The Susceptibility of Fertility Measures to Random Fluctuations

The total fertility rate (TFR) is the standard measure used to compare the fertility levels of populations over time and across space. This measure has been used to document the recent rapid fertility decline in developing countries (Mexico, North Africa, for example) and very low TFRs in some European countries have led researchers to explore the phenomenon of "lowest-low" fertility (Kohler and Billari 2002). Arguably, however, the TFR has significant limitations, primarily a sensitivity to an increase (or decrease) in age at childbearing, that may inaccurately present a deflated (or inflated) value (Bongaarts and Feeney 1998). The Bongaarts-Feeney tempo adjustment technique (where the TFR is adjusted based on the change on the mean age at childbearing) has therefore become an important tool in the demographer's toolbox. Use of the Bongaarts-Feeney measure has been largely restricted to countries experiencing below replacement fertility and a notable increase in the age at first birth. Postponement of first births is often the focus of tempo studies as mean age at birth increases generally result from an increasing age at first birth while spacing between births is not assumed to change. Second births may occur at older ages but it is generally assumed that this is as a result of the older ages of first births. Results of Bongaarts-Feeney type analyses suggest that fertility in the "lowest-low" countries may not be as low as the traditional TFR has indicated (Bongaarts and Feeney 1998, Bongaarts 1999, Sobotka 2004). One of the primary assumptions in the application of the Bongaarts-Feeney technique is that women are postponing their fertility at an equal rate across age groups (ie no cohort differences in fertility postponement) (Bongaarts and Feeney 1998, Kohler and Philipov 2001). Conceptually this assumption is problematic as there are likely to be cohort differences in fertility behavior – the expansion of women's employment and increased opportunities for higher education, for example, will likely have a different effect on women in their early twenties than on women nearing the end of their reproductive years (Van Imhoff and Keilman 2000).

In an effort to increase the generalizability of the TFR adjustment technique by taking into consideration age specific variation in fertility postponement, Kohler and Philipov in their 2001 article, developed an alternative strategy for adjustment. The Kohler Philipov technique incorporates variation in postponement – age-period interaction – but reduces to the Bongaarts Feeney approach if the timing changes are assumed to be invariant. Kohler and Philipov argue that while mean age at birth may be increasing it is likely that the variation around the mean age at birth is similarly increasing. Neglecting the impact of increased variation around the mean age of birth leads to bias in the tempo adjusted measure which then produces incorrect measures of period quantum. Conceptually the argument seems valid, but the data requirements and the complexity of the adjustment technique make the Kohler Philipov model comparatively difficult to implement. Furthermore, it is not necessarily clear, however, that the Kohler Philipov measure greatly improves the adjustment provided by Bongaarts Feeney and the mathematical complexity of the technique reduces its widespread usage and easy interpretation.

Each measure, traditional TFR, BF- adjusted TFR and the KP-adjusted TFR are nearly always presented as fixed values (as opposed to random variables where variation is present). However in some cases there are only small differences in the adjusted measures which may be the result of random variation. The purpose of this research is to evaluate the random variation found in the TFR and the mean age at birth, and the resulting variation in the two adjusted measures, BF-TFR and KP-TFR. Natural variation is likely to occur in the mean age at birth and the variation in the mean age at birth which will cause variation in the adjusted TFR measures. Calculating confidence intervals for the measures will help to determine if the adjusted measures are capturing and presenting statistically different fertility behaviors than the traditional TFR and than each other.

Data and Methods

Variances will be simulated[1] using birth counts representing those associated with different analysis situations – survey data, small country, medium sized country, and large country. The beta distribution will be used to represent the distribution of births. Evaluating the standard deviations of different populations contributes to an improved understanding of the usefulness of macro-level fertility measures in several different contexts. Additionally, births of a small country, Sweden (using census data) and those found in survey data (using Guatemala DHS data) will serve as case-studies where the measures and their variances will be calculated and compared.

The analysis will also be expanded to evaluate countries during periods of notable birth fluctuations – Eastern Germany during the 1970s after the introduction of the birth control pill allowed for rapid family size reductions and Taiwan in year 2000, the Dragon Year and the millennium, an optimal year for birth according to traditional Chinese beliefs. Descriptive analysis (not shown here) highlight the increase in mean age at childbearing in both populations - supporting the application of methods that adjust for maternal age changes (ie; Bongaarts-Feeney or Kohler-Philipov adjustments) – however, contextual evidence (Dorbritz 2008, Goodkind 1991, 1993) suggests that real changes in quantum (an decrease in completed family size in Eastern Germany and an increase in period fertility in Taiwan) were occurring at the same time. The results from the analyses of Eastern Germany and Taiwan will be included in the final research paper.

