

Tempo and the TFR

Máire Ní Bhrolcháin

Published online: 4 June 2011
© Population Association of America 2011

Abstract Tempo effects in period fertility indicators are widely regarded as a source of bias or distortion. But is this always the case? Whether tempo change results in bias depends, in the view advanced here, on the measure used, the meaning of bias/distortion, and the objective of analysis. Two ways of construing bias in period measures are suggested, and their relevance is discussed in the context of five broad purposes for measuring period fertility: describing and explaining fertility time trends, anticipating future prospects, providing input parameters for formal models, and communicating with nonspecialist audiences. Genuine timing effects are not biasing when period fertility is the explanandum but are distorting when the aim is to estimate cohort fertility. Alternatives to tempo adjustment are available that are a more defensible solution to the issue of timing change. Tempo adjustment could be more fruitfully considered a form of modeling rather than empirical measurement. The measurement of period fertility could be improved by relying more on a statistical approach and less on indicators based on stable assumptions. Future progress will depend on integrating research on measurement with substantive investigation.

Keywords Measurement · Period fertility · Total fertility rate · Tempo adjustment · Forecasting

Introduction

Are conventional measures of period fertility biased? If so, what should we do about it? These questions are of long standing, and have renewed practical importance currently because of the delayed childbearing of recent decades in industrialized countries. Bongaarts and Feeney (1998, 2000, 2006) have responded to this

M. Ní Bhrolcháin (✉)
Centre for Population Change and Division of Social Statistics, University of Southampton,
Southampton SO17 1BJ, UK
e-mail: mnb2@soton.ac.uk

phenomenon with an innovative methodological solution. What they regard as tempo-related bias in period indicators of fertility can, they suggest, be removed by adjusting for tempo. The Bongaarts-Feeney approach to the total fertility rate (TFR) has attracted much interest, and the ideas have been extended in several ways (Ediev 2008; Inaba 2007; Kohler and Ortega 2002a, b; Kohler and Philipov 2001; Zeng and Land 2001, 2002). Demographic opinion, however, is divided on the merits of tempo adjustment, the procedure having been the subject of much debate (Barbi et al. 2008; Frejka and Calot 2001; Kim and Schoen 2000; Lesthaeghe and Willems 1999; Schoen 2004; Smallwood 2002b; Sobotka 2004a; van Imhoff 2001; van Imhoff and Keilman 2000). Yet the idea that tempo effects are a source of bias in period fertility measures permeates much recent discussion and remains influential. For example, 9 of the 19 papers in the Frejka et al. (2008) state-of-the art collection on fertility trends in Europe refer to tempo effects as distorting, as do Billari et al. (2006); Bongaarts (2008); Gauthier and Philipov (2008); and Goldstein et al. (2009).

Methodological commentary has focused hitherto on the properties and performance of the Bongaarts-Feeney measure and related indices. This article takes a different approach, qualifying the assertion that tempo effects are distorting, and examining the rationale for tempo adjustment in principle. Demography has a can-do pragmatism, particularly in the area of measurement and estimation. While this pragmatic approach is highly productive in applied settings, the downside is a propensity for conceptual casualness. The present article attempts to unpack and examine the tight mesh of ideas underpinning conventional fertility measurement and to highlight dimensions of the issue that have been neglected in recent debates.

That period fertility measures are influenced by timing effects is generally agreed. But are tempo effects a source of bias/distortion? The answer depends on three things: the period measure used, the meaning of bias/distortion, and the objective of measuring period fertility in any particular case. The first and second of these issues are discussed briefly in the next two sections. The core of the article, in the section that follows, addresses the third. I identify five principal reasons for measuring period fertility, consider for each in turn whether tempo effects are biasing, and assess, if so, whether adjustment solves the problem. A concluding section summarizes the article's findings in tabular form and discusses some strategic aspects of fertility measurement.

Period Measures of Fertility

Period measures of fertility are of two main types: rates of varying levels of specificity (e.g., specific for age, parity, duration) and summary measures obtained by cumulating such rates additively or multiplicatively. Cumulating period rates in this way may be regarded simply as providing a basic, single-figure statistical summary indicator (Guillot 2006; Ní Bhrolcháin 1994). In this guise, the style of summary is arbitrary and subject only to evaluation on statistical grounds. The TFR could be replaced by, for example, TFR/35, representing the average age-specific rate during the year, without any change in meaning or statistical content.

The TFR is usually seen as something more than this. It is defined classically as the mean family size of a hypothetical cohort experiencing throughout its lifetime

the age-specific rates of a period (see, e.g., Pressat and Wilson 1988). Hence, the TFR and indices constructed along similar lines, such as period parity progression ratios, are known as *synthetic cohort* or *hypothetical cohort* measures.

The TFR has no theoretical status, having been devised for purely practical ends. As an indicator of period fertility, it has several defects that are touched on only briefly in this article (but see Ní Bhrolcháin 1992, 2008). Nevertheless, the TFR is extremely useful in practice because it provides an informative, age-standardized reading on a population's fertility at a point in time, using relatively simple data inputs. It is widely employed as a conventional, pragmatic indicator of the period level of fertility, although little thought is normally given to what precisely "the level of fertility" at a point in time may mean. Reference to the TFR throughout this article is, unless otherwise specified, to the indicator itself without further interpretation.

Bias or Distortion in Period Measures

In the literature on tempo adjustment, the TFR is regarded as biased or distorted by tempo effects. Terminology varies, with some sources using the term *bias* alone or interchangeably with *distortion* (Bongaarts 2002; Bongaarts and Feeney 1998; Kohler and Ortega 2002b; Kohler and Philipov 2001; van Imhoff and Keilman 2000; Zeng and Land 2001, 2002) and other, particularly more recent, sources referring exclusively to distortion (Bongaarts 1999, 2008; Bongaarts and Feeney 2000; Frejka et al. 2008; Frejka and Ross 2001; Kohler et al. 2002; Kohler and Ortega 2002a; Schoen 2004; Sobotka 2004a). The two terms are used interchangeably in the recent fertility literature, mean the same thing, and are considered synonymous in the present article.

Tempo-induced bias is prominent in recent writing on fertility, but its precise meaning is usually not explicit and needs clarification. Sources on tempo adjustment are of little help, since they offer no formal definitions of bias/distortion in the TFR other than to equate it with the presence of timing change. For example, Bongaarts and Feeney (1998:272) state that the "tempo component equals the distortion that occurs due to timing changes." Bongaarts and Feeney (2006:116) elaborate further by defining a tempo distortion as "an inflation or deflation of a period quantum or tempo indicator, such as birth, . . . that results from a rise or fall in the mean age at which the event occurs." However, it is not clear with respect to what the "inflation or deflation" occurs. Statistical bias is not the issue, as Zeng and Land (2002) observe, since there is no question of a probability distribution for the TFR. How can we sensibly construe the notion of bias/distortion? Two meanings of bias in the TFR appear to be implicit in recent discussion: confounding and measurement bias. Confounding is at issue when the TFR is used to assess *period trends* in fertility, and measurement bias is relevant when the TFR is seen as representing the *level* of fertility.

