WHY WE MEASURE PERIOD FERTILITY

MÁIRE NÍ BHROLCHÁIN

ABSTRACT

Four reasons for measuring period fertility are distinguished: to explain fertility time trends, to anticipate future fertility, to construct theoretical models and to communicate with non-specialist audiences. The paper argues that not all measures are suitable for each purpose, and that tempo adjustment may be appropriate for some objectives but not others. In particular, it is argued that genuine timing effects do not bias or distort measures of period fertility as dependent variable. Several different concepts of bias or distortion are identified in relation to period fertility measures. Synthetic cohort indicators are a source of confusion since they conflate measurement and forecasting. Anticipating future fertility is more akin to forecasting than to measurement. Greater clarity about concepts and measures in the fertility arena could be achieved by a stronger emphasis on validation. Period incidence and occurrence-exposure rates have a straightforward interpretation. More complex period fertility measures are meaningful only if a direct or indirect criterion can be specified against which to evaluate them. Their performance against that criterion is what establishes them as valid or useful.

Southampton Statistical Sciences Research Institute
Applications & Policy Working Paper A07/04

University of Southampton

Why we measure period fertility

Máire Ní Bhrolcháin
Division of Social Statistics
University of Southampton

.

October 2006

Revised June 2007

Acknowledgments: I thank David Freedman, Laurent Toulemon and Mike Murphy for helpful comments on a previous draft. Any errors that remain are mine.

.

Abstract

Four reasons for measuring period fertility are distinguished: to explain fertility time trends, to anticipate future fertility, to construct theoretical models and to communicate with non-specialist audiences. The paper argues that not all measures are suitable for each purpose, and that tempo adjustment may be appropriate for some objectives but not others. In particular, it is argued that genuine timing effects do not bias or distort measures of period fertility as dependent variable. Several different concepts of bias or distortion are identified in relation to period fertility measures. Synthetic cohort indicators are a source of confusion since they conflate measurement and forecasting. Anticipating future fertility is more akin to forecasting than to measurement. Greater clarity about concepts and measures in the fertility arena could be achieved by a stronger emphasis on validation. Period incidence and occurrence-exposure rates have a straightforward interpretation. More complex period fertility measures are meaningful only if a direct or indirect criterion can be specified against which to evaluate them. Their performance against that criterion is what establishes them as valid or useful.

Keywords

Fertility; period fertility; measurement; tempo adjustment; timing; validation; forecasting.

1. Introduction

Four main purposes for specifying period fertility measures can be identified: to explain fertility time-trends, to anticipate future fertility, to construct theoretical models, and to convey information on fertility trends to non-specialist audiences. The measures most suitable for each of these objectives, and the criteria for assessing them, differ. Fertility indices that are adjusted for period change in the timing of childbearing—tempo adjusted measures—may be appropriate for some purposes, but not others. No one fertility index or set of indices is best suited to all purposes. The unexpectedly low levels reached by fertility in developed countries in recent decades have provoked much discussion of fertility prospects (see e.g. Lesthaeghe and Willems 1999; Golini 1998; Bongaarts 2002; Lutz et al 2003; Morgan 2003). That debate has centred partly on timing effects and also on measurement, stimulated by the elegant and sophisticated adjustment to the total period fertility rate proposed by Bongaarts and Feeney (1998).

The issue of indicators has been a matter of debate—arguably because of a lack of clarity about the variety of reasons for measuring period fertility and about how fertility indices should be evaluated. A further source of difficulty is that the "fertility" to be measured is widely thought of as, in some sense, the average number of children women have, a formulation which in a period context gives rise to measures based on the synthetic cohort principle. Such indices are peculiarly unsuited to fertility in its period aspect. A final difficulty is that the ideas of quantum and tempo are thought to be straightforwardly applicable to period fertility phenomena, whereas they are, in fact, poorly defined in a period context. The recent literature on adjusting fertility measures for tempo effects has little to say about any of these difficulties. By and large it ignores the differing objectives of period analysis, the likely multiplicity of corresponding indices, the intellectual hazards of thinking in terms of synthetic cohort indicators, and the problematic nature of the period concepts of quantum and tempo.

The present paper addresses period fertility measurement in low fertility populations. The focus is on period fertility indicators because it is these that present the greatest difficulties in relation to measurement. Indices that represent fertility in consecutive calendar periods are less transparent in meaning, and more contentious, than measures of cohort fertility. Note, however, that the relative merits of a period vs. a cohort perspective on fertility are not at all central but incidental to the paper's concerns. The starting point is, rather, the uncontroversial fact that measures of period fertility are widely used in the demographic literature; period measures are used by scholars on each side of the period/cohort debate, as well as by those who are agnostic on the subject. The paper directs attention to a less well recognized issue—that period measures serve a variety of purposes, and that the attributes desirable in a period fertility index depend on the specific purpose they are intended to serve. A key objective of the paper is to examine how far tempo adjustment is appropriate, in principle, for each of the four objectives identified. A further aim is to emphasize the neglected issue of validation—period fertility indices must be evaluated against an external criterion in order to establish their utility. A discussion section summarizes and elaborates further on the arguments presented.

2. Objectives of fertility measurement

2.1 Explaining fertility trends

A natural starting point might appear to be the measurement of period fertility for purely descriptive purposes. But aside from published compilations of statistical series, and in textbook exposition, it is doubtful that period trends in fertility are often examined with purely descriptive intent. One or other of the purposes identified above are almost always implicit or explicit in work on period fertility series. Hence, we consider first period fertility measured for the purpose of explaining time-trends. To seek a well-specified measure of this kind is not to imply that explaining period trends takes precedence over explicating change in cohort fertility. It assumes only that the attempt to explain fertility trends in period mode is a valid and potentially useful research objective. The need for such measures is, thus, independent of whether we view period trends as essentially a reflection of cohort forces or vice versa, though the indicators chosen will often reflect assumptions about the relative role of period and cohort influences on fertility rates.

If an indicator is sought purely for prediction purposes, it does not matter what form it takes as long as it performs well as a predictor. But to understand the causes of change in period fertility, we need a form of measurement that reflects period fertility in its role as explanandum or dependent variable—*if* our objective is to construct an explanatory theory of real world processes rather than simply to "explain" variance empirically by means of a regression equation. The properties desirable in a measure of fertility as dependent variable may be quite different from the desiderata in period fertility indices that serve other functions. In particular, there is no *a priori* reason why a measure of period fertility as explanandum should be expected to function as a predictor of future or long-run trends. In a theoretical explanatory context, a measure of period fertility as dependent variable needs to meet both demographic-statistical and substantive criteria. Ideally, specification of fertility as dependent variable should flow from an explicit behavioral model or theory, and in any case choice of measure usually embodies some assumptions about underlying processes, whether acknowledged or not. Leridon (2006) notes, however, that there is no generally accepted and successful theory of fertility change, and that this view is shared by several other leading commentators in the field (Lesthaeghe and Surkyn 1988; van de Kaa 1996; Kirk 1996). Hence, we have as yet little by way of established theory to guide us in specifying suitable measures. Nevertheless, some practical guidelines can be suggested.