Results

The simulations reveal large differences in the magnitudes of the standard deviations across the different populations. Table 1 presents the simulated values of the measures and of specific components of the measures.

Population Size	Delta Prime (KP)	Standard Deviation Prime (KP)	Mean Prime (KP and BF)	TFR BF	TFR KP
Survey	0.0084	1.3319	0.1029	0.2183	0.4215
Small	0.0009	0.1367	0.01	0.0201	0.034
Medium	0.0003	0.0474	0.0034	0.0068	0.0118
Large	0.0001	0.0204	0.0015	0.0031	0.005

It is not surprising that the KP TFR has a standard deviation that is twice as large as the standard deviation of the BF TFR. Calculation of the KP TFR requires the use of several random variables, standard deviation of the fertility schedule, change of the

[1] For comparison's sake standard errors will also be calculated directly using approximate closed form solutions. The derived formulae are provided in the appendix.

standard deviation of the fertility schedule and the difference in mean age. KP TFR assumes only the change in mean age is random. Also notable is the large standard deviation found in measures calculated from survey data. The margin of error for the TFR BF in the survey data situation is almost half a birth. In some countries where survey data has been used as the only source of birth data to calculate TFR and adjusted TFR such large variability may lead to misinterpretation of fertility change or lead to researchers claiming that fertility has changed when, in fact, there is no significant difference in magnitude.

Case of Sweden:

Sweden provides an example of an analysis situation where tempo adjusted fertility measures have been helpful in examining why fertility rates have changed over time. Figure 1 presents the TFR in addition to the adjusted fertility rates, BF TFR and KP TFR and an alternate version of the KP TFR. After a decreasing fertility rate, the traditional TFR shows a fertility "boom" in the early 1990s which peaked at 1992. A decrease in fertility follows. Coincident with the fertility boom is a leveling off of the mean age at first birth (Figure 2) while during the periods where the TFR is decreasing there is an increase in the mean age at first birth (MAFB). The change in MAFB signals the potential presence of tempo effects indicating the importance of adjusting the TFR for changes in tempo before comments on the actual level of fertility (quantum) can be made.

The BF adjusted measure shows a fertility boom occurring a few years later than the traditional TFR shows. The KP adjusted measure instead "rise[s] and fall[s] almost synchronously" with the traditional TFR leading to the conclusion that the changes that boom was actually a change in quantum rather than a reaction to changes in timing (Kohler and Philipov 2001).

Each adjusted measure is smoothed – meaning that to calculate the measure data from several years before and after the year of interest is averaged. The BF relies on the average of the change in MAFB one year ahead and one year behind the year of interest while the KP is calculated using the average change of *delta* (a measure of change in variation and change in the mean) from two years ahead and two years behind the year of interest. This smoothing may impact the location and magnitude of the apparent boom/bust and may have a substantial impact on the representation of TFR by the adjusted measures. The smoothing discussion is beyond the scope of the paper but is briefly addressed in the appendix and will be further explored in subsequent related analyses.

Figure 1:

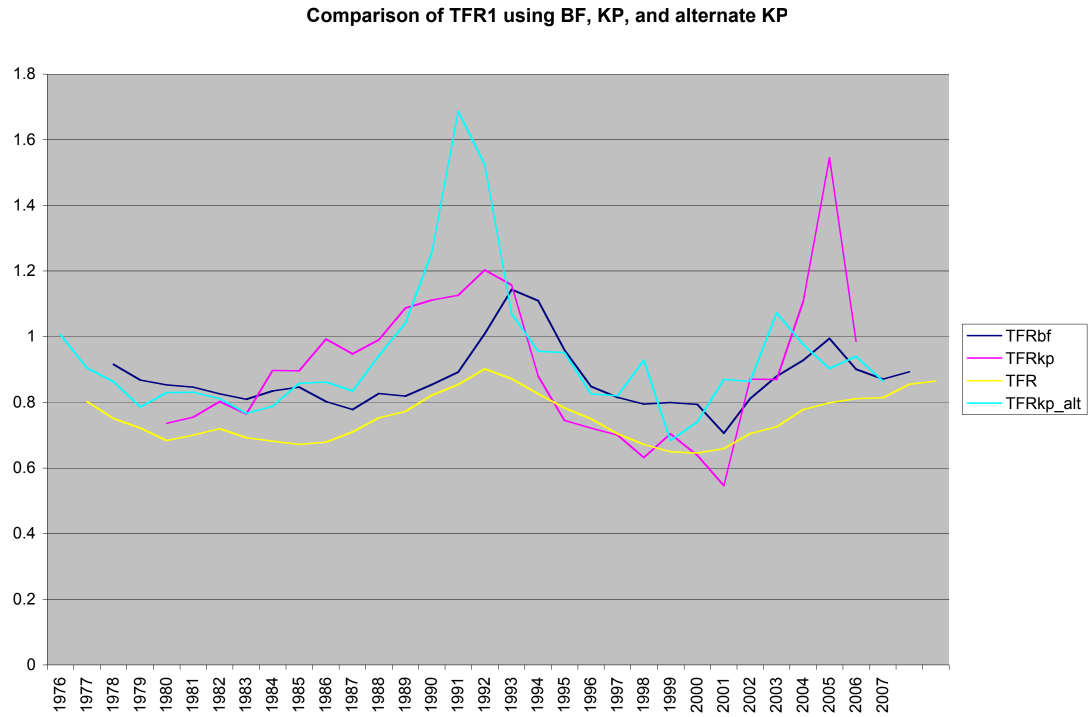


Figure 2:

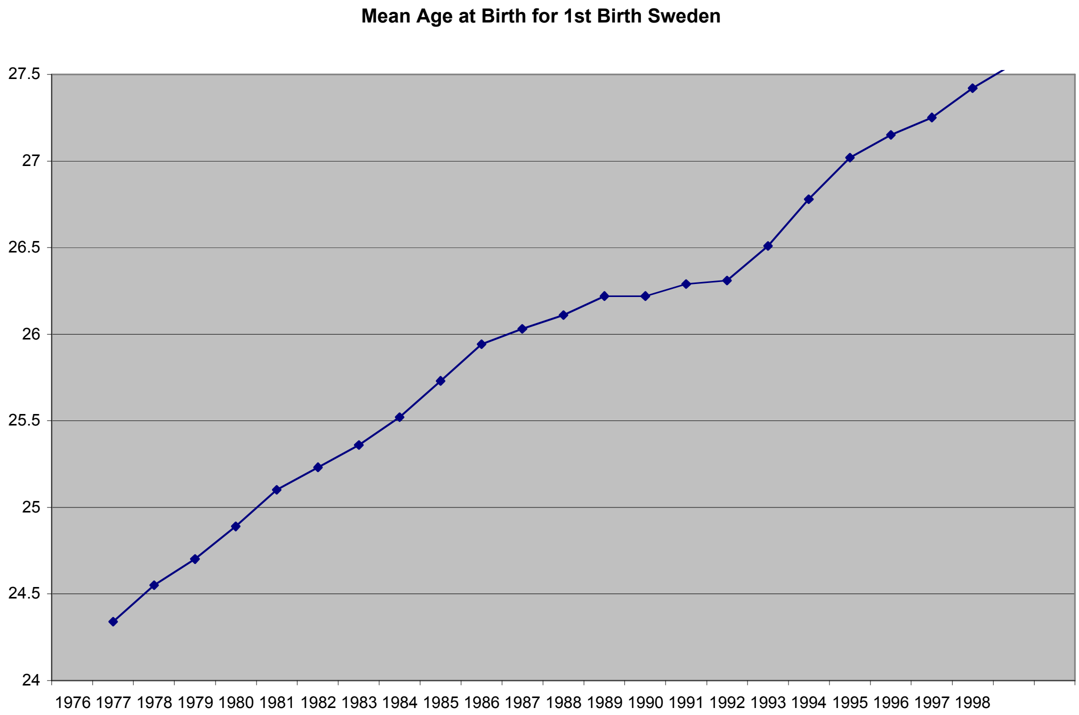
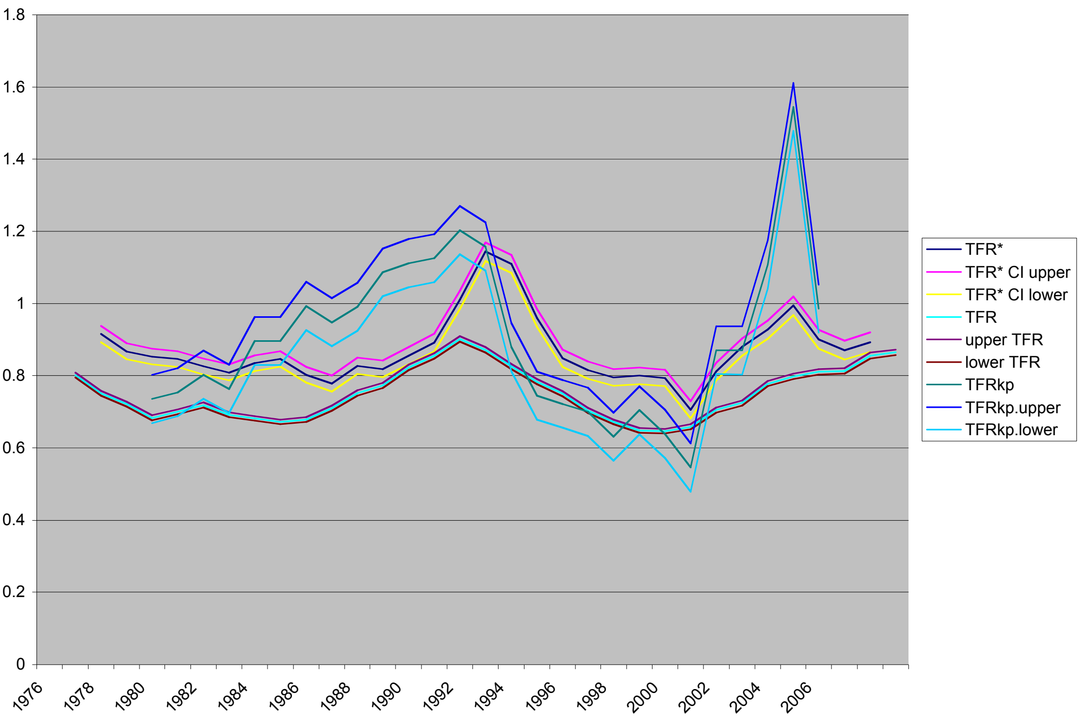