Bias/Distortion Understood as Confounding

The claim that the TFR is biased/distorted by period tempo change can be understood as meaning that *change* in the TFR gives a biased/distorted measure of

period *change* in fertility. Leaving aside the question of level, the change in fertility from t to $t + d$ may be inaccurately measured by $\text{TFR}_{t+d} - \text{TFR}_t$ or by $\text{TFR}_{t+d} / \text{TFR}_t$, because of confounding.¹ Since the TFR is standardized only for age, it can be influenced by period change in the exposure distribution in two ways: the composition of the female population can vary from year to year (a) by parity and (b) by age or duration within parity. Since birth rates vary by each of these factors, change in the period TFR may reflect compositional change in these respects rather than in specific birth rates. Confounding of calendar time with the exposure distribution is designated here as bias of type A. Change in both level and tempo can give rise to such confounding, and both are thus a source of type A bias. A pure change in level has this effect because it can alter the parity composition of the population at risk. Timing change is a source of confounding because advance or delay in births of a given order alters the parity composition of the population at risk, as well as the distribution by age or duration since previous birth of women of a given parity. The effect is particularly marked when the age at first birth changes.

Note, however, that it is only the compositional results of tempo change—sometimes described as spurious timing effects—that give rise to confounding. They should be distinguished from genuine tempo effects, which manifest as a shift along the age or duration axis of parity-specific period birth rates. In subsequent discussion, tempo usually refers to genuine or true tempo effects, and this will be clear from the context.

Measurement Bias

The TFR might also be considered biased/distorted in the sense that it suffers from measurement bias. Such bias exists when there is a systematic difference between a measure of something and its true value. If the TFR is distorted in this sense, there must be some real-world phenomenon or theoretical construct that it misrepresents. What is this? What is it that a period fertility indicator is *supposed* to measure? The literature on adjusting period fertility indicators for tempo, like the discipline as a whole, is less than clear on the subject. If measurement is biased, three interpretations are possible, designated here B1 to B3.

Definitional (B1)

One reading is that any measure of period fertility that is influenced by genuine timing effects is distorted by definition, and that it is a construct “period fertility” that is wrongly measured by the TFR. Some scholars appear to espouse a view close to this, seeing tempo effects, in and of themselves, as biasing to period measures. For

¹ Bias in an estimate of change can be assessed separately from bias in the measure itself. A biased measure of the period level of a phenomenon may nevertheless give an unbiased estimate of absolute or relative period change. Suppose the level of fertility in t can be meaningfully represented by a single figure, f_t , and is measured with bias by an indicator f_t^* . Under case A, if $f_t^* = a + f_t$, measured absolute change between t and $t + d = f_{t+d}^* - f_t^* = a + f_{t+d} - a - f_t = f_{t+d} - f_t$. Under case B, if $f_t^* = kf_t$, measured relative change $= f_{t+d}^* / f_t^* = kf_{t+d} / kf_t = f_{t+d} / f_t$. Hence the estimated absolute (case A) and relative (case B) change is unbiased even though the measures themselves, in each case, are biased. Clearly, a measure of change using an unbiased measure will itself be unbiased.

example, the earlier quotation from Bongaarts and Feeney implies that timing change is by definition distorting (see also Bongaarts 2002, 2008; Bongaarts and Feeney 2000; Kohler and Philipov 2001; Zeng and Land 2002). This is a defensible position for which a case might, in principle, be made. But it would have to be put forward and justified explicitly because, as I will show in a later section, one can equally argue that genuine tempo effects are intrinsic to the fertility of a period.

Cohort Estimator (B2)

A second way of seeing distortion in the TFR as measurement bias is in Norman Ryder's sense: that period parameters are "distorted reflections of cohort behavior" and that the time series of period and cohort values do not coincide when fertility timing is changing (Ryder 1964:79, 1980). This implicit construal of bias/distortion—viewing period measures as erroneous indicators of real cohort values—is fairly common in the literature on tempo adjustment (Bongaarts 2008; Kohler and Ortega 2002a; Schoen 2004; Smallwood 2002b; Sobotka 2003; van Imhoff 2001; van Imhoff and Keilman 2000).

The conventional period TFR measures (real) cohort fertility accurately only if age-specific rates are either fixed or randomly distributed around a given period's values. Such stability is rare. The TFR and analogous synthetic indicators are therefore unquestionably biased as estimates of real cohort parameters. This type of bias is exacerbated by timing change. Furthermore, it is present in empirical populations even when timing is constant. Thus, the period TFR is a distorted version of associated cohort values for two reasons: because fertility timing changes and because rates are not fixed.

Theoretical (B3)

A third interpretation is within a theoretical framework. In a theoretical population with fixed age-specific rates, the period TFR is an unbiased measure of cohort completed fertility, but it becomes biased as a cohort indicator when a theoretical tempo change is introduced. As just noted, the TFR in empirical populations is a biased indicator of cohort total fertility regardless of whether timing is changing. In most discussion of tempo adjustment, tempo change is treated as the exclusive cause of bias in period indices, a view that makes sense only if it refers to the relationship between the period and cohort TFR in theoretical rather than in empirical populations.

Objectives of Fertility Measurement

Hand (2004:267) observed that "(t)he use to which an index will be put will be the determining factor in its construction." Hence, the attributes desirable in an indicator flow from its intended purpose. Although fundamental, the principle has been absent from recent discussion of fertility measurement. In this section, I aim to address that omission, by focusing on the leading objectives of measuring period fertility, considering whether, in each case, timing effects distort period measures, and, if so,

whether adjustment solves the problem. Five main objectives can be identified: to explain time-trends in fertility, to anticipate future trends, to describe fertility, to provide input parameters for formal population models, and to convey information on fertility trends to nonspecialist audiences.

Explaining Fertility Trends

We can set about explaining fertility time-trends in two ways. The strongest form of explanation is a theory entailing a substantive model—that is, a model that represents as far as possible the real-world processes generating the fertility rates analyzed (Cox 1990; Freedman 1985; Hand 2004). An alternative form of explanation is via empirical models, a more common strategy in social science; these models aim to explain as much as possible of the variance in a dependent variable, but they are not designed to represent the *modus operandi* of the underlying phenomena.² The arguments that follow apply broadly to both approaches, but some distinctions are drawn in discussion.

The question of whether timing effects are distorting when period fertility is the explanandum is addressed here in two stages, considering first the TFR *per se* as a measure, and then the arguments for and against adjusting the TFR.

Period Synthetic Indicators as Dependent Variables

Adjustment for bias/distortion is proposed primarily for the period TFR and other synthetic indicators. A first question, then, is whether indicators of this kind are suitable as dependent variables when explaining period trends.³ I think not, for two reasons. First, the TFR is a single figure indicator, but period change is multidimensional rather than unidimensional and thus cannot be represented by a single figure. A summary indicator, however, may be adequate for analyses of the long run that aim to explain gross changes in level. Second, the TFR and analogous indicators are in synthetic cohort form, referring to the cumulative experience of a lifetime, and are therefore in an inappropriate metric to represent the phenomena of a single period (Ní Bhrolcháin 1992). As a result, they are unsuitable as dependent variables in any substantive model of the underlying process in its period aspect. As noted earlier, synthetic indicators can be treated as simple statistical summaries, and in that guise might be suitable as dependent variables for empirical as distinct from behavioral models. However, the problem remains that no single measure can accurately represent fertility in time series when trends differ between exposure categories.

² A behavioral model is designed to represent underlying mechanisms and processes and could, in principle, be quite parsimonious and less complex than an empirical model. For an example, see Hernes's (1972) model of entry into first marriage.

³ When the objective is to explain change in the period TFR, be it discursively or by means of a statistical or mathematical model, the TFR is the dependent variable. By contrast, when the period TFR is regarded as an indicator of (future) completed cohort fertility or as a predictor of future fertility trends, it functions as an independent variable.

Are Timing Effects a Source of Bias in Period Fertility as Dependent Variable?

