> 37.5%).

4.2 Distribution Function in Absolute Prediction Error

Figure 1 further provides a comprehensive comparison over the five methods by depicting their distribution functions (i.e. cumulated density functions) in absolute PE. When fertility is non-parity specific, the distribution curve of the Freeze-BF2 method lies mostly above the others (especially of the conventional methods), regardless of the range of truncation percentile. By observing the distributional curve of the Freeze-BF2 method, we note that no more than 40% (60%) of absolute prediction errors are categorized "very good" ("good") while more than 20% are "poor". This reveals the distributional details of the corresponding average numbers shown in Table 3.

When fertility is of first birth, in contrast, the distribution curve of the Proportion-Inflation method lies mostly above the other curves, except when the truncation percentile is in the range of $[75\%, 85\%)$. In the case of truncation percentile below 65% (i.e. in the first three ranges), about 70% (90%) of absolute prediction errors by the Proportion-Inflation method are categorized "very good" ("good") while less than 5% (none) are "poor" ("very poor"). In the truncation percentile range of [65%, 75%), the curves of the Freeze-BF2 and Proportion-Inflation methods are virtually indistinguishable, but with a trunction percentile of $[75\%, 85\%)$ the Freeze-BF2 method dominates. Essentially similar but slightly worse results are found when fertility is of second birth; all curves lean further to the right.

When fertility is of birth order 3 and above, however, all methods exhibit a certain proportion of absolute prediction errors larger than 70% since no curve reaches 1.0 within the range presented in the figure. At least 10% of absolute prediction errors are larger than 37.5%, the threshold beyond which a PE is categorized "very