One criterion relevant to choice of measure is scale—detail in measurement will vary with time-scale and the level of generality sought. For broad brush treatment or long range perspectives such as those of e.g. Frejka and Ross (2001) or Caldwell and Schindlmayr (2003)—where gross change or differentials are the focus of interest—explanatory ideas will probably be general and systemic, and a total fertility rate (TFR) of some kind may well be adequate. Where the TFR works in this context, it is not because it represents “family size”, in any real sense, in successive calendar years but because it summarizes annual age-specific fertility rates, and so is a general indicator of level. The one-dimensional representation of year on year variation in level alone may well be adequate if the aim is to account for a shift from a total fertility rate in the region of, say, 5 to one of around 2. But for short to medium run series, in low fertility settings, a greater degree of resolution is to be expected both in fertility as the object of explanation and in the explanatory factors proposed. In this context, a total fertility indicator—whether conventional or adjusted in some way—is a coarse dependent variable, and more refined measurement is called for. The principal refinement is that rates or probabilities are used and that these are specific by parity, parity dependence being a key feature of fertility behavior in contracepting populations. Sizeable differentials in birth propensity occur by parity, and the time-path of parity specific rates can vary substantially, particularly the contrast between low vs. higher order births (Ryder 1986; Feeney and Yu 1987; Ní Bhrolcháin 1987; Rallu and Toulemon 1994; Morgan 2003). In addition, progressions of differing orders appear to be subject to different influences (Bulatao 1981; Namboodiri 1981; Isaac et al 1982; Yang 1994; Andersson 2001). Rates or probabilities specific by parity and/or duration since previous birth also have the important substantive property that they reflect directly the sequential nature of the family building process, and hence are close to the behavioral specifics of the processes giving rise to aggregate change.

A critical behavioral issue in the context of fertility measurement—the substantive counterpart to the statistical-demographic period vs. cohort contrast—is the role of personal time and historical contingency. How far are childbearing decisions influenced by (a) current state (e.g. parity, age, duration) vs. previous fertility history (beyond that represented in current state) vs. intentions or expectations regarding either future or overall lifetime fertility and (b) in a historical context, by current, past or expected future social and economic circumstances? A pure period-based behavioral model such as Ryder's (1973: 504) suggestion that couples may make childbearing decisions sequentially—month by month or year by year—will lead us to choose fertility indicators that are strictly period-based, and to look for determinants of this kind also.[1] An alternative is the cohort approach explicitly adopted by Butz and Ward (1979) whose indices of period quantum[2] (or level) and timing, in both their ex-post and ex-ante forms, reflect the supposition that annual childbearing decisions are taken relative to an overall desired or expected family size. Lee (1980: 208) is also clear about his behavioral assumptions—that the ultimate desired family size of any cohort is not fixed but may change over time, and that at each age "the annual birth rate is a fixed proportion of additional desired fertility." Both of these see period rates as essentially driven by a cohort target, fixed in the case of Butz and Ward but, for Lee, subject to alteration through the life course. In recent years, research on the behavioral underpinnings of fertility change has, however, not advanced greatly, and fewer linkages have been made between behavioral assumptions, on the one hand, and either measurement schemes or aggregate trends, on the other.

The principal demographic-statistical requirement for measures of period fertility as dependent variable is that they are a valid representation of year on year change[3]. There is, however, no independently ascertainable, true value of period fertility against which to evaluate a proposed period fertility indicator. An assessment of the validity of the time trends depicted cannot therefore be made directly. We can, however, specify some

---

[1] Ryder has been a strong advocate of the cohort approach to fertility but confessed to doubts on the question when confronted with aggregate change in US fertility combined with the findings of the US fertility surveys (Ryder 1973).
[2] The quantum idea essentially refers to the level of fertility, but with any timing component removed.
[3] Note that throughout the paper it is assumed that fertility rates and indices are measured accurately in the sense that numerators and denominators are measured without error.

minimal conditions. Where used as a dependent variable, an indicator of period fertility should be uninfluenced by nuisance factors—that is, any substantial demographic influences on fertility rates that we acknowledge as occurring but do not wish to explain. The effect of such factors can be removed in several ways—by increased specificity, by standardizing, or by some form of regression. Thus, age specific rates have the rationale or removing age structure effects—rates vary substantially by age, and age structure is accepted for many analytical purposes as a given rather than as something to be explained. Beyond age, the distribution of women by parity is an important potential nuisance factor, in that birth rates vary by parity, and parity specific rates often follow different trends. The biasing effect of the parity distribution can be removed by parity specific rates. Parity specific indicators are, as we have seen, desirable also on substantive grounds.

Further disaggregation or standardization by age and/or duration may also be required, for analogous reasons. Age-parity or parity and duration-specific rates, as well as period parity progression ratios can be used for the purpose, though data availability can be a limitation. In low fertility societies with substantial rates of childbearing outside of marriage, first birth rates should be specific by age, as should the first period parity progression ratio. But duration-specificity is almost certainly a better choice than age for these indicators for birth orders two and above—from the first birth onwards, duration appears a more natural dimension of personal time in behavioral terms than age, since it represents better the sequencing of family formation. Furthermore, age-parity specific rates of order two and above are strongly influenced by those of earlier orders, and particularly of the first birth. With age-parity specific rates, apparent timing influences in aggregate second and later birth rates may reflect wholly or partly timing shifts in the rates of previous birth orders, while such confounding is absent when duration since previous birth is the measure of elapsed time for birth orders 2+. Note that in removing the effects of age and/or duration, we implicitly accept as given the variation in rates by age and/or duration and so decline to account for these.

Is tempo adjustment appropriate for measures of period fertility as dependent variable? The argument for such correction is that period measures are biased or distorted by timing change. In fact, considered as a dependent variable, the conventional TFR can misrepresent period change even when timing is stable, since it is standardized only for age and can be influenced by changes in the parity distribution. If on the other hand we acknowledge the parity dependence of rates and represent period fertility as dependent variable by rates specific by parity and by age for the first birth and parity and duration for second and later births, or measures standardized for these factors—are any remaining timing effects a source of bias or distortion? On the present view, the answer is no. The reason is two-fold and the argument is illustrated via the start of childbearing, which is subject to the largest timing effects.

First, part of what is thought of as distortion due to timing change is really a compositional effect and can be removed by standardizing for parity and either or both age and duration. For example, a progressive delay in first births will result in an increase in the proportion of women at younger ages who are childless, and so at risk of a first birth. In such conditions, time-trends either in age specific rates or in unconditional first birth rates by age (incidence rates) reflect not only change in the birth propensities of childless women—the true phenomenon of interest—but also changing proportions childless at each age, due to later timing—a compositional effect. The same holds in reverse when first births are advanced. This compositional but timing-related effect is, however, eliminated by confining age specific first birth rates to childless women. More generally, methods have been available for some time that remove this spurious timing-related influence from period measures—age-parity specific rates for first birth together with parity- and duration-specific rates for later births, and period life tables that synthesize these (Henry 1953; Feeney and Yu 1987; Ní Bhrolcháin 1987; Murphy and Berrington 1993; Rallu and Toulemon 1994).

Second, genuine timing effects—manifest, in the case of first birth, as a shift along the age axis in the birth rates of the childless—are not distorting to measures of period fertility as dependent variable, because real tempo change is part of what we are, or should be, trying to explain. For example, the baby boom of the late 1950s and 1960s was partly due to accelerated childbearing (Butz and Ward 1979; Ryder 1980). If part of the explanation is that post-war prosperity, full employment and high wages induced couples to marry earlier and have children sooner than in preceding periods, that faster pace of family formation needs to be represented on the left hand side of the equation. Similarly, later childbearing is one aspect of what we need to explain in relation to current fertility trends in developed societies—and to do so, the full effect of genuine timing change should be retained in measures of fertility as dependent variable. The same applies to changes in variance discussed by Kohler and Philipov (2001)—they too need to be accounted for in substantive terms rather than removed as a nuisance factor from period fertility as dependent variable.