The arguments in the preceding section apply to the TFR in all its forms: additive, multiplicative, adjusted, or unadjusted. Nevertheless, let us set aside objections to TFR-like indices and consider the issue of bias/distortion in the TFR when period fertility is the explanandum. Of the types of bias identified earlier, two are potentially relevant to trends in the period TFR as dependent variable: confounding (bias A) and definitional (bias B1). Bias of types B2 and B3 cannot be present because when we are trying to explain trends in period fertility, the dependent variable measures a period phenomenon and is not a cohort estimator.

Year-on-year movements in the TFR can be biased by compositional factors because of confounding between the exposure distribution and calendar time—a spurious timing effect (bias A). The solution here is not adjustment, which is designed to remove genuine timing effects, but to use indices either specific for parity and age, or for parity and duration for orders two and higher, or standardized for these factors. Such indicators, and measures summarizing them—period parity progression ratios, the period parity progression based TFR, and regression-standardized parity-specific rates (Feeney and Yu 1987; Hoem 1993; Ní Bhrolcháin 1987; Rallu and Toulemon 1994)—are free of confounding bias. But, occurrence-exposure rates of this type, and measures derived from them, are influenced by genuine tempo change and thus might be thought of as definitionally distorted (bias B1). That is a mistake. True tempo effects do not distort period indicators in the role of explanandum. Rather, genuine tempo effects are an integral component of the period fertility trends we have to explain, as are the changes in variance highlighted by Kohler and Philipov (2001). Explanation means explaining the whole of the change in period fertility, not just part of it. To remove tempo effects from period fertility as explanandum would denude it of an intrinsic and often substantial component of change (Ní Bhrolcháin 1992, 1994). The greater the timing shift in a period, the worse the impact of adjustment because a larger part of what is happening in a period, from an explanatory angle, is removed. Wachter (2005) sees tempo adjustment as a way of standardizing for tempo. I argue that a period fertility indicator that is standardized for tempo is misspecified, if the objective is to explain period trends. The issue is not at all comparable to age standardization in mortality analysis. In explaining variation in death rates, age structure effects are regarded as extraneous and not a target of explanation. In explaining change in period fertility, by contrast, identifying the causes of tempo change *is* part of the objective.

An extreme example, the Year of the Fire Horse in Japan, illustrates the argument. Japan saw a dip of 27% in the TFR in 1966, due to avoidance of births in a year that, according to popular astrology, would be an inauspicious year of birth for a girl. To explain the phenomenon, it would be absurd to use as the dependent variable a time series of tempo-adjusted TFRs. Any timing element is integral to the impact of folk belief on that year's fertility. (Of course, if estimating the effect of the folk belief on cohort fertility, cohort outcomes would be the dependent variable.) In the case of the speed premium in Sweden, tempo is a large part of the period effect, and its precise detail allows a strong case to be made for a causal influence of legislation relating to maternity-pay provision on fertility (Andersson 1999; Hoem 1990; Ní Bhrolcháin and Dyson 2007). The same applies to less-pronounced period fluctuations. For

example, accelerated childbearing was a sizeable ingredient of the baby boom of the late 1950s and 1960s in the United States, the United Kingdom, and several other developed societies (Butz and Ward 1979; Ní Bhrolcháin 1987; Ryder 1980). If part of the explanation of the baby boom is that postwar prosperity, full employment, and high wages gave rise to accelerated marriage and childbearing, that faster pace of family formation must be represented on the left hand side of the equation. Similarly, later childbearing is a central feature of what is to be explained in the period trends of the last few decades in developed societies.

An analogy may be useful. Consider a car traveling for a fixed duration of time. Its speed varies during the journey: rounding a sharp bend or going uphill, it slows down, while on the straight or downhill, it travels faster. Speed may vary also depending on terrain, traffic, the driver's inclinations, and so on. The task of explaining period fertility trends is analogous to accounting for the sequence of speeds at which the car travels. Saying that a well-standardized period fertility indicator is distorted is like saying that a measure of the car's speed at an arbitrarily chosen point in the journey, or when the car is changing speed, is mistaken. It may well give a biased estimate of average speed over the journey as a whole, but it gives an accurate account of the car's speed at the point at which this was measured. If we think in terms of "underlying" speed or average speed during a journey, and whether and how it can be inferred from speed along the route, we are measuring something other than speed at a particular time-point. In addition, we would have to either construct models and make assumptions for the purpose, or investigate the properties of a large number of such journeys, to generate an empirical basis for the estimate.

The analogy is not perfect, but it helps to spotlight some key points. To adjust the measure of speed during periods of acceleration and deceleration would clearly misrepresent the recorded time sequence of speeds during the journey. Quantifying and explaining that sequence—comparable to measuring and explaining period fertility trends—is a different problem from measuring and explaining average speed or distance traveled during the journey—a task analogous to estimating cohort or longer-run fertility levels. Schoen (2004) has used the car analogy for a different purpose—to argue for the importance of cohort fertility—and assumes that the driver has an intended destination, although one that may alter during the journey. Here, the analogy is between the car's trajectory and aggregate fertility movements, and no assumption is needed about intentions regarding destination, speed, or duration of the journey.

Period Quantum and Tempo

The argument thus far is that genuine timing effects are intrinsic to period fertility as dependent variable, but that does not rule out the disaggregation of period fertility into tempo and quantum components, each to be the focus of explanation. Separate estimates of period level and timing effects for explanatory purposes could be argued for if several conditions were to hold. These conditions are that in period mode: (1) the quantum of fertility and its timing are separable in a quantitative sense; (2) quantum and tempo measures reflect distinct aspects of the underlying behavioral process (when we are seeking a behavioral explanation rather than an empirical one); and (3) that quantum and tempo respond differently to change in social, economic, and other determining factors.

Are tempo and quantum separable in a quantitative sense? Indices can be specified that, on their face, represent the quantum and tempo aspects of period fertility (e.g., Bongaarts and Feeney 1998; Butz and Ward 1979; Foster 1990; Kohler and Ortega 2002a; Ryder 1980). However, although separable in theory, quantum and tempo tend to covary in practice. The most familiar evidence of this is the near-universal tendency of cohort fertility series to reflect the fluctuations in corresponding period series, but with a lesser amplitude.⁴ Time series presented by Schoen (2004) illustrate the point further. Schoen's Table 1 gives a number of quantum indicators for the United States during 1917–2001: the conventional TFR, two versions of the Bongaarts and Feeney adjusted TFR, the Butz and Ward average completed fertility (ACF) measure,⁵ together with the mean age at childbearing and corresponding (true) cohort total fertility for part of the series. From these, the period timing indices associated with each quantum measure have been obtained. Corresponding to the ACF, we have the timing measure TFR / ACF (Butz and Ward 1979; Schoen 2004). Two specifications of a Bongaarts and Feeney timing index are calculated for each of the two versions of the BF TFR: additive (BF timing = $TFR - TFR^*$) and ratio versions (BF timing = TFR / TFR^*).

In all cases, the timing and quantum indicators are positively associated. The correlation between ACF and its associated timing index is .83 (1917–1997); and for the four versions of the Bongaarts and Feeney timing and quantum indices, the correlations are between .32 and .45 (1918–1997). On the ACF evidence, periods of faster timing are also periods of higher levels of fertility. This is true also of the Bongaarts and Feeney indices, although to a lesser extent, possibly because the TFR_{adj} indicator can be erratic (Schoen 2004; Smallwood 2002b). Overall, these figures demonstrate that although separable to some degree, period quantum and tempo are not independent (see also van Imhoff and Keilman 2000).