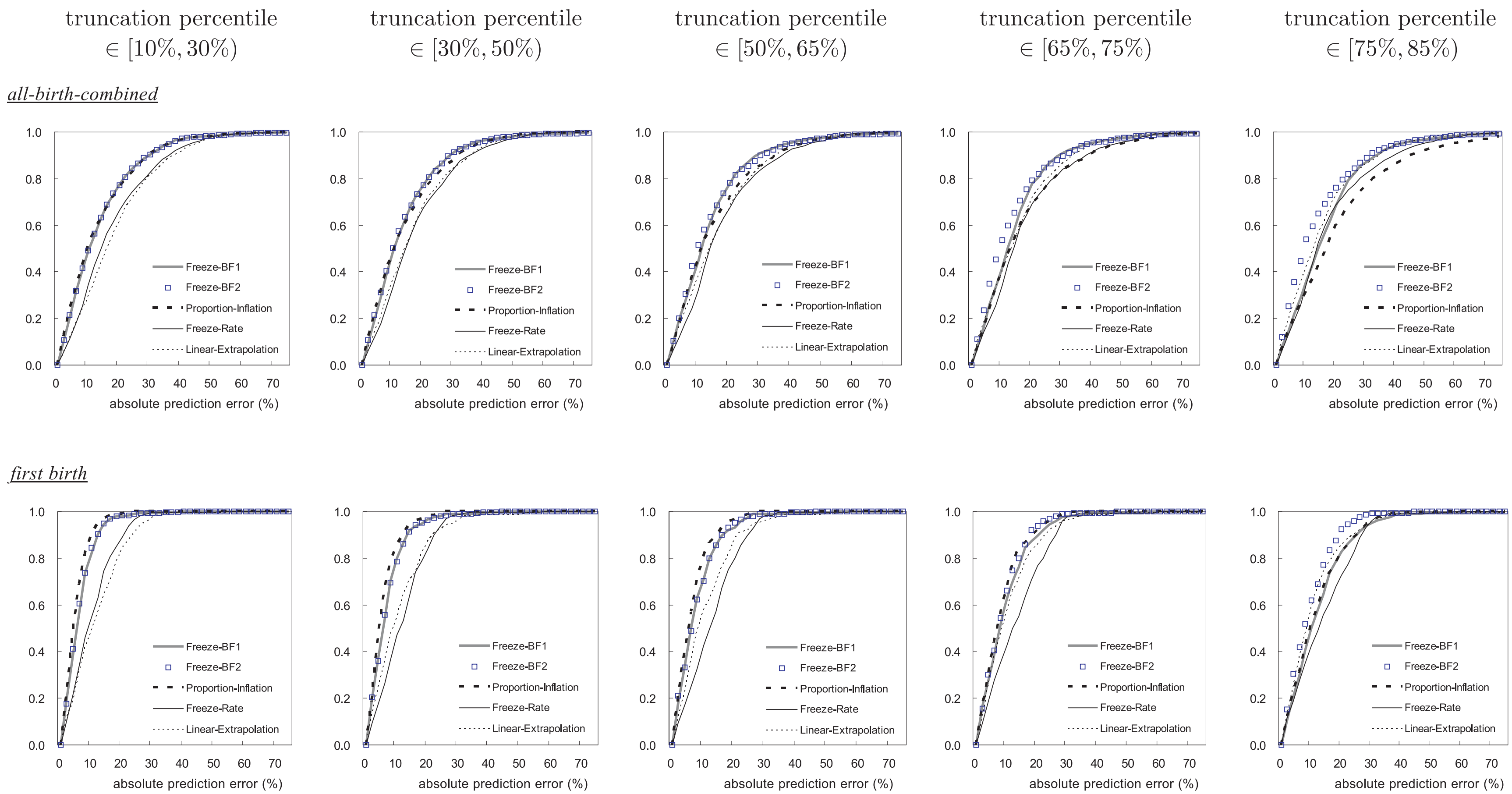

Figure 1: Cumulated Density Functions of Absolute Prediction Errors by Method

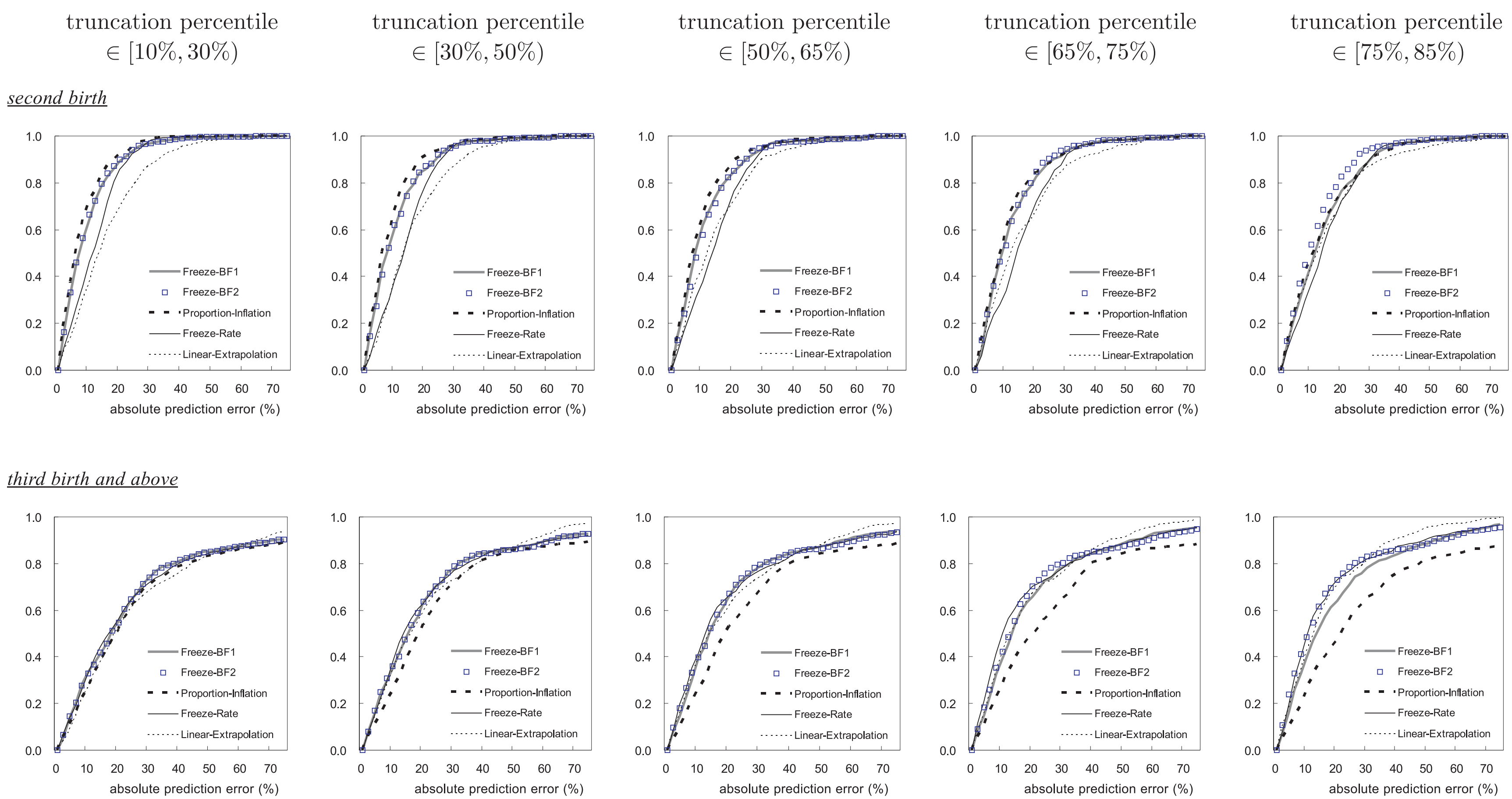

Figure 1: continued

poor", regardless of method or truncation percentile range, suggesting that none of the methods investigated in this paper can provide satisfactory predictions for high birth orders. Furthermore, the distribution curve of the Linear-Extrapolation method lies mostly below that of the Freeze-BF2 method for lesser prediction errors but lies above when larger prediction errors are cumulated. The fact that "very poor" predictions are less frequently observed when using the Linear-Extrapolation method explains why this method yields the lowest mean absolute PEs in Table 3.

4.3 Further Examination Across Cohorts

As mentioned in Section 2, a poor performance by tempo-adjusted methods is expected when there exists a strong quantum effect. In this subsection, we analyze prediction errors in further detail by dividing birth cohorts into three subgroups: cohorts 1910–30, cohorts 1935–50, and cohorts 1950–65. Note that we select Canada, Netherlands, Sweden, and the U.S. as representative countries because of their wide data ranges, but not every subgroup contains cohort information from all of the four representative countries due to data availability. Figure 2 presents their CTFR and MAC curves by birth order, providing background information about the quantum and tempo effects for the three subgroups. Respectively, women in these three subgroups experienced their main childbearing ages during the baby-boom, baby-bust, and post baby-bust periods, accompanied with an increasing, decreasing, and relatively flat CTFR. In addition, the corresponding MAC curves show that they were advancing, in the transition from advancement to postponement, and postponing childbearing. Overall, the change in the all-birth-combined CTFR is mainly decided by that of birth order 3 and above. The extent of change in the CTFR increases