An analogy may help to illustrate the argument. 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. 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 at a given stage along the way, we are measuring something other than speed at a particular time-point. We would, in addition, either have to 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 may be sufficient to highlight several key points. Our task in explaining period fertility trends is analogous to accounting for the speed of the car at successive points during its journey. Explaining episodes of acceleration and deceleration is comparable to explaining tempo effects on period fertility. These two together are, however, a different problem from either measuring or 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, though one that may alter during the journey. In the present case the analogy is between the car's trajectory and aggregate fertility movements, and no assumption is needed about intentions regarding either destination, speed, or duration of the journey.

Tempo adjustment of period fertility as dependent variable could potentially be argued for if several conditions were to hold: that in period mode, the quantum of fertility and its timing are separable in a quantitative sense, that they reflect distinct aspects of the underlying behavioral process, and that they respond differently to change in social, economic and other determining factors. It is not at all clear that these conditions are currently met in demography. There appears to be little agreement about whether the quantum idea is applicable in a period context and if so what it means (see e.g. the varying definitions and views of Butz and Ward 1979; Ryder 1980; Pressat and Wilson 1988; Murphy 1993; Bongaarts and Feeney 1998; van Imhoff and Keilman 1999; van Imhoff 2001; Bongaarts 2002; Koehler and Ortega 2002; Schoen 2004). And while several specifications of period quantum and tempo measures have been suggested (e.g. Butz and Ward 1979; Ryder 1980; Bongaarts and Feeney 1998; Kohler and Ortega 2002) we have little evidence thus far that these numerical quantities correspond to a real, substantive aspect of the processes giving rise to fertility in period mode. That is, we have little solid evidence that at the individual level decisions about timing and quantity are made independently rather than being a joint process. On the question of response, instances can be found of changes in timing in reaction to socio-economic determinants—the Swedish speed premium effect being a very clear-cut case (Hoem 1990; Andersson 1999; Andersson et al 2006). But it is not obvious that such cases are exclusively due to timing effects, nor that currently available indices of timing would represent them accurately.[4] Nor has it ever been shown that where socio-economic factors influence timing, these have nothing but a tempo effect. Empirical evidence demonstrating that quantum and tempo are genuinely distinct aspects of the behavioral processes underlying period fertility movements would be required to clarify the matter further.

Establishing a link between behavioral processes and demographic indices—whether of quantum, tempo or the two combined—is an essential step in arguing for the relevance of the indices concerned for explanatory purposes. If the two cannot be firmly linked, then either the indicator(s) proposed, or the explicit or implicit behavioral concepts are of unproven value in explaining time trends, even though the indices may have instrumental value as predictors.

How can we tell whether an indicator is suitable as a dependent variable—that is, how can we validate the measure? No independent criterion is available by which to assess how well a period measure, or a set of such measures, reflects temporal change in

[4] For example, the duration-specific indicators for birth orders 2 and above of Hoem (1990) and Andersson et al (2006) give much clearer evidence of the timing shifts associated with the Swedish speed premium effect than do the adjusted mean ages of Figure 4b of Kohler and Philipov (2001). The latter would not, alone, be convincing evidence of a speed premium effect whereas the analyses by duration are very compelling indeed. Of course, the Swedish maternity pay regulations are framed in terms of time since previous birth, and so measures that are specific by duration since previous birth naturally fit the structure of the maternity pay incentive.

fertility. But an indirect check on the validity of a measure of period fertility as dependent variable *is* available—viz. explanatory success. An indicator or set of indicators of period fertility as explicandum can be considered useful or valid to the extent that it is embedded in an empirically successful explanation of period trends. As Ryder suggested, we will know we have the right measures when we have a good explanation of time trends. An effective explanation of post-war fertility trends in developed countries would be based, at least in part, on some form of quantitative model of time series data, since an array of ever changing influences are almost certainly at work. However, systematic attempts at explaining aggregate fertility trends via statistical or econometric models are and have been rare. The Easterlin approach has not been successful in meeting empirical tests (Waldorf and Byun 2005) and there has been little work on aggregate movements in fertility to follow that of Butz and Ward (1979) and Lee (1980). Hence we currently have no well-formulated explanatory framework, backed up by solid evidence, to underwrite particular approaches to period fertility measurement, even though good substantive grounds exist for e.g. advocating parity specific measures. Nevertheless, the key point here is that a convincing explanation, and not a check against cohort values, is the appropriate criterion for evaluating an indicator of period fertility as dependent variable.

2.2 Anticipating future fertility

A further purpose for constructing indicators based on period fertility is to anticipate future fertility. This is often the explicit or implicit rationale for studying current or recent trends. The outlook for fertility is a natural preoccupation in a profession whose most sought-after applied function is providing information on future population trends. It is also a natural focus of interest when fertility levels are causing policy concern, as is currently the case in many low fertility societies (Lutz and Skirbekk 2005; McDonald 2006). Most of the difficulties that beset period fertility measurement arise from attempting to expand its time reference into the future. Indicators that reflect some aspect of the (prospective) longer run are clearly more difficult to create from the data of a single period, and inherently less satisfactory, than are measures intended to reflect purely current reproductive performance.

Period-based measurement can be geared to reflecting fertility prospects in three ways. One objective is to estimate the fertility of cohorts—discussed here in terms of birth cohorts, though marriage cohorts and parity cohorts may also be of analytical interest. Second, period fertility may be examined for indications about future trends, in a broader and less specific sense. Forecasting or projection is a third approach to the fertility of the future.

2.2.1 Cohort fertility

Where data are available on completed childbearing, measuring cohort fertility is straightforward. A variety of indices of both level and timing can be specified, up to the limits of data quality and sample size. Practical problem may arise regarding e.g. whether immigrants should be included, but the indices themselves are straightforward to define. However, where the childbearing of the cohorts in question is incomplete, period rates are used in the estimation process and it is then that difficulties arise.[5] Leaving aside projection, which is discussed in a later section, conversion of period fertility into cohort

[5] van Imhoff (2001: 24-5) expresses the matter thus: "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…"

terms takes two forms: calculating synthetic cohort measures and demographic translation.

2.2.1.1 Synthetic cohort indicators

The simplest and most common demographic device for bridging the period-cohort gap is the synthetic cohort indicator. Such measures have a dual function in demography, interwoven in such a way as to cause much confusion. On the one hand, they can be thought of as quantifying the fertility performance of a period through a summary indicator that has a statistical rather than a demographic role. On the other, they convert the rates of a single period into an estimate of experience that in reality extends over many years, and in some cases over a lifetime. To fulfil this second role, the assumption is required that the age and/or duration-specific rates of a given period obtain at successive ages/durations[6]. The best known hypothetical cohort fertility index is the conventional TFR, normally presented as an estimate of the mean family size of an imaginary cohort experiencing a particular period's age-specific fertility rates throughout their childbearing years. Other synthetic cohort indicators include measures of level such as period parity progression ratios in their various forms, or timing indicators such as the standardized mean age at childbearing[7] or period-based birth intervals (Henry 1953; Whelpton 1954; Feeney and Yu 1987; Ní Bhrolcháin 1987; Rallu and Toulemon 1994; Kohler and Ortega 2002).