Are period quantum and tempo behaviorally distinct? Estimates of tempo and quantum effects such as those of Butz and Ward (1979), Ryder (1980), Foster (1990), and Bongaarts and Feeney (1998) are interesting. But, thus far, there have been no studies that investigate whether the mathematical constructs in these analyses correspond to real-world processes. The question is relevant at both micro and macro levels. At the individual level, the question is whether decisions about tempo and quantum are taken separately or jointly. At the aggregate level, whether the tempo/quantum divide corresponds to underlying processes is harder to determine, but it can be studied indirectly by investigating both the independence of tempo and quantum, and how far they are jointly or separately influenced by causal factors.

Ryder's perspective on the subject is instructive. In a paper that is best known for having estimated the quantum and tempo of cohort fertility and for having analyzed them into their components, Ryder (1980) concluded that the two are not independent. In his view, quantum and tempo "are to some degree manifestations of the same underlying behavior," and "we cannot, in principle, make a statistical

⁴ That tempo and quantum covary has long been recognized. The covariation has been referred to by some recent authors as "tempo-quantum interaction."

⁵ Butz and Ward's ACF, Schoen's preferred measure of period quantum, is the period TFR inflated/deflated by an index representing period tempo.

separation of the tempo and quantum facets of fertility” (Ryder 1980:44–45). For Ryder—the pioneer in estimating tempo and quantum—the two were interlinked at both the individual and at the aggregate levels.

Are period quantum and tempo influenced either by different factors or differentially by the same factors? If yes, then they may reflect genuinely distinct processes. If not, they are a single, undifferentiated entity. While instances can be found of changes in timing in reaction to socioeconomic determinants—the Swedish speed premium effect being a very clear-cut case (Andersson 1999; Andersson et al. 2006)—it is not obvious that such instances are exclusively due to timing effects nor that currently available indices of timing would represent them accurately. Research is still needed to address the real-world justification for separating level and timing, at both individual and aggregate levels.

Validation

Ultimately, measures used for any purpose need formal validation. No independent criterion is available against which to validate period measures in explanatory studies, but we do have an indirect check on the performance of a period measure as dependent variable: namely, explanatory success. Indicators of period fertility as explanandum would be validated by an empirically successful explanation of period trends in which they were embedded—a form of construct validity. As Ryder has suggested, we will know we have the right measures when we have a good explanation of time trends. Thus far, we lack convincing, well-documented explanations for period trends that could help to adjudicate between different measurement approaches. Nevertheless, success in explanation, rather than a check against cohort values, is the appropriate criterion for evaluating indicators of period fertility as dependent variable.

Anticipating the Future

A second major objective for measuring period fertility is to anticipate future population parameters. This is often the explicit or implicit rationale for studying recent trends and is entirely natural in the discipline. Period-based fertility measures are used in several ways to get a handle on population prospects: to estimate the fertility of cohorts, to look for indications of future trends in a less specific sense, and to encapsulate population growth prospects. Each of these objectives is considered here in turn.

Cohort Fertility

When data are available on completed childbearing, measuring cohort fertility is straightforward, subject only to the limits of data quality and sample size. However, when cohort childbearing is incomplete, period rates are used in the estimation process, and that presents some challenges. In a paper that has influenced this article, van Imhoff (2001:24–25) remarked: “A particularly important struggle faced by demographic analysts is how to arrive at statements about family formation processes from a cohort perspective . . . from data that are collected on an annual

basis. . . .” Synthetic cohort indicators are the principal means by which period fertility is converted into cohort terms. Demographic translation is another approach but will not be discussed here (Keilman 1994; Ryder 1964).

The divergence between time series of period and cohort indicators is one of the best known facts of demography. Nevertheless, the discipline appears prone to a curious doublethink that effectively treats synthetic indices such as the period TFR as cohort estimators. Evidence of the conflation of the two is found in the long-standing criticism that period indices constructed additively occasionally produce results that are impossible in a real cohort (see, e.g., Bongaarts 2002; Bongaarts and Feeney 1998; Ryder 1990; Sobotka 2003; van Imhoff and Keilman 2000). The causes of such apparent anomalies have been known for decades: variable age-specific rates and tempo change. However, that the feature is still mentioned as a defect reveals that the TFR continues to be thought of in some sense as a cohort estimator.

The period TFR, then, has long been known to be an unreliable indicator of the mean family size of associated cohorts. The discrepancy is often illustrated graphically by the much larger swings in period than in cohort total fertility. Considered as a cohort estimator, the TFR suffers from measurement bias (type B2), owing to both the variability of age-specific rates through time and tempo change. Genuine timing effects *are* a source of bias when classic period indices, such as the TFR, are used to estimate corresponding cohort parameters. Hence, methods to remove tempo effects are perfectly justifiable, unlike the case in which period measures are the focus of explanation. A successful method of eliminating timing effects would yield period indicators that better approximate cohort quantities. Does tempo adjustment achieve this in practice—that is, how well do tempo-adjusted indicators perform as cohort estimators?

Validation

Evidence is mixed on the validity of tempo-adjusted measures as estimators of completed cohort parameters. There are scattered data—proportions having a first birth in the United States or proportions marrying in France—suggesting that adjusted period values can be closer to cohort figures than are the period equivalent, although limited evidence is available thus far (Bongaarts and Feeney 2006: Figures 8–13; Sobotka 2003; Winkler-Dworak and Engelhardt 2004). But time series of annual adjusted TFRs do somewhat less well, not showing a reliable improvement over the conventional TFR in approximating to the total fertility of associated cohorts (Schoen 2004; Smallwood 2002b; Sobotka 2003, 2004a; van Imhoff and Keilman 2000). Thus, evidence validating the effectiveness of tempo adjustment in reducing measurement bias in the period TFR as an indicator of cohort fertility is at best equivocal. Adjustment, in any case, cannot be a complete solution since it cannot correct the bias resulting from the nonconstancy of age-specific rates.

Note that validation against real cohort outcomes is not in keeping with the original rationale advanced by Bongaarts and Feeney (1998) for tempo adjustment (although they also comment that in adjusting for tempo they have “extrapolated the experience of a single year into the future and ascertained the implied cohort fertility” [p. 289]). Nevertheless, such a validity check accords with the thinking of

other demographic scholars, for whom the essential purpose of tempo adjustment is to better approximate cohort values (Kohler and Ortega 2002a; Morgan et al. 2009; Schoen 2004; Smallwood 2002b; Sobotka 2003; van Imhoff 2001; van Imhoff and Keilman 2000). Also, in more recent work, Bongaarts and Feeney (2006:145) come closer to this standpoint, seeing the adjusted period measures as “approximate measures of lagged cohorts” when “patterns of change . . . are close . . . to the translation assumptions.”

Future Fertility

Period fertility indicators may be used to appraise fertility prospects more generally. A broad reference to the future is found in the discourse of fertility measurement in various ways. It appears to be what van Imhoff (2001:24) had in mind when he said that by “level of fertility,” we mean something like “how many children do people have, on average.” It also seems essentially what is meant by widely used references to “true” or “underlying” fertility, or the completed fertility “implied by” current rates (at least according to one reading—another will be considered in a later section).