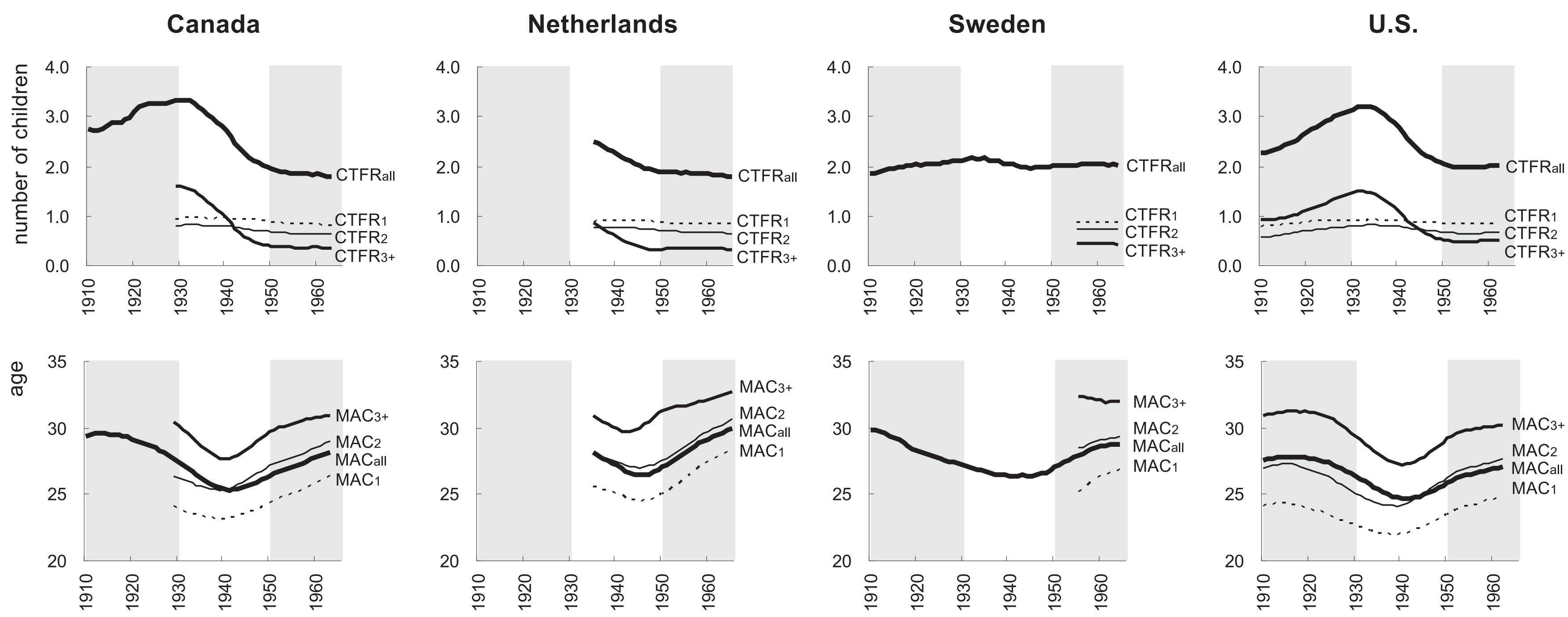

Figure 2: CTFR and MAC by Birth Order for Canada, Netherlands, Sweden, and the U.S.

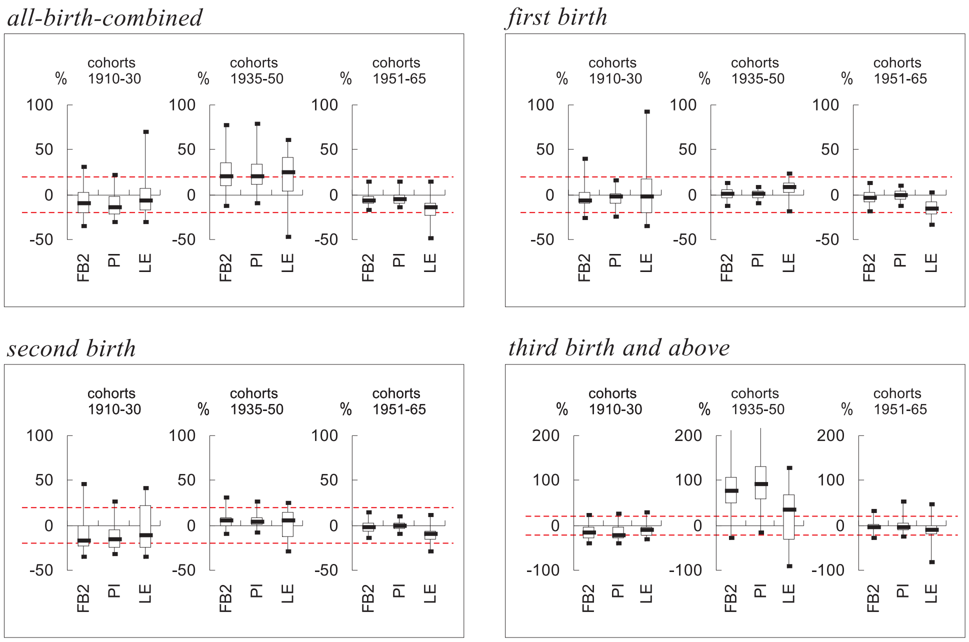

Figure 3: Distribution Comparison Across Three Cohort Subgroups

Note: 'FB2', 'PI', and 'LE' represent the Freeze-BF2, Proportion-Inflation, and Linear-Extrapolation methods, respectively, and the truncation percentile is in the range of $[10\%, 30\%)$. Broken lines in each panel mark the $+20\%$ and -20% thresholds beyond which a PE is categorized "poor".

as the birth order gets higher for the first two subgroups, but all CTFR curves remain relatively flat for subgroup 3. More specifically, the absolute average change in quantum across the cohort scale is much larger for subgroup 2 than for subgroup1 in the birth order 3+ and thus all-birth-combined cases, but slightly smaller in the birth order 1 and birth order 2 cases. In contrast, the absolute average change in tempo across the cohort scale is larger for subgroup 3 than for subgroup 1, regardless of birth order.