[6] See Rallu and Toulemon (1994) for a period fertility life-table specific for both age and duration since previous birth.
[7] The standardized mean age at childbearing can be seen as a synthetic cohort measure, since it represents the mean age at birth of an imaginary cohort experiencing the age specific rates of a given period.

In the literature on tempo adjustment, the charge is leveled at the TFR that it is biased or distorted whenever the timing of fertility is changing (Bongaarts and Feeney 1998; van Imhoff 2001; Zeng and Land 2001, 2002; Kohler and Philipov 2001; Bongaarts 2002; Sobotka 2004). However, terminology varies, with some sources using the concept of bias, either alone or interchangeably with the idea of distortion (Bongaarts and Feeney 1998; Bongaarts 2002; Bongaarts and Feeney 2002; van Imhoff and Keilman 1999; Kohler and Philipov 2001; Zeng and Land 2001, 2002) and others referring exclusively to the notion of distortion (Bongaarts 1999; Bongaarts and Feeney 2000; Frejka and Ross 2001; Kohler et al. 2002; Schoen 2004; Sobotka 2004). The terms will be used interchangeably throughout the paper. The bias in question is not statistical, in that there is no question of a probability distribution for the TFR (see also Zeng and Land 2002, note 1). The most reasonable way of construing discussion in this area is that it is measurement bias that is at issue. However, it is not at all clear what recent commentators believe to be biased or distorted by the TFR but two interpretations are possible.

A first possibility is that any period measure of fertility that is influenced by timing effects is considered to be distorted by definition. This position would have to be proposed and supported explicitly, however, since it is not self-evidently true that a change in tempo necessarily distorts a measure of period fertility. Some scholars (Bongaarts and Feeney 1998, 2000; Zeng and Land 2002) appear to espouse a view somewhere close to this—but to sustain the point, they would have to justify the position, and tackle the alternative view that genuine timing effects are integral to, and not distorting of, period fertility when considered as a dependent variable.

The claim that the period TFR is "distorted" goes back to Ryder (1964), who meant that period total fertility was a distorted version of cohort values. This gives us a second way of interpreting the notion of distortion as measurement bias and appears to be what many commentators have in mind in referring to tempo distortions. However, the assertion that bias is present only when timing is changing reflects a theoretical rather that an empirical construal of the TFR. As a measure of a real, as distinct from a theoretical, cohort quantity, the TFR is, in fact, *always* mistaken and not solely when fertility timing is changing. The conventional TFR is equivalent to cohort fertility only where age-specific rates are either fixed or randomly distributed around a given period's values. Such stability is rare. Hence, the conventional period TFR is almost always biased as a real-world cohort estimator since it corresponds only by chance to the mean family size of any cohort at risk during the period.

Thus, bias in the TFR as an indicator of real cohort values is present whether or not the timing of fertility is changing. Shifts in fertility tempo may of course increase the bias—that is, create a greater discrepancy between period and corresponding cohort TFRs, particularly if only a timing change occurs with no change in cohort quantum. There are, therefore, two sources of bias in the TFR as an measure of real cohort fertility—one due to the non-fixity of age specific rates and the other due to tempo change. Is adjustment of the period TFR for tempo effects warranted, in principle, in this context? The answer is probably yes, *if* the objective is to get closer to cohort mean family size than the figure given by the period TFR, and if a case can be made that the timing change is a short-run phenomenon. The same goes for other synthetic indicators, such as period parity progression ratios, *if* the objective is to get closer to the cohort equivalent. The case for adjustment is not that tempo effects are always offset in later periods, but that they often are to some degree—though the extent to which compensating movements in rates occurs is unpredictable (Lesthaeghe 2001; Frejka and Calot 2001; Billari and Kohler 2004). While tempo adjustment can do nothing to correct for the bias due to non-fixity of the rates, it can sometimes be expected to reduce the bias in period synthetic indicators, considered as estimates of cohort quantities. The argument for tempo adjustment is pragmatic and empirical—based on the behavior of fertility series in the past, as well as judgment about the likelihood that a current shift is mainly a timing phenomenon—rather than theoretical, though currently available adjustment procedures are derived on theoretical grounds. The case against is that we have no way at present of distinguishing a short term timing shift from a long-run change in level. Whatever method is used, biases will remain in tempo adjusted period measures considered as estimates of real cohort quantities, both because rates schedules rarely if ever remain fixed and because timing changes may be neither as systematic nor as persistent as recently proposed adjustment procedures assume.

2.2.1.2 Demographic translation

A further type of conversion, demographic translation, goes somewhat beyond the synthetic cohort principle and treats the transfer from period to cohort as a systematic problem to be solved. Translation seeks a determinate, formal mathematical relationship between the level and timing of period and cohort fertility, whether specific by age (Ryder 1964; Foster 1990) or by order of birth (Keilman 1994). Translation can proceed from period to cohort or vice versa. For accuracy, demographic translation requires smooth patterns of change much simpler than those found in observed fertility series. It fulfils a useful function in providing an account of overall fertility trends that dampens the more volatile period picture.

The appropriate criterion against which to evaluate a period indicator intended to estimate cohort fertility is clearly the fertility of real cohorts. The conventional period TFR performs badly on this test, having long been known to be a poor estimator of cohort mean family size. The discrepancy has often been illustrated graphically by the much larger swings in the period TFR than in cohort total fertility (for recent examples see e.g. van Imhoff and Keilman 2000; Frejka and Calot 2001; Schoen 2004). It is this shortcoming that has, in part, motivated the search for improved ways of converting between period and cohort, among them tempo adjustment. Adjustment methods vary in how well they estimate cohort quantities. Bongaarts and Feeney do not see their adjusted TFR as an estimator of cohort fertility but several commentators have construed it in this way (van Imhoff and Keilman 1999; van Imhoff 2001; Kohler and Ortega 2002; Smallwood 2002; Schoen 2004; Sobotka 2003). Evidence suggests the adjusted TFR has limited accuracy on an annual basis (van Imhoff and Keilman 1999; van Imhoff 2001; Smallwood 2002; Schoen 2004) but Bongaarts and Feeney (1998) show that an average of the annual adjusted TFR values tracks cohort fertility well. Neither the Kohler and Philopov (2001) nor the Kohler and Ortega (2002) measures appear to have been evaluated against the cohort equivalents. Butz and Ward's (1979) Average Completed Fertility index, advocated by Schoen (2004) as a measure of period quantum, performs better than the alternatives considered by Schoen in approaching cohort values. But it has several acknowledged limitations as a period measure, requiring in particular that cohort fertility be already known, and thus conditioning on the future. It can only be used prospectively as an estimator of cohort fertility by substituting annual intentions data for completed cohort fertility. Its performance in that mode has not been evaluated.

Demographic translation has the unambiguous objective of translating between cohort and period formats, and so the criterion of success in each case is perfectly clear. Ryder's classic translation procedure approximates more closely than the period TFR to cohort values over the medium term, but is nevertheless quite inaccurate. Indeed, Ryder (1964) acknowledged that its empirical performance is severely limited by the mathematical approximations required (see also Murphy 1993; van Imhoff 2001). On one view, the task of translation is impossible since "in real life the factors involved in explaining the link between period and cohort quantum are so complex and subtle… that we will never be able to describe it completely" (van Imhoff 2001: 25). If this is so, then translation in either direction, though an interesting theoretical problem, and offering a useful tool for projection and modeling purposes, cannot be empirically successful since period influences are irregular, unpredictable, and appear not to have cyclical features.