On one interpretation, tempo-adjusted measures are a guide to longer-run fertility in some nonspecific sense, although not cohort fertility. This is how tempo adjustment has been viewed in practice by some demographic analysts—as carrying implications for long-run trends in fertility (Bongaarts 2002; Frejka and Ross 2001; Kohler and Ortega 2002a; Lesthaeghe and Willems 1999; Morgan and Berkowitz King 2001). However, evidence validating the TFR_{adj} as a predictor of longer-run fertility is as yet rather weak, as in the cohort case, and doubts have been expressed as to whether fertility is likely to reach the levels implied by the tempo-adjusted TFR (Frejka and Calot 2001; Frejka and Ross 2001; Lesthaeghe and Willems 1999; Sobotka 2004a).

Population Growth Prospects

Another role for period fertility measures is to express population growth prospects, reflected in the convention of regarding a TFR of 2.1 as replacement fertility. The replacement-level TFR idea assumes stable conditions just like its predecessors, the net and gross reproduction rates (NRR and GRR). It has been known since the 1940s that a single period’s TFR is no indication of future growth prospects. And despite the severe criticisms to which reproduction rates have been subject (Dorn 1950; Hajnal 1947, 1959; Stolnitz and Ryder 1949; Whelpton 1946), we continue to use the TFR as a kind of reproduction rate.

It is well known that replacement level fertility is empirically invalid as a concept because the TFR can be below replacement for decades without resulting in population decline (Smallwood and Chamberlain 2005). The reason is, of course, that the long-run, stable assumptions are almost never valid. Hajnal (1959:29) captured the essence of the matter by questioning the sense of “measuring the reproductivity of an actual population that is not stable.” To suggest adjusting the TFR to improve the current estimate of replacement prospects is taking the period estimate of replacement fertility too seriously. Besides, such a proposal goes

beyond the evidence. There is no evidence that the TFR_{adj} improves on the classic TFR as an indicator of long-run population replacement. Validation of the adjusted TFR as an indicator of replacement is lacking hitherto.

Alternatives to Tempo Adjustment

When period measures are used to estimate cohort or other future outcomes, there is no argument in principle against tempo adjustment. What matters is the empirical fit between an adjusted indicator and its target. Tempo-adjusted period measures have had limited success in approximating cohort quantities better than unadjusted versions, and the record in predicting future fertility movements is patchy, also. No evidence is at hand on the performance of the TFR_{adj} as an indicator of replacement.

How does adjustment fare against alternatives? Population science has some well-tried procedures for explicitly gauging likely future trends. One approach is to assess possible futures from detailed analysis of the recent past. Such close analysis has prompted several scholars to question the realism of scenarios implied by adjusted measures (Frejka and Calot 2001; Lesthaeghe and Willems 1999; Sobotka 2004b; van Imhoff 2001). The formal counterpart of this approach—a full population projection—is a further, time-honored alternative to tempo adjustment. Forecasting has several hard-to-beat advantages as a guide either to completed cohort fertility or to future fertility levels. The estimates are presented as projections rather than as measures, their inherent uncertainty is acknowledged, and assumptions about future movements in component rates are made explicit. In anticipating the future, a forecasting methodology will always trump tempo adjustment because it is less constrained, is more flexible, and can incorporate a range of possible forms of timing change, not just the Bongaarts-Feeney variety. Forecasting assumptions can be made selectively and with judgment. They are limited in applied contexts only by their perceived realism. Importantly, forecasts of cohort fertility can use the cumulated fertility of incomplete cohorts at the base period. By contrast, tempo adjustment is a one-trick pony—a rigid application of the same transformation of period rates, year by year—and dispenses with the entire past of the fertility process but for each pair of adjacent periods. To estimate the ultimate mean family size of incomplete cohorts or future period fertility, an explicit forecast is a more transparent, versatile, and powerful vehicle (cf., Kohler and Ortega 2002a, b; Lesthaeghe and Willems 1999; Schoen 2004; van Imhoff 2001). In particular, replacement prospects are gauged more convincingly by a realistic forecast than by any single period's TFR, adjusted or not.

Describing Fertility Trends and Differentials

Descriptive work is ubiquitous in demography, in reports on levels, trends, and differentials. Whether the TFR is biased by timing effects in this context depends on the objective. If description is undertaken as the background either to explaining time trends or to anticipating the future, the arguments of the preceding discussion apply. Description, however, may have less specific aims. The TFR is often used to convey gross differences in fertility level between countries, for example. Such comparisons suffer from confounding (bias A), which can be removed by using

indicators specific for parity and age or duration. Nevertheless, when contrasts are large—between, say, TFRs of 6 versus 4, or 3 versus 2—comparisons are unquestionably informative. When differentials are smaller, timing effects may be responsible for a larger part of cross-national differences. Whether tempo adjustment can illuminate such cases depends on solid evidence validating it for *some* purpose. Meanwhile, concrete indicators of known significance seem more fruitful. For example, the standardized mean age at first birth is a more informative adjunct to the period TFR than TFR_{adj} , which is opaque as to the source of timing change. Lutz and Skirbekk's (2005) analysis of tempo-related policies illustrate this point: their proposed interventions target not an abstract “tempo effect” but the mean age at first birth and, through it, the mean age at childbearing.

Formal Models

Period fertility indices are also deployed in a theoretical way. This is not measurement at all, but a form of population modeling. The TFR as classically conceived—the mean family size of a hypothetical cohort subject to the rates of a particular period—is a theoretical parameter, not an empirical measure. A synthetic cohort has no empirical counterpart, and its hypothetical mean family size measures no real-world entity. Routine use of these indicators may have engendered a belief in the existence of real counterparts to the measures employed, what Wilson and Oeppen (2003) termed “reification.” Synthetic indicators are quite unlike apparently analogous measurements such as speed. Miles per hour, although stated in hypothetical terms (distance that would be traveled in an hour at a given speed), can be converted into distance traveled during a very short interval at the time of measurement. That quantifies something real. Nothing similar about real-world births or birth rates can be derived from the period TFR, in its classic interpretation.

Conceptual approaches interpreting the TFR_{adj} and allied indicators as theoretical, rather than empirical, entities include those of Zeng and Land (2001) and of Rodriguez (2006), who see the Bongaarts and Feeney TFR_{adj} as the cohort mean family size in a theoretical population with age-order specific rates of the period and subject continuously to the tempo change of the period. A theoretical construal is implicit rather than explicit in accounts attributing the presence of distortion in the TFR exclusively to tempo change (Bongaarts and Feeney 1998, 2006; Kohler and Ortega 2002a; Kohler and Philipov 2001). It is only in theoretical populations that bias in the TFR as a measure of cohort fertility is confined to periods when tempo is changing. The model underlying the TFR_{adj} is one among a large variety of potential models of aggregate change in fertility. Tempo adjustment in the mortality arena has, following Vaupel (2002), come to be regarded in this way, and the correspondence is made explicit between specific period life expectancy indicators and the processes generating real-world mortality change (Barbi et al. 2008; Vaupel 2009). Recognition of the essentially theoretical nature of Bongaarts and Feeney's tempo-adjustment ideas could open up a fruitful new area in population mathematics: the development of a range of theoretical models of temporal change in fertility. The utility of such models would depend on how well they reflected the empirical processes giving rise to fertility change.

In a theoretical context, tempo adjustment is unexceptionable. Procedures to adjust for the tempo-associated distortion can be designed around the particular type of tempo shift assumed to operate. They are uncontroversial when estimating the cohort fertility of a theoretical population experiencing the conditions specified. However, proposing a theoretical indicator as a measure of an empirical phenomenon requires, as a prerequisite, evidence that the model conditions under which the adjusted indicator is meaningful hold empirically.