Figure 3 displays prediction error distributions of the Freeze-BF2, Proportion-Inflation, and Linear-Extrapolation methods for three subgroups of cohorts when the truncation percentile is in the range of $[10\%, 30\%)$. For clarity, prediction error distributions of the Freeze-BF1 and Freeze-Rate methods are not shown, and figures regarding other truncation percentile ranges are omitted due to the similarity in pattern. They are available upon request from the authors. Each box-and-whisker plot provides five summary statistics of a PE distribution: the bottom and top of the box are the lower and upper quartile, the band near the middle of the box is the median, and the ends of the whisker represent the minimum and maximum. In each panel we mark the +20% and −20% thresholds beyond which a PE is categorized "poor" with broken lines.

Figure 3 provides several informative observations by which one can conclude that the performance of CTFR estimators is mainly influenced by the quantum effect rather than the tempo effect. First of all, the most significantly deviated PE distribution occurs in the case of third birth and above for subgroup 2; quite a lot of PEs fall outside the ±20% interval (i.e. in the "poor" or "very poor" regions). Second, the PE distribution deviates much more for subgroup 2 than for subgroup 1 in the all-birth-combined and birth order 3+ cases but slightly less in the first and second birth cases. These two observations are consistent with the aforementioned patterns in quantum across subgroups and birth orders. In contrast, the deviation mostly falls within the ±20% interval for subgroup 3 regardless of birth order, given the fact that women in this subgroup experienced a rather mild quantum change but quite a strong change in tempo. Also note that a Linear-Extrapolation estimate, compared with Freeze-BF2 and Proportion-Inflation estimates, is more likely to deviate from the actual CTFR for subgroup 3 regardless of birth order, which is not unexpected since there is no tempo adjustment designed in this estimator. To sum up, Figure 3 suggests that: when the quantum effect is relatively mild, the tempo-adjusted methods should take priority over the conventional methods in selective implementation. In the presence of a strong quantum change, although the tempo-adjusted methods are outperformed by the conventional methods, there is in fact no method that reliably provides satisfactory CTFR predictions.

5 Summary

This paper establishes a formal relationship between period and cohort measures with which one can appropriately respond to the literature that casts doubts on the usefulness of period measures as cohort estimators. Specifically, when adjusted or unadjusted period measures are utilized in a many-to-one manner rather than in the conventional one-to-one manner, the related applications are not limited to repeatable events. Furthermore, based on the formal relationship, this paper proposes three tempo-adjusted predictors of cohort quantum which are easy to implement. Empirical evidence from Canada, the U.S., and 23 European countries suggests that our predictors provide satisfactory estimates and outperform the conventional Freeze-Rate and Linear-Extrapolation methods in most cases, particularly when the observed cohort experience is truncated at a very young age. As for cases where there exists a strong quantum effect, our detailed analysis shows that there is so far no ideal method (at least among the five investigated in this paper) whose prediction of the CTFR is statistically reliable, implying that future studies along this line should devote some effort to overcoming the deviation of prediction caused by big quantum changes.

Appendix

A1 Proof of Proposition 2

Proof. By the assumption of a constant period pattern for $p(a,t)$ at any time t,

$$p(a,t) = p(a - R(t), t_0) \quad \text{with} \quad R(t) = \int_{t_0}^{t} r(k)\, dk.$$

This states that $p(a,t)$ at time t has the same shape as $p(a,t_0)$ at an arbitrary chosen time t_0, but has shifted along the age axis by $R(t)$ years. For convenience, let $t_0 = c$. Also, let

$$u = a - R(c+a) = a - \int_{c}^{c+a} r(k)\, dk$$

so that $u = 0$ if $a = 0$ and $u = \beta' = \beta - R(c+\beta)$ if $a = \beta$. Differentiating both sides of the equation gives $du = \big[1 - r(c+a)\big]\, da$. Thus,

$$\begin{aligned} \int_0^{\beta} w(a, c+a)\, da &= \int_0^{\beta} \big[1 - r(c+a)\big]\, p(a, c+a)\, da \\ &= \int_0^{\beta'} p(u,c)\, du = 1 \end{aligned}$$

can be verified by the property mentioned in Equation (1). Note that β' can be a large finite number such that $p(a,c) = 0$ for all $a \geq \beta'$ by choosing an appropriate β value. □