2.2.2 Future fertility

Beyond explicit attempts to estimate cohort levels of reproduction, period fertility may be the vehicle for a more general discussion of fertility prospects—either indirectly or, if directly, with a fairly inexact time reference. Such discussion can involve measurement to varying degrees. At one extreme, analyses, arguments and views may be offered on possible futures, without any special attention to the choice of indicator. At the other, the subject may be addressed via an approach to measurement that implicitly but clearly reflects a reference to future reproductive performance. A broad reference to the future is found in the discourse of fertility definitions and measurement in a variety of ways. It appears to be what e.g. van Imhoff (2001) has in mind when he says 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 expressions that refer to the completed fertility “implied by” current rates. The concepts of “true” or “underlying” level of fertility could also be argued to refer to longer-run fertility levels in some non-specific sense (at least on one reading—another will be considered in a later section). These terms are never applied to simple incidence or occurrence-exposure rates, but carry the connotation of a mean family size, or parity progression ratio, or other synthetic quantity. In contrasting a measured with a true or underlying quantity, they implicitly draw a distinction between the temporary fertility conditions of the current period and longer term fertility levels. Also the real as opposed to synthetic realization of such quantum indicators occurs over long stretches of time rather than in individual periods.

One interpretation of the Bongaarts and Feeney adjusted TFR, and related measures, is that they serve this function—i.e. are an attempt to infer longer run fertility in some non-specific sense, though not cohort fertility. Bongaarts and Feeney suggest, for example, that the adjusted TFR gives “a better answer to the question of how many births women will have if current childbearing behavior continues into the future” (Bongaarts and Feeney 1998: 285). While certainly based on the rates of a single period, and in that sense a pure period measure, it can be argued that assumptions about future fertility are inherent to the index, as of all synthetic cohort measures that are interpreted as such. This also appears to be a reasonable inference from the fact that tempo adjustment has been understood, in practice, as carrying implications for long-run trends in fertility (Lesthaeghe and Willems 1999; Morgan and King 2001; Kohler and Ortega 2002; Bongaarts 2002). The upward correction to recent period TFRs resulting from tempo adjustment has been interpreted as implying that period fertility rates are temporarily low and hence as implicitly predicting a future recovery in fertility, though doubts have been expressed as to the likelihood or extent of such recuperation (Lesthaeghe and Willems 1999; Frejka and Calot 2001; Sobotka 2004).

If the objective is to get some idea of longer-run mean family size, adjusting for timing change can be justified on the same pragmatic grounds as apply to estimating cohort fertility. The record suggests that declines or rises in period rates associated with changes in the age at childbearing are often though not always compensated for in subsequent years. If grounds can be found for interpreting a short-run trend as primarily a timing change, adjusting for timing effects may be, in principle, a reasonable procedure to adopt in attempting to get closer to longer-run quantum estimates. However, as in the cohort case, such adjustment can be considered a type of forecasting and methods need to be evaluated for their success in this respect.

While the appropriate criterion for evaluating period estimators of cohort parameters is clear-cut, the task of evaluating period measures with an extended time reference but conceived in other terms—measures of longer run fertility levels, of underlying completed fertility levels, or of period quantum—is more complex. An index intended to reflect longer-run fertility levels in some sense can be evaluated by some kind of longer-run averaging procedure (see e.g. the averaging procedure of Bongaarts and Feeney 1988). Both the ideas of underlying fertility and of fertility quantum appear to be ill-defined as empirical entities in a period context—as noted earlier, there seems to be little agreement as to what quantum means in relation to periods though the idea is perfectly clear when applied to a cohort. The meaning of such concepts could be considerably clarified by specifying empirical criteria against which to evaluate them.

2.2.3 Projection

Projection or forecasting is the final way of anticipating the future. As a source of information on either a generalized notion of current and future fertility levels or cohort fertility in particular, projection or forecasting has several merits. The estimates produced are presented as projections rather than as measures, the inherent uncertainty of the estimates is acknowledged, and assumptions about future movements in rates must be made explicit. In addition, where cohort fertility is the target, the accumulated fertility experience of incomplete cohorts up to the base year of the projection can be incorporated into estimates of their future completed fertility. Treating the anticipation of future fertility as a forecasting problem, rather than as something to be inferred from current period rates, appears both more transparent and empirically more realistic (on this point see also Lesthaeghe and Willems 1999 and Schoen 2004).

Improving the accuracy of fertility forecasts is a continuing concern for applied demography. A potentially powerful approach that has been pursued less actively than it might be is the projection of parity progression ratios on either a period or cohort basis. Feeney (1985) appears to have been the first to propose this strategy and some useful findings have been reported by Toulemon and Mazuy (2001) and by Sobotka (2004). Kohler and Ortega (2002) combine this approach with tempo adjustment, using their adjusted parity progression measures as a basis for projection under several scenarios. Again, these need to be evaluated for forecast accuracy. Schoen (2004, footnote 3) expresses reservations about the discretion of the analyst in Kohler and Ortega's procedures, and appears to imply that identifying timing effects from period rates is a measurement rather than a forecasting problem. However, if period fertility measures are intended to reflect future fertility—as in any attempt to infer a cohort timing effect from period rates—they are in practice an attempt to forecast and so necessarily require investigator discretion.

The criterion against which fertility projections should be evaluated is unambiguous—fertility out-turn, whether cohort or period. Nevertheless, assessing the comparative performance of fertility projections over time is potentially complex, requiring analysis not only by initial date, and duration of the projection, but also by period (see Keilman's (1990) analysis of forecast errors in a framework equivalent to an age-period-cohort analysis). Fertility projection has had limited success in low fertility societies and there is little evidence that fertility projections based on a cohort approach are any more accurate than those based on period lines (Booth 2006; Keilman 1990; Keilman and Pham 2004;

Lee 2004). In addition to evaluation for forecast accuracy, fertility projections could usefully be compared with other methods—adjustment, translation—of anticipating future fertility.

2.3 Theoretical models

A third way in which indices of period fertility can be deployed is theoretical. This is really not a measurement activity at all, but a form of population modeling. The representation of the TFR as the mean family size of a hypothetical cohort is an instance. In a theoretical population with fixed age specific rates, it estimates cohort fertility without bias. Under changing tempo, however, the conventional TFR is biased as an measure of cohort fertility, in the sense that it is not equivalent to the cohort mean family size in a theoretical population subject indefinitely to the rates, and to the timing shift, of a given period. In a theoretical context, procedures to adjust for tempo change can be designed around the particular type of tempo shift assumed to operate. These are clearly appropriate if the objective is to estimate theoretical cohort fertility. This could be an alternative way of construing the Bongaarts and Feeney adjusted TFR—as the cohort mean family size in a theoretical population with the age-order specific rates of a given period and subject continuously to the tempo change of that period—and is the preferred interpretation of Zeng and Land (2001) and of Rodriguez (2006).

It can be argued that the concept widely used in demography of the mean family size "implied by" the rates of a given period is essentially of this type: a theoretical construct rather than an empirical measure. Clearly nothing in the real world is literally implied, in a logical sense, by the rates of a single period, since real world populations are neither stationary nor, for the most part, stable. However, we might concede that orders of magnitude are probably implicit in period rates—women who are of childbearing age in a period when the TFR is, say, 6 are extremely unlikely to have a mean family size of 2, for example, and vice versa. But estimates at the degree of resolution that is usually sought in attempts to refine period fertility measures cannot be considered to be logically implied by the rates of a single period, except in a theoretical context.