Communication and Public Information

A final reason for choosing an index of period fertility is to convey information on fertility trends to nonspecialist audiences. Such communication needs may influence scientific practice, imperceptibly, because of the strong ties between academic demography and official statistical agencies.

A tempo-adjusted TFR appears unnecessary in this more popular context. Although crude, the TFR, if contextualized, can satisfy most nontechnical users' needs, at least for policy purposes. When tempo change is prominent, the classic TFR can be accompanied by information on timing, particularly on age at first birth. It may be suggested that unadjusted period synthetic indicators mislead the uninitiated about their likely lifetime experience. But classic or adjusted, period synthetic measures need not and should not be presented as implicit forecasts. To convey future likely experience, statistical offices should rely on explicit forecasts and not on period synthetic indices.

William Brass believed that the preference for single-figure summary indicators, especially the TFR, arose from extra-scientific pressures to simplify communication, thus displacing technically superior approaches using multiple indicators (Brass 1990: 455). The TFR's apparent intelligibility to nonspecialists is irrelevant, though, when choosing measures for technical purposes. Physicists and astronomers do not justify their measurement procedures by whether lay people can understand them. Nor should we. Besides, a projected completed family size is just as user-friendly as a classic or tempo-adjusted version.

Discussion

Bongaarts and Feeney's ideas on tempo adjustment have been and continue to be influential. I have considered the arguments in principle for tempo adjustment first by clarifying the statistical meaning of the terms *bias* or *distortion* as applied to the period TFR. I identified two distinct types of bias implicit in recent discussion: confounding and measurement bias. I then distinguished the main purposes for which period fertility is measured; asked whether timing effects are biasing in each case; and, if so, considered what kind of bias/distortion is in question and whether adjustment is an effective way to deal with it. Table 1 summarizes my conclusions.

I have argued that genuine timing effects are not a source of distortion but integral to period fertility as dependent variable, and so adjustment is neither necessary nor appropriate when explaining period change. In nonspecialist communication, tempo effects can be conveyed well by existing methods. In a descriptive context,

Table 1 Summary of conclusions on tempo adjustment of period fertility measures and alternative solutions

Purpose	Are Genuine Timing Effects a Source of Bias in Period Measures?	Is There a Case for Tempo Adjustment?	Appropriate Method of Validation	Has Tempo Adjustment Been Validated?	Alternative Solutions
Explanation	No	No, but separation into period quantum and tempo not ruled out.	Development of a successful explanation of trends to which specified measure(s) integral.	No	Ideally, measures specific for parity and age or duration, or indices standardized for these, but depends on time scale.
Anticipation Estimate cohort values	Yes, probably worsen fit between period and cohort or longer-term indicators.	Depends entirely on empirical success.	Compare estimates with true (empirical) cohort parameters.	Limited. Results mixed thus far.	Project/forecast cohort fertility explicitly.
Predict longer run fertility					
Evaluate growth prospects	Yes, stable assumptions even less valid under changing tempo.		Evaluate measures as empirical predictors of longer run trends.	No	Detailed analysis of time series of recent trends. Project/forecast explicitly.
Description	Depends on purpose	Depends on objective of investigator.	Depends on purpose.	No	Mathematics of population reproductivity needs to be developed further. Depends on ultimate purpose. TFR together with standardized mean age at first birth highly informative.
Population Modeling	Yes under classic assumptions, but alternative models specifiable.	Adjustment neither required nor ruled out by theory.	Only mathematical consistency required. ^a	Only mathematical consistency required.	Adjustment model one of a large number of possible models. Investigation of further formal models of temporal change would be fruitful.
Communication	Depends on presentation.	Inadvisable; little justification.	Not relevant.	Not relevant.	Present a greater range of indicators. Rely more on long-term projections of fertility.

^a However, if a theoretical indicator is applied to an empirical situation and treated as a measure, it has to be shown that the assumed conditions hold sufficiently generally for the use of the measure to be defensible.

adjustment will often be irrelevant. By contrast, genuine tempo effects *are* a source of bias when period synthetic measures are used as a proxy for the cohort equivalents or to predict longer-term fertility, or in some theoretical scenarios. Tempo adjustment is, therefore, a candidate technique when period rates are employed to estimate completed cohort fertility or to predict future fertility. Does it work well in that context? Thus far, the evidence is mixed. Tempo-adjusted indicators have been evaluated primarily against cohort parameters. Adjustment has not yet been shown to improve reliably on period measures as estimators or predictors, although there are instances in which it appears to do reasonably well. Alternatives to adjustment are available—notably disaggregated measures, particularly parity-specific rates—to replace the reliance on single-figure summary indicators. Estimates of the final fertility of incomplete cohorts and of future likely period fertility should be based routinely and explicitly on forecasts (Kohler and Ortega 2002a, b; Lesthaeghe and Willems 1999; Schoen 2004). We especially need to avoid the conflation of measurement and forecasting embodied in the synthetic cohort technique. Close analysis of recent trends of disaggregated indicators is clearly a prerequisite both for explanatory purposes and for assessing future prospects.

Bongaarts and Feeney originally envisioned the TFR_{adj} as giving a better reading of current fertility than the classic TFR, but they were not explicit about the intended use of TFR_{adj} : whether it is proposed as measuring period fertility as explanandum, predictor, description, theoretical entity, or vehicle of communication. To evaluate whether one measure of period fertility gives a better reading than another, we need to know the objective of making the measurement, so as to identify an empirical criterion of how well the proposed measure performs. Paucity of empirical validation is a central weakness in the current case for tempo adjustment.

Much of the conceptual difficulty associated with tempo adjustment stems from reliance on a single-figure synthetic period indicator. This limiting choice casts a shadow on conventional methodology. Its historical origins are instructive. The NRR and GRR were routinely used as indices of time trends up to the 1940s, but rapid shifts in fertility in the 1930s and 1940s brought the realization that the stable assumptions underwriting their quantitative relevance did not hold empirically. The heterogeneity of fertility series with respect to both parity and personal time was also recognized. A variety of alternatives to the NRR and GRR were proposed—measures specific for parity- and age or duration, cohort analysis, and period parity progression ratios (Hajnal 1947, 1959; Henry 1980; Stolnitz and Ryder 1949; Whelpton 1946, 1949, 1954). But these methods were sidelined by the ascent of the TFR, a reproduction rate in all but name, its leading role being in Brass's view attributable to “[s]implicity, convenience and propaganda” (Brass 1990:456). It is time for the discipline to cast off the conceptual-methodological straightjacket imposed by a historical legacy that overvalues bureaucratic and short-term convenience at the expense of scientific requirements.

Parity-specific measures that retain genuine tempo effects—for example those of Feeney and Yu (1987), Ní Bhrolcháin (1987), and Rallu and Toulemon (1994) as well as the regression-standardized indicators of Hoem (1993)—remain the best current alternative to the TFR when the objective is to explain period fertility. And while, on the present argument, period parity progression ratios, being synthetic, are

problematic in the role of explananda, they may be useful as standardized summaries. That parity-specific measures are demanding of data has doubtless contributed to their relative neglect. But current resources could be better exploited: indirect methods are available (Feeney 1991; Henry 1980), the “own children” method has been shown to produce reasonable estimates (Luther et al. 1990; Murphy and Berrington 1993), and age- and parity-specific rates can be reconstructed in the absence of comprehensive micro data (Chamberlain and Smallwood 2004; Heuser 1976; Smallwood 2002a). Parity-specific rates also offer a promising but underutilized approach to forecasting (Feeney 1985; Kohler and Ortega 2002a, b; Sobotka 2004b; Toulemon and Mazuy 2001).