Figure 3:



Incorporating the standard deviations of the measures into the analysis we see that the small peaks and valleys, those of the late 1980s and mid 1990s surrounding the boom, found in each measure of the TFR are not significantly different from each other. However, the large peak – the fertility boom – varies significantly in magnitude and duration (or width of the hump) depending on which type of adjustment is applied. The BF adjusted measure indicates the presence of tempo effects distorted the period TFR while the KP adjusted measure suggests that the boom was primarily a change in quantum.

Case of Guatemala:

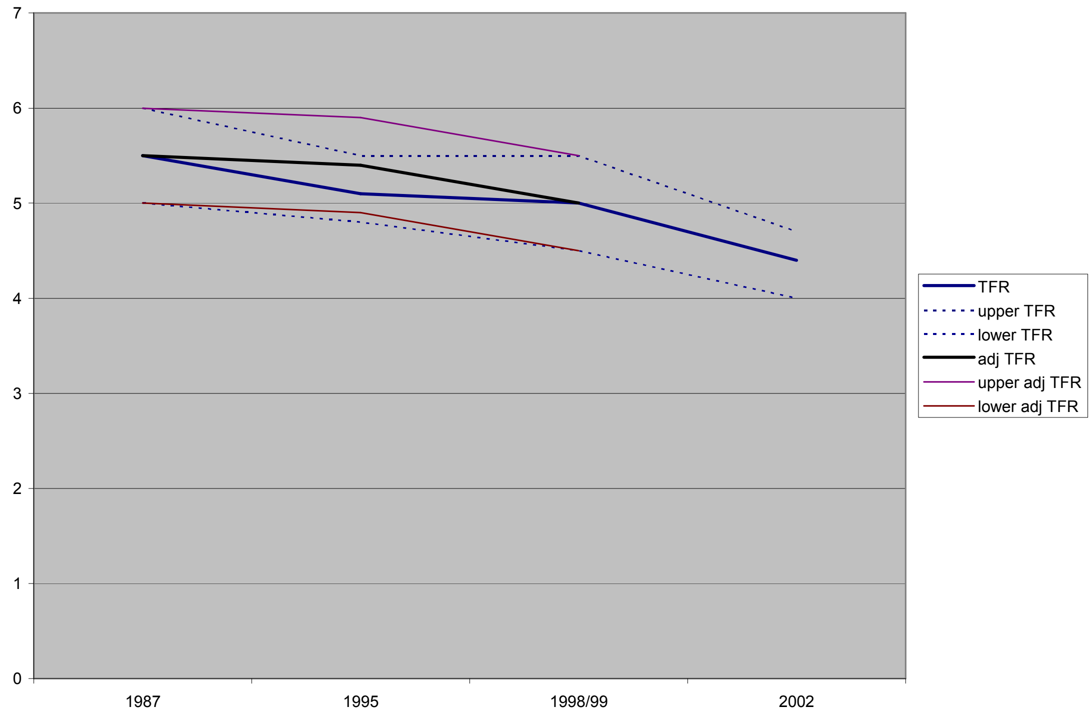
Guatemala provides an alternative case for analysis of fertility measures. Because Guatemala does not collect detailed data on births it is impossible to directly apply the analysis techniques used to analyze Sweden or other populations with detailed birth registries. To analyze fertility in Guatemala researchers rely on survey data, specifically data from Demographic and Health Surveys. TFR is then calculated using 5 year age groups, resulting in seven age categories, and births over the previous three year period are counted to achieve a larger birth count. However, because of the severe data limitations it is not possible to calculate a measure as complex as the KP TFR primarily because an accurate measure of the variation in the fertility schedule is problematic to calculate from data that is grouped by age. BF TFR can be calculated and may actually provide insight into instances where it appears as though the fertility decline has stalled[2].