Evaluating indices of fertility that are defined within a theoretical population model is essentially a matter of checking mathematical derivations, and possibly also the theoretical coherence and utility of a particular hypothetical measure. No empirical criterion is relevant for evaluating a measure construed as reflecting a hypothetical entity within a population model. An empirical yardstick becomes necessary only when a theoretical specification is regarded as measuring a real world process.

2.4 Communication and public information

A final reason for choosing an index of period fertility is to convey information on fertility trends to non-specialist audiences of various kinds. Time trends in fertility have practical consequences that matter to policy makers, service providers, the business community, journalists and the public at large. Information is supplied to these various audiences by government statisticians and demographers through regular updates and commentary on fertility trends. Such dissemination is not without its hazards (Teitelbaum 2004). Period fertility, being up to date and changing faster than the cohort equivalent, tends to be the focus of such interest, and so an indicator, or set of indicators, of current childbearing is needed for popular dissemination.

The period total fertility rate fulfils that role in most developed countries and appears reasonably adequate for the purpose, though the crude birth rate or the general fertility rate or any of a number of other indices could serve the same purpose and do so in more restricted vital registration systems. The natural criteria for assessing fertility indicators for communication purposes are how easy they are to produce and how accurately the indicator chosen can be interpreted for and by non-specialist audiences. The TFR has clear advantages in that respect, though these should be distinguished from its technical qualities as an indicator. A fertility index deployed to such practical ends need not have a theoretical pedigree, any more than e.g. the Human Development Index, or poverty indicators, or the retail price index, do. Nor need it be as refined as those used for scientific research. It is not obvious that tempo adjusted measures are needed in this more popular context since changes in timing can be conveyed by reporting time trends in mean/median age at birth or, for preference, at first birth. Also, non-specialist users may often not be primarily interested in fertility *per se*. Rather, medium to long-run broad population prospects may be of greater pertinence for general policy and public information purposes and these depend not only on fertility but also on age structure, mortality and migration. National policy decisions on e.g. planning facilities and services, maternity provision, family policy, immigration, or pension provision require information going far beyond recent trends in a summary fertility index—projections of annual birth numbers, of cohort fertility, of population size and structure and so on. For this reason, an unrefined index of fertility may be quite adequate for most non-technical users' purposes, since it will usually be supplemented by other kinds of data relevant to future prospects, at least in a policy context.

3. Discussion

Several arguments relating to recent debates on fertility measurement are advanced in this paper. First, there are a variety of reasons for measuring period fertility, and not all measures are suited to each purpose. Period measures are relative not so much to the observer (van Imhoff 2001) as to the purpose in hand. Recent debates on fertility measurement have tended to overlook the multiple objectives for which period fertility measures can be employed. Second, a major distinction is between period measures intended, on the one hand, to represent period fertility as dependent variable, and, on the other, to anticipate longer run fertility levels, whether of cohorts or in a more diffuse sense. Third, the accuracy of indicators of each type needs to be evaluated, and clarity is required on the appropriate criterion in each case. Indicators based on concepts that lack a convincing direct or indirect criterion may be of limited value as measures of empirical conditions.

The monitoring and measurement of fertility trends mostly has the explicit or implicit purpose of attempting to divine the future, usually by approximating long run mean family size in some sense. This is a perfectly valid aim and is not surprising, given the applied orientation of much demographic activity. One consequence is, however, that measurement and forecasting are conflated in demographic thinking about fertility. We should recognize that what we think of as measurement is often really a form of forecasting and that merging these distinct activities results in unnecessary confusion.

Evaluating how best to measure period fertility as dependent variable, and how to analyze the factors associated with aggregate time trends in a statistically systematic way, is an important research objective. It, and the measurement effort it entails, is a distinct problem from that of assessing future fertility prospects, whether on a cohort or period basis. For example, if we wish to specify a measure of period fertility as dependent variable, the question of whether trends in different age groups compensate for each other now, or will do so in the future, is completely irrelevant to the specification. The essential contrast is between specifying appropriate measures of period fertility as an *outcome* of current and past demographic and socio-economic factors versus measuring it as either a *predictor* or *determinant* of future fertility and population parameters.

Another consequence of demography's future orientation is that although commentary on past fertility trends is extensive and although many explanatory schemes have been put forward (Morgan and Taylor 2006), a statistically systematic approach to explaining aggregate trends in low fertility societies has been neglected in demography, and with it the specification of measures of period fertility as dependent variable. Further possible reasons for this neglect include the limited success of attempts to model time series in the past (Lee 2004), the shortcomings of conventional micro-economic theory which tends to inform such work—inappropriately, since it attempts to explain an aggregate phenomenon in terms of micro-level theory (Murphy 1992)—the neglect of theory better suited to the realities of fertility variation, especially parity-specificity (Leibenstein 1974; Namboodiri 1981), the recent dominance of micro-level analytical approaches, and the absence of overlap between an interest in formal demography and either structural modeling or explanatory investigation more broadly. Like other disciplines, demography has a division of labor. Analysts with an interest in technical and measurement issues tend not to be concerned with the substantive factors that drive time-trends, and vice versa. The recent literature on tempo adjustment is, accordingly, largely focused on the technical properties of the period indices discussed. Little or no attention has been given to the suitability of the fertility measures discussed as dependent variables. Finally, there has recently been a decided lack of attention to the substantive processes underlying aggregate fertility change—the question of how decisions are made in personal time, and how these are influenced by historical change, both short and long-term. For example, the idea of postponement has been adopted widely in demography in the last decade or so to describe recent trends in fertility in developed societies. No clear behavioral model has been put forward to give substance to the concept, and little or no attention has been given to testing the implied behavioral model against alternatives. Behavioral mechanisms have been overlooked despite the massive scale of the move to later childbearing, and although the forecast implicit in the postponement idea of a recuperation in fertility has been challenged (Lesthaeghe and Willems 1999; Frejka and Calot 2001). The behavioral underpinnings of fertility change need to come back on the agenda of fertility studies and with it research on the formal modeling of aggregate change.

Demography has a long tradition of using synthetic cohort indicators, interpreted via stationary population assumptions. But a measure summarizing one period’s rates, however transformed, is a slim basis for predicting the future. Fertility time trends do not follow any known deterministic law. If we want to estimate future completed fertility, it appears short-sighted to ignore the accumulated fertility of the recent past, whether in period or cohort form—but that is what synthetic cohort measures do, even when adjusted for shifting tempo. Attempts to anticipate future trends would be better served by an empirical search, by means of statistical or econometric models, for systematic relationships between the fertility performance of the past—whether in period or cohort format—and that of the future.