Demographic measurement appears to inhabit a methodological ghetto, sealed off from the broader statistical world of measurement and validation. The conventional reliance on period measures rooted in stable population assumptions has outlived its usefulness. A less mechanical and more empirical approach to demographic measurement is required that is, above all, statistically defensible. The record of investigations conducted along these lines is decidedly positive. Demonstration of the speed-premium effect in Sweden by a set of regression standardized indicators had no need of stable assumptions (Andersson 1999; Hoem 1993). None of the three most informative decompositions of tempo and quantum—those of Butz and Ward (1979), Foster (1990), and Ryder (1980)—required stable population assumptions, nor did Lee’s (1980) moving target theory. These demographically aware mathematical and statistical analyses demonstrate the potential payoff to a statistically less constrained, more open-minded, and more sophisticated approach to demographic measurement and analysis.

The essence of the period tempo issue is that we cannot distinguish short- or medium-term period timing shifts from longer-term quantum changes. A high priority for fertility research is, therefore, to find empirical methods of telling these apart, either contemporaneously or retrospectively. Analyzing the rates alone is unlikely to solve the problem. The solution almost certainly requires detailed knowledge of the substantive factors driving fertility trends in any particular case. The history of science has precedents. For example, in the development of a scale for temperature, the study of underlying processes—how various substances behaved across a range of temperatures—was essential to establishing a sound measuring instrument (Chang 2004). Similarly, technical and substantive features of fertility change are interwoven. Future progress will hinge on integrating research on measurement with investigation of the causes of fertility change at the aggregate level.

Acknowledgments I thank David Freedman, Mike Murphy, Laurent Toulemon, Chris Wilson and several anonymous referees for helpful comments on earlier versions of this article. Any errors are mine.

References

- Andersson, G. (1999). Childbearing trends in Sweden 1961–1997. *European Journal of Population*, 15, 1–24.
- Andersson, G., Hoem, J. M., & Duvander, A.-Z. (2006). Social differentials in speed-premium effects in childbearing in Sweden. *Demographic Research*, 14(article 4), 51–70. doi:10.4054/DemRes.2006.14.4

- Barbi, E., Bongaarts, J., & Vaupel, J. W. (Eds.). (2008). *How long do we live? Demographic models and reflections on tempo effects*. Berlin and Rostock, Germany: Springer and Max Planck Society.
- Billari, F. C., Liefbroer, A. C., & Philipov, D. (2006). *The postponement of childbearing in Europe: Driving forces and implications*. *Vienna Yearbook of Population Research*, 4, 1–17.
- Bongaarts, J. (1999). The fertility impact of changes in the timing of childbearing in the developing world. *Population Studies*, 53, 277–289.
- Bongaarts, J. (2002). The end of the fertility transition in the developed world. *Population and Development Review*, 28, 419–443.
- Bongaarts, J. (2008). What can fertility indicators tell us about pronatalist policy options? *Vienna Yearbook of Population Research*, 6, 39–55.
- Bongaarts, J., & Feeney, G. (1998). On the quantum and tempo of fertility. *Population and Development Review*, 24, 271–291.
- Bongaarts, J., & Feeney, G. (2000). On the quantum and tempo of fertility: Reply. *Population and Development Review*, 26, 560–564.
- Bongaarts, J., & Feeney, G. (2006). The quantum and tempo of life-cycle events. *Vienna Yearbook of Population Research*, 4, 115–151.
- Brass, W. (1990). Cohort and time period measures of quantum fertility: Concepts and methodology. In H. A. Becker (Ed.), *Life histories and generations*, Vol. 2 (pp. 455–476). Utrecht, The Netherlands: ISOR.
- Butz, W. P., & Ward, M. P. (1979). Will US fertility remain low? A new economic interpretation. *Population and Development Review*, 5, 663–688.
- Chamberlain, J., & Smallwood, S. (2004). Estimates of true birth order for Scotland, 1945–1999. *Population Trends*, 117, 27–42.
- Chang, H. (2004). *Inventing temperature: Measurement and scientific progress*. Oxford, UK: Oxford University Press.
- Cox, D. (1990). Role of models in statistical analysis. *Statistical Science*, 5, 169–174.
- Dorn, H. (1950). Pitfalls in population forecasts and projections. *Journal of the American Statistical Association*, 45, 311–334.
- Ediev, D. (2008). *On the theory of distortions of period estimates of the quantum caused by the tempo changes (European Demographic Research Paper 3/2008)*. Vienna, Austria: Vienna Institute of Demography.
- Feeney, G. (1985). *Parity progression projection*. In *International population conference, Florence 1985*, Vol. 4 (pp. 125–136). Liège, Belgium: International Union for the Scientific Study of Population.
- Feeney, G. (1991). Fertility decline in Taiwan: A study using parity progression ratios. *Demography*, 28, 467–479.
- Feeney, G., & Yu, J. (1987). Period parity progression measures of fertility in China. *Population Studies*, 41, 77–102.
- Foster, A. (1990). Cohort analysis and demographic translation. *Population Studies*, 44, 287–315.
- Freedman, D. A. (1985). Statistics and the scientific method. In W. Mason & S. Fienberg (Eds.), *Cohort analysis in social research* (pp. 343–366). New York: Springer Verlag.
- Frejka, T., & Calot, G. (2001). Cohort reproductive patterns in low fertility countries. *Population and Development Review*, 27, 103–132.
- Frejka, T., & Ross, J. (2001). Paths to subreplacement fertility: The empirical evidence. *Population and Development Review*, 27(Supplement: Global Fertility Transition), 213–254.
- Frejka, T., Sobotka, T., Hoem, J. M., & Toulemon, L. (Eds.). (2008). Summary and general conclusions: Childbearing trends and policies in Europe. *Demographic Research Special Collection*, 7. doi:10.4054/DemRes.2008.19.2
- Gauthier, A., & Philipov, D. (2008). *Can policies enhance fertility in Europe? Vienna Yearbook of Population Research*, 6, 1–16.
- Goldstein, J., Sobotka, T., & Jasilioniene, A. (2009). The end of “lowest-low” fertility? *Population and Development Review*, 35, 663–699.
- Guillot, M. (2006). Tempo effects in mortality: An appraisal. *Demographic Research*, 14(article 1), 1–26. doi:10.4054/DemRes.2006.14.1
- Hajnal, J. (1947). The analysis of birth statistics in the light of the recent international recovery of the birth-rate. *Population Studies*, 1, 137–164.
- Hajnal, J. (1959). The study of fertility and reproduction: A survey of 30 years. In *Thirty years of research in human fertility* (pp. 11–37). New York: Milbank Memorial Fund.
- Hand, D. J. (2004). *Measurement. Theory and practice*. London, UK: Arnold.
- Henry, L. (1980). *Fertility of marriage: A new method of measurement (UN/ESCAP Population Studies Translation Series No. 3, Trans.)*. New York: United Nations (Original work published 1953).