Figure 4 presents the TFR and BF TFR over the past 20 years in Guatemala. The traditional TFR suggests that a decline in fertility began in the 1980s followed by

---

[2] While low and decreasing fertility is of concern to researchers and policymakers focused on developed countries, those who are interested in developing country demographics focus instead on stalled fertility transitions.

a stall during the early 1990s. The adjusted measure, however, suggests that there was no stall but instead the fertility decline began at a later date.



After including the confidence intervals for both the TFR and BF TFR however, we see that not only is there no significant difference between the adjusted and the actual measure but we also see that there is little difference over time. It is therefore difficult to determine if demographic transition has actually begun.

Discussion and Conclusion

The purpose of this analysis was to evaluate random variation in the traditional TFR and two popular tempo-adjustment techniques, Kohler-Philipov and Bongaarts-Feeney. Random variation exists in macro-level measures of fertility but is not often considered when different measures are evaluated. However, incorporating confidence intervals into a fertility analysis does not necessarily enhance the analysis. Including confidence intervals may be an important step in the analysis of survey data however even in a country with as few births as Sweden, including confidence intervals for the traditional and adjusted TFR did not provide much additional information.

Results from the Sweden analysis suggest that the smaller peaks and valleys found in non-smoothed fertility measures are likely due to random variation. The large peaks and valleys, however are indicative of significant changes to the level of fertility. The Guatemalan analysis, on the other hand, suggests that nearly all reported fertility change over the past 15 year period is possibly due to random variation. Given the important political, economic and social changes that have occurred in Guatemala during the same time period it is more likely that significant changes in fertility rates actually have occurred but the crudeness of these macro-level measures combined with limited survey data inhibits the true fertility behavior from being discovered. Honing macro-level measures to accommodate the special cases where survey data is the only available data would likely provide important insight into the fertility change of many developing countries and may also improve our understanding of the measures themselves. Furthermore, expanding the analysis to include populations/time periods where known quantum changes occurred, like Eastern Germany after the expansion of the birth control pill and Taiwan at the Dragon Year, will permit more detailed examination of the measures.

References

Bongaarts, J. and G. Feeney. 1998. "On the Quantum and Tempo of Fertility." *Population and Development Review*. 24(2): 271-291.

Bongaarts, J. 1999. "The Fertility Impact of Changes in the Timing of Childbearing in the Developing World." *Population Studies.* 53(3): 277-289

Dorbritz, J. 2008. "Germany: Family diversity with low actual and desired fertility." *Demographic Research* 19(17): 557-598.

Goodkind, D. M. 1991. "Creating new traditions in modern Chinese populations: aiming for birth in the Year of the Dragon." *Population and Development Review* 17:663-686.

Goodkind, D. M. 1993. "New zodiacal influences on Chinese family formation: Taiwan, 1976." *Demography* 30:127-142.

Kohler, H. P. and F. C. Billari. 2002. "The emergence of lowest-low fertility in Europe during the 1990s." *Population and Development Review* 28:641-680.

Kohler, H. P. and D. Philipov. 2001. "Variance effects in the Bongaarts-Feeney formula." *Demography* 38:1-16.

Sobotka, T. 2004. "Is lowest-low fertility in Europe explained by the postponement of childbearing?" *Population and Development Review* 30:195-220.

Van Imhoff, E. and N. Keilman. 2000. "On the quantum and tempo of fertility: Comment." *Population and Development Review* 26:549-553.

Appendix

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.