The ubiquity in demography of some version of the period TFR is, for all its faults, neither accidental nor arbitrary. It is, after all, the counterpart of cohort mean family size, which ultimately reflects population reproductivity. It is "good value" as an indicator, being standardized for *the* major demographic influence on vital rates, while requiring relatively simple inputs. It gives an excellent idea of crude orders of magnitude and so is useful for broad-based comparisons. It is unexceptional in a theoretical context. Finally, it has an attractively meaningful interpretation for individuals, and so is a successful device in communicating with general audiences. That it is easily understood by non-specialists is often cited as an advantage, but this has no bearing on its merits as a scientific tool. Where, as currently in developed societies, fertility response through time is distinctively articulated across age and life-stages, the crudity of the TFR—at least in its role as dependent variable—needs to give way to detailed measures that are more convincing, and more defensible, in statistical and behavioral terms. Besides, the TFR performs poorly as a forecaster in real-world situations where precision matters (period parity progression ratios may well do better, but have not been evaluated for forecasting purposes). If the question is whether long-term mean family size will be closer to e.g. 1.8 than to 2.1 or 1.6 vs. 1.3, we require a demographic technology that goes beyond estimating out-turn by the assumption that this year's rates are fixed. Redefining the question at issue as a forecasting rather than a measurement problem could be a productive way forward (see also Lesthaeghe and Willems 1999; Schoen 2004).

The sense in which the terms bias and distortion are used in the literature on adjustment is not always clear. Two different concepts of bias or distortion have been distinguished in the present paper. Measurement bias is present when the indicator chosen systematically gives a mistaken reading of, or misrepresents, the phenomenon in question. The claim that the TFR is biased when timing is changing appears to be rooted in this concept, and can be interpreted in three possible ways. One is that a period measure influenced by timing effects is by definition biased. A second reading is that the TFR is biased relative to real cohort values, and a third, relative to theoretical cohort values. In the second of these, bias is present whether timing is changing or not, and in the third, only when timing is changing. Measures of period fertility as dependent variable are not designed as indicators of cohort values, but can be biased in a further sense, by the presence of nuisance factors, if inadequately standardized. The paper has argued that timing change is not a source of bias but integral to period fertility as dependent variable—part of what we should be trying to explain.

The synthetic cohort device is deeply engrained in demographic thinking. Ironically, it is called on precisely because the hypothesis on which it is based—stable rates—is rarely valid. Period fertility fluctuates, sometimes sharply, and the period TFR cannot therefore be relied on to reflect long run completed family size. This has given rise to concepts such as period quantum and "underlying" or "true" fertility, or the total fertility "implied by" current rates, that essentially assert a contrast between the apparent and the real. Such terms are widely employed in fertility analysis and are probably subject to a process of reification (Wilson and Oeppen 2003). But they are ill-defined and lack a clear empirical reference. The view taken here is that, if meaningful at all, these constructs refer to longer-run fertility in some sense. Such concepts can be useful only if there is clarity about their intended status as theoretical vs empirical entities and, where intended as empirical concepts, that a criterion is specified against which to evaluate indices intended to represent them. Validation is central to clarifying the meaning and intended purpose of period fertility measures, as well as to establishing their utility.

The central arguments of the present paper are as follows.

1. Period fertility is measured and analyzed for a number of objectives: to explain past trends, to predict or anticipate future fertility, as input to theoretical models, and to communicate with non-specialist audiences. Any proposal for an index of period fertility should specify which of these objectives it is intended to serve.
2. The properties desirable in a period fertility indicator depend on the purpose for which it is intended. An index of period fertility as explanandum need not, and should not be expected to, function as an indicator or predictor of future fertility trends. An index of fertility intended for scientific analysis need not be readily understood by non-specialist audiences.

3. Establishing the utility of a period fertility measure requires that it is validated by an appropriate external criterion. Explanatory success is the appropriate criterion of a measure of period fertility considered as dependent variable. By contrast, approximation to cohort values, or to longer run fertility outcomes in a more diffuse sense, is the appropriate yardstick where anticipation of future fertility is the objective.
4. In period mode, unlike cohort mode, the concepts of "quantum" and "tempo" are not clearly defined. There is no evidence that these constructs are empirically independent. Nor is there evidence that they correspond to distinct aspects of real-world decisions.
5. Genuine timing effects that are not due to compositional effects are integral to period fertility as dependent variable. At present there is no justification for removing real timing effects from indices of period fertility in its role as explanandum. To defend the procedure, timing and quantum would have to be shown to be distinct behavioral processes in period mode, and the proposed indicators should be demonstrated to accurately measure the underlying behavioral components.
6. Tempo adjustment is implicitly an attempt to forecast the future rather than to measure the here and now.
7. The apparent rationale for tempo adjustment derives from the tendency in classic demographic thinking to conflate the present and the future, to confuse measures and forecasts. Demographers often interpret the question of what current fertility

"is" as meaning "what fertility is and is likely to be in the future". Synthetic cohort type indicators embody and reinforce this confusion.

8. If we wish to know what future or longer-run levels of fertility will be, we should project or forecast explicitly.

9. The notion of the mean family size "implied by" current rates has meaning only in a theoretical context. In an empirical context, it is close to meaningless and highly misleading.

4. References

Andersson, G. 1999. “Childbearing trends in Sweden 1961-1997,” *European Journal of Population* 15: 1-24.

Andersson, Gunnar. 2000. “The impact of labour-force participation on childbearing behaviour: Pro-cyclical fertility in Sweden during the 1980s and the 1990s,” *European Journal of Population* 16: 293-333.

Andersson, Gunnar, Jan M. Hoem and Ann-Zofie Duvander. 2006. "Social differentials in speed-premium effects in childbearing in Sweden," *Demographic Research* 14: 51-70.

Billari, Francesco and Hans-Peter Kohler. 2004. “Patterns of low and lowest-low fertility in Europe,” *Population Studies* 58: 161-176.

Bongaarts, John. 1999. “The fertility impact of changes in the timing of childbearing in the developing world,” *Population Studies* 53: 277-289.

Bongaarts, John. 2002. “The end of the fertility transition in the developed world,” *Population and Development Review* 28: 419-443.

Bongaarts, John and Griffith Feeney. 1998. “On the quantum and tempo of fertility,” *Population and Development Review* 24: 271-291.

Bongaarts, John and Griffith Feeney. 2000. “On the quantum and tempo of fertility: Reply,” *Population and Development Review* 26: 560-564.

Bulatao, Rodolfo A. 1981. “Values and disvalues of children in successive childbearing decisions,” *Demography* 18: 1-25.

Butz, William P. and Michael P. Ward. 1979. “Will US fertility remain low? A new economic interpretation,” *Population and Development Review* 5: 663-688.

Caldwell, John C. and Thomas Schindlmayr. 2003. “Explanations of the fertility crisis in modern societies: a search for commonalities,” *Population Studies* 57: 241-263.

Feeney, Griffith and Jingyuan Yu. 1987. “Period parity progression measures of fertility in China,” *Population Studies* 41: 77- 102.

Feeney, Griffith. 1985. “Parity progression projection,” in: *International Population Conference, Florence 1985,* Vol. 4. Liège: International Union for the Scientific Study of Population: 125-136.

Foster, A. 1990. "Cohort analysis and demographic translation," *Population Studies* 44: 287-315.

Frejka, Tomas and Gérard Calot. 2001. "Cohort reproductive patterns in low fertility countries," *Population and Development Review* 27: 103-132.

Frejka, Tomas and John Ross. 2001. "Paths to subreplacement fertility: The empirical evidence," *Population and Development Review* 27 (*Supplement: Global Fertility Transition)*: 213-254.

Golini, Antonio. 1998. "How low can fertility be? An empirical exploration," *Population and Development Review* 24: 59-73.

Hoem, J. 1990. "Social policy and recent fertility change in Sweden," *Population and Development Review* 16: 735-747.