A2 Proof of Proposition 3

Proof. Assuming that the period pattern for $p(a,t)$ is constant for time $t \in [A, \beta]$, we apply the same technique as in the proof of Propostion 2 but let $t_0 = c + A = T$. Therefore,

$$p(a,t) = p(a - R(t), c+A) \quad \text{with} \quad R(t) = \int_{c+A}^{t} r(k)\, dk.$$

Let

$$u = a - R(c+a) = a - \int_{c+A}^{c+a} r(k)\,dk$$

so that $u = A$ if $a = A$ and $u = \beta'' = \beta - R(c+\beta)$ if $a = \beta$. Differentiating both sides of the equation gives $du = \big[1 - r(c+a)\big]\,da$. Thus,

$$\begin{aligned} \int_A^\beta w(a, c+a)\,da &= \int_A^\beta \big[1 - r(c+a)\big]\,p(a, c+a)\,da \\ &= \int_A^{\beta''} p(u, c+A)\,du = \int_A^{\beta''} p(u, T)\,du. \end{aligned}$$

Note that β'' can be a large finite number such that $p(u, T) = 0$ for all $u \geq \beta''$ by choosing an appropriate β value. $\square$

References

Bongaarts, John and Griffith Feeney (1998), "On the Quantum and Tempo of Fertility", *Population and Development Review*, 24(2), 271–291.

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.

Goldstein, Joshua R. and Thomas Cassidy (2010), "Cohort Postponement and Period Measures", MPIDR Working Paper WP 2010-015.

Heuser, Robert L. (1976), *Fertility Tables for Birth Cohorts by Color: United States 1917–73*, DHEW Publication No. (HRA) 76-1152. Rockville, MD: National Center for Health Statistics.

Keilman, Nico (1994), "Translation Formulae for Non-Repeatable Events", *Population Studies*, 48(2), 341–357.

——— (2000), "Demographic Translation: From Period to Cohort Perspective and Back", Memorandum No. 06/2000, Department of Economics, University of Oslo.

Kohler, Hans-Peter and José Antonio Ortega (2002), "Tempo-Adjusted Period Parity Progression Measures, Fertility Postponement and Completed Cohort Fertility", *Demographic Research*, 6, 92–144.

Kohler, Hans-Peter and Dimiter Philipov (2001), "Variance Effects in the Bongaarts-Feeney Formula", *Demography*, 38(1), 1–16.

Li, Nan and Zheng Wu (2003), "Forecasting Cohort Incomplete Fertility: A Method and An Application", *Population Studies*, 57(3), 303–320.

Ní Bhrolcháin, Máire (2011), "Tempo and the TFR", *Demography*, 48, 841–861.

Ryder, Norman B. (1964), "The Process of Demographic Translation", *Demography*, 1(1), 74–82.

——— (1990), "What Is Going to Happen to American Fertility", *Population and Development Review*, 16(3), 433–454.

Schoen, Robert (2004), "Timing Effects and the Interpretation of Period Fertility", *Demography*, 41(4), 801–819.

Smallwood, Steve (2002), "The Effect of Changes in Timing of Childbearing on Measuring Fertility in England and Wales", *Population Trends*, 109, 36–45.

Sobotka, Tomáš (2003), "Tempo-Quantum and Period-Cohort Interplay in Fertility Changes in Europe: Evidence from the Czech Republic, Italy, the Netherlands and Sweden", *Demographic Research*, 8, 151–214.

van Imhoff, Evert (2001), "On the Impossibility of Inferring Cohort Fertility Measures from Period Fertility Measures", *Demographic Research*, 5, 23–64.

van Imhoff, Evert and Nico Keilman (2000), "On the Quantum and Tempo of Fertility: Comment", *Population and Development Review*, 26(3), 549–553.

Willekens, Frans and Nazli Baydar (1984), "Age-period-cohort models for forecasting fertility", *Working Paper No. 45, NIDI, The Hague.*