- Hernes, G. (1972). The process of entry into first marriage. *American Sociological Review*, 37, 173–182.
- Heuser, R. 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.
- Hoem, J. (1990). Social policy and recent fertility change in Sweden. *Population and Development Review*, 16, 735–747.
- Hoem, J. (1993). Classical demographic methods of analysis and modern event-history analysis. *IUSSP International Population Conference Proceedings*, 3, 281–291.
- Inaba, H. (2007). Effects of age shifts on the tempo and quantum of non-repeatable events. *Mathematical Population Studies*, 14, 131–168.
- Keilman, N. (1994). Translation formulae for non-repeatable events. *Population Studies*, 48, 341–357.
- Kim, Y. J., & Schoen, R. (2000). On the quantum and tempo of fertility: Limits to the Bongaarts-Feeney adjustment. *Population and Development Review*, 26, 554–559.
- Kohler, H.-P., & Ortega, J. A. (2002a). Tempo-adjusted period parity progression measures, fertility postponement and completed cohort fertility. *Demographic Research*, 6(article 6), 91–144. doi:10.4054/DemRes.2002.6.6
- Kohler, H.-P., & Ortega, J. A. (2002b). Tempo-adjusted period parity progression measures: Assessing the implications of delayed childbearing for cohort fertility in Sweden, the Netherlands and Spain. *Demographic Research*, 6(article 7), 145–190. doi:10.4054/DemRes.2002.6.7
- Kohler, H.-P., & Philipov, D. (2001). Variance effects in the Bongaarts-Feeney formula. *Demography*, 38, 1–16.
- Kohler, H.-P., Billari, F. C., & Ortega, J. A. (2002). The emergence of lowest-low fertility in Europe during the 1990s. *Population and Development Review*, 28, 641–680.
- Lee, R. D. (1980). Aiming at a moving target: Period fertility and changing reproductive goals. *Population Studies*, 34, 205–226.
- Lesthaeghe, R., & Willems, P. (1999). Is low fertility a temporary phenomenon in the European Union? *Population and Development Review*, 25, 211–228.
- Luther, N. Y., Feeney, G., & Zhang, W. (1990). One-child families or a baby boom? Evidence from China's 1987 One-per-Hundred Survey. *Population Studies*, 44, 341–357.
- Lutz, W., & Skirbekk, V. (2005). Policies addressing the tempo effect in low-fertility countries. *Population and Development Review*, 31, 699–720.
- Morgan, S. P., & Berkowitz King, R. (2001). Why have children in the 21st century? Biological predisposition, social coercion, rational choice. *European Journal of Population*, 17, 3–20.
- Morgan, S. P., Zhigang, G., & Hayford, S. R. (2009). China's below-replacement fertility: Recent trends and future prospects. *Population and Development Review*, 35, 605–629.
- Murphy, M., & Berrington, A. (1993). Constructing period parity progression ratios from household survey data. In M. Ní Bhrolcháin (Ed.), *New perspectives on fertility in Britain* (pp. 17–32). London, UK: HMSO.
- Ní Bhrolcháin, M. (1987). Period parity progression ratios and birth intervals in England and Wales 1941–1971. *Population Studies*, 41, 103–125.
- Ní Bhrolcháin, M. (1992). Period paramount? A critique of the cohort approach to fertility. *Population and Development Review*, 18, 599–629.
- Ní Bhrolcháin, M. (1994). Past history, synthetic indicators and period fertility. *Population (an English Selection)*, 6, 116–120.
- Ní Bhrolcháin, M. (2008). Five reasons for measuring period fertility (S3RI Applications and Policy Working Papers A08/05). Southampton, UK: Southampton Statistical Sciences Research Institute. Retrieved from <http://eprints.soton.ac.uk/52411>
- Ní Bhrolcháin, M., & Dyson, T. (2007). On causation in demography: Issues and illustrations. *Population and Development Review*, 33, 1–36.
- Pressat, R., & Wilson, C. (Eds.). (1988). *The dictionary of demography*. Oxford, UK: Blackwell.
- Rallu, J.-L., & Toulemon, L. (1994). Period fertility measures: The construction of different indices and their application to France, 1946–89. *Population (an English Selection)*, 6, 59–94.
- Rodriguez, G. (2006). Demographic translation and tempo effects: An accelerated failure time perspective. *Demographic Research*, 14(article 6), 85–110. doi:10.4054/DemRes.2006.14.6
- Ryder, N. (1964). The process of demographic translation. *Demography*, 1, 74–82.
- Ryder, N. (1980). Components of temporal variations in American fertility. In R. W. Hiorns (Ed.), *Demographic patterns in developed societies* (pp. 15–54). London, UK: Taylor and Francis.
- Ryder, N. (1990). What is going to happen to American fertility? *Population and Development Review*, 16, 433–454.
- Schoen, R. (2004). Timing effects and the interpretation of period fertility. *Demography*, 41, 801–819.

- Smallwood, S. (2002a). New estimates of trends in births by birth order in England and Wales. *Population Trends*, 108, 32–48.
- Smallwood, S. (2002b). The effect of changes in timing of childbearing on measuring fertility in England and Wales. *Population Trends*, 109, 36–45.
- Smallwood, S., & Chamberlain, J. (2005). Replacement fertility, what has it been and what does it mean? *Population Trends*, 119, 16–27.
- Sobotka, T. (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(article 6), 151–214. doi:10.4054/DemRes.2003.8.6
- Sobotka, T. (2004a). Is lowest-low fertility in Europe explained by the postponement of childbearing? *Population and Development Review*, 30, 195–220.
- Sobotka, T. (2004b). *Postponement of childbearing and low fertility in Europe*. Amsterdam, The Netherlands: Dutch University Press.
- Stolnitz, G. J., & Ryder, N. B. (1949). Recent discussion of the net reproduction rate. *Population Index*, 15, 114–128.
- Toulemon, L., & Mazuy, M. (2001). Cinq projections de fécondité fondées sur une hypothèse de stabilité de comportements [Five projections of fertility based on a stable behavior hypothesis]. *Population*, 56, 647–656.
- van Imhoff, E. (2001). On the impossibility of inferring cohort fertility measures from period fertility measures. *Demographic Research*, 5(article 2), 23–64. doi:10.4054/DemRes.2001.5.2
- van Imhoff, E., & Keilman, N. (2000). On the quantum and tempo of fertility: Comment. *Population and Development Review*, 26, 549–553.
- Vaupel, J. W. (2002). Life expectancy at current rates vs. current conditions. *Demographic Research*, 7 (article 8), 365–378. doi:10.4054/DemRes.2002.7.8
- Vaupel, J. W. (2009). Lively questions for demographers about death at older ages. *Population and Development Review*, 35, 347–356.
- Wachter, K. W. (2005). Tempo and its tribulations. *Demographic Research*, 13(article 9), 201–222. doi:10.4054/DemRes.2005.13.9
- Whelpton, P. K. (1946). Reproduction rates adjusted for age, parity, fecundity, and marriage. *Journal of the American Statistical Association*, 41, 501–516.
- Whelpton, P. K. (1949). Cohort analysis of fertility. *American Sociological Review*, 14, 735–749.
- Whelpton, P. K. (1954). *Cohort fertility*. Princeton, NJ: Princeton University Press.
- Wilson, C., & Oeppen, J. (2003). On reification in demography. In J. Fleischhacker, H. A. de Gans, & T. K. Burch (Eds.), *Population, projections and politics* (pp. 113–129). Amsterdam, The Netherlands: Rozenberg.
- Winkler-Dworak, M., & Engelhardt, H. (2004). On the tempo and quantum of first marriages in Austria, Germany, and Switzerland: Changes in mean age and variance. *Demographic Research*, 10(article 9), 231–264. doi:10.4054/DemRes.2004.10.9
- Zeng, Y., & Land, K. (2001). A sensitivity analysis of the Bongaarts-Feeney method for adjusting bias in observed period total fertility rates. *Demography*, 38, 17–28.
- Zeng, Y., & Land, K. (2002). Adjusting period tempo changes with an extension of Ryder's basic translation equation. *Demography*, 39, 269–285.