Hoem, Jan M., Alexia Prskawetz, and Gerda Neyer. 2001. "Autonomy or conservative adjustment? The effect of public policies and educational attainment on third births in Austria, 1975-1996," *Population Studies* 55: 249-261.

Isaac, Larry, Philips Cutright, Elton Jackson and William R. Kelly. 1982. "Period effects on race- and parity-specific birth probabilities of American women, 1917-1976: A new measure of fertility," *Social Science Research* 11: 176-200.

Keilman, Nico. 1990. "Analyzing ex-post observed errors in a series of population forecasts," *Zeitschrift für Bevölkerungswissenschaft* 17: 411-432.

Keilman, Nico. 1994. "Translation formulae for non-repeatable events," *Population Studies* 48: 341-357.

Kim, Young J. and Robert Schoen. 1999. "On the quantum and tempo of fertility: limits to the Bongaarts-Feeney adjustment," *Population and Development Review* 26: 554-559.

Kirk, Dudley. 1996. "Demographic transition theory," *Population Studies* 50: 361-387.

Kohler, Hans-Peter and Dimiter Philipov. 2001. "Variance effects in the Bongaarts-Feeney formula," *Demography* 38: 1-16.

Kohler, Hans-Peter and Jose Antonio Ortega. 2002. "Tempo-adjusted period parity progression measures, fertility postponement and completed cohort fertility," *Demographic Research* 6: 92-144.

Kohler, Hans-Peter, Francesco C. Billari and Jose Antonio Ortega. 2002. "The emergence of lowest-low fertility in Europe during the 1990s," *Population and Development Review* 28: 641-680.

Lee, Ronald. D. 1980. “Aiming at a moving target. Period fertility and changing reproductive goals,” *Population Studies* 34: 205-226.

Lee, Ronald. 2004. “Quantifying our ignorance: stochastic forecasts of population and public budgets,” in Linda J. Waite (ed.) *Aging, Health, and Public Policy: Demographic and Economic Perspectives*, Supplement to *Population and Development Review*.

Leridon, Henri. 2006. “Demographic effects of the introduction of steroid contraception in developed countries,” *Human Reproduction Update* 12: 603-616.

Lesthaeghe, Ron. 2001. “Postponement and recuperation: Recent fertility trends and forecasts in six Western European countries.” Interface Demography, Vrije Universiteit Brussel. IPD-WP 2001-1.

Lesthaeghe, Ron and Johan Surkyn. 1988. “Cultural dynamics and economic theories of fertility change,” *Population and Development Review* 14: 1-45.

Lesthaeghe, Ron and Paul Willems. 1999. “Is low fertility a temporary phenomenon in the European Union?” *Population and Development Review* 25: 211-228.

Leibenstein, Harvey. 1974. “An interpretation of the economic theory of fertility: promising path or blind alley?” *Journal of Economic Literature* 12: 457-479.

Lutz, Wolfgang, Brian C. O’Neill and Sergei Scherbov.. “Europe’s population at a turning point,” *Science* 299: 1991-1992.

Lutz, Wolfgang and Vegard Skirbekk. 2005. “Policies addressing the tempo effect,” *Population and Development Review* 31: 699-720

McDonald, Peter. 2006. “Low fertility and the state: The efficacy of policy,” *Population and Development Review* 32: 485-510.

Morgan, S. Philip. 2003. “Is low fertility a twenty-first-century demographic crisis?” *Demography* 40: 589-603.

Morgan, S. Philip and Rosalind Berkowitz King. 2001. “Why have children in the 21st century? Biological predisposition, social coercion, rational choice,” *European Journal of Population* 17: 3-20.

Morgan, S. Philip and Miles G. Taylor. 2006. “Low fertility at the turn of the twenty-first century,” *Annual Review of Sociology* 32: 375-399.

Murphy, Mike. 1992. “Economic models of fertility in post-war Britain – A conceptual and statistical re-interpretation,” *Population Studies* 46:235-258.

Murphy, Mike. 1993. “Time-series approaches to the analysis of fertility change,” in M. Ní Bhrolcháin (ed.) *New Perspectives on Fertility in Britain*. London: HMSO, 51-66.

Murphy, Mike and Ann Berrington. 1993. “Constructing period parity progression ratios from household survey data,” in M. Ní Bhrolcháin (ed.) *New Perspectives on Fertility in Britain*. London: HMSO, 17-32.

Namboodiri, N. Krishnan. 1981. "On factors affecting fertility at different stages in the reproduction history: an exercise in cohort analysis," *Social Forces* 59: 1114-1129.

Ní Bhrolcháin, Máire. 1987. “Period parity progression ratios and birth intervals in England and Wales 1941-1971,” *Population Studies* 41: 103-125.

Rallu, Jean-Louis and Laurent Toulemon. 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 (6): 85-110.

Ryder, Norman. 1964. “The process of demographic translation,” *Demography* 1: 74-82.

Ryder, Norman. 1973. “A critique of the National Fertility Study,” *Demography* 10: 495-506.

Ryder, Norman. 1980. "Components of temporal variations in American fertility," in R.W. Hiorns (ed.). *Demographic Patterns in Developed Societies*. London: Taylor and Francis, pp. 15-54.

Ryder, Norman. 1986. “Observations on the history of cohort fertility in the United States,” *Population and Development Review* 12: 617-643.

Schoen, Robert. 2004. “Timing effects and the interpretation of period fertility,” *Demography* 41: 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, Tomas. 2004. “Is lowest-low fertility in Europe explained by the postponement of childbearing?” *Population and Development Review* 30: 195-220.

Sobotka, Tomas. 2004. *Postponement of Childbearing and Low Fertility in Europe*. Amsterdam: Dutch University Press.

Teitelbaum, Michael S. 2004. “The media marketplace for garbled demography,” *Population and Development Review* 30: 317-327.

Toulemon, Laurent and Magali Mazuy. 2001. “Cinq projections de fécondité fondées sur une hypothèse de stabilité de comportements,” *Population* 56: 647-656.

Van de Kaa, Dirk. 1996. “Anchored narratives: the story and findings of half a century of research into the determinants of fertility,” *Population Studies* 50: 389-432.

van Imhoff, Evert 2001. “On the impossibility of inferring cohort fertility measures from period fertility measures.” *Demographic Research* 5(2): 22-64.

van Imhoff, Evert and Nico Keilman. 1999. “On the quantum and tempo of fertility: comment,” *Population and Development Review* 26: 549-553.

Waldorf, Brigitte and Pillsung Byun. 2005. “Meta-analysis of the effect of age structure on fertility,” *Journal of Population Economics* 18: 15-40.

Whelpton, Pascal K. 1954. *Cohort Fertility*. Princeton: Princeton University Press.

Wilson, Chris and Jim Oeppen. 2003. “On reification in demography,” in: J. Fleischhacker, H.A. de Gans and T.K. Burch (eds.) *Population, Projections and Politics*. Amsterdam: Rozenberg, 113-129.

Yang, Quanhe. 1994. “Determinants of the decline in parity progression ratios in China, 1979-1984,” *International Family Planning Perspectives* 20: 100-107.

Zeng, Yi and Kenneth Land. 2001. “A sensitivity analysis of the Bongaarts-Feeney method for adjusting bias in observed period total fertility rates,” *Demography* 38: 17-28.

Zeng, Yi and Kenneth Land. 2002. “Adjusting period tempo changes with an extension of Ryder’s basic translation equation,” *Demography* 39: 269-285.