The standardized TFR

Jan M. Hoem

Stockholm University Demography Unit
and Max Planck Institute for Demographic Research, Rostock

and

Cornelia Mureşan

Babes-Bolyai University, Cluj-Napoca

Abstract: The period-based Total Fertility Rate is probably the most commonly used single measure of a population's fertility level, but it regularly meets with disdain from professional demographers, mainly because it only controls for the population's age distribution and not for any other subdividing feature, such as the parity distribution, ethnic composition, or educational attainment. In the present contribution we show how the usual TFR can be standardized for the population's distribution across any selected subdivision. We use the data of the Romanian Gender and Generations Survey to illustrate how this can be done, and in the process we illuminate how even standardized TFRs computed from a data set of this smallish size order ($n \approx 6000$) is a blunt tool that cannot easily be used to reveal an important feature of demographic trends such as the postponement of fertility in recent decades. The replacement of the convenient single measure (TFR) also comes at the cost of the introduction of a plethora of other summary measures.

1. Introduction

Professional demographers typically are less than enthusiastic about the use of the period-based Total Fertility Rate (TFR) as a measure of a population's fertility level, mainly because it controls the data for the population’s age distribution but not for any other sub-division of the population, such as parity, ethnic group, educational attainment, or geographical location. In this paper we show that this particular criticism can be overcome through a new approach to event-history analysis. By way of illustration we produce period TFR values based on the data for women in the first panel round of the Romanian Gender and Generations Survey (GGS), standardized for parity, educational level, and the rural/urban character of the woman’s place of birth. The method can be extended directly to cover any other distributional dimension. Thus the most direct methodological criticism of the TFR can be discarded: it *is* possible to control for other factors than age attained.

This suggests that the TFR may have a potential in fertility analysis similar to that of a standardized rate in mortality analysis. Nevertheless, we need to raise the issue whether computing a standardized TFR provides a sufficiently efficient use of the data, or whether a different method can give better value for money. We show that some questions are better answered by other means when the data come from a semi-large sample like that of the typical first-round GGS, which at best has some 6000 female respondents. In particular we find that the approach that produces the standardized TFR will not really establish whether there has been a systematic postponement of childbearing in Romania since the fall of communism. This is not because the method cannot be used for this purpose in principle, but because a very real postponement is drowned in random variation. To establish the childbearing postponement we must either use a much bigger data set (as is available in official statistics) or apply a more efficient different method (which in fact can also be developed from event-history analysis). In this manner, the criticism of the TFR is shifted from a complaint about the lack of general standardization to the lack of efficiency in the use of the information actually available in the GGS data.

To establish periods of fertility postponement is essential to a concern with the underestimation of population fertility by the TFR during such periods, a concern which has received a lot of attention in recent years and which has led to attempts to correct for tempo effects (Bongaarts and Feeney, 1998; Sobotka, 2004, Chapter 4; and their successors, including Goldstein et al., 2009).

2. Childbearing postponement revealed in official statistics

Romanian official statistics provide age-specific fertility rates for single-year age groups for each calendar year. Mureşan et al. (2008) have used such data to produce annual TFRs and mean ages at childbearing (their Figures 2a and 2b, redrawn in our Figure 1) and annual age profiles of fertility (their Figure 5a, redrawn in our Figure 2). According to these diagrams, childbearing was postponed in Romania progressively since 1994 (but not during the preceding years). The TFR declined between 1987 and 1996, and then subsequently remained largely constant. Note that the periods of fertility postponement and of TFR decline do not coincide. Such data are available for many countries.

Figure 1. Total fertility rate and mean age at childbearing. Romania 1960-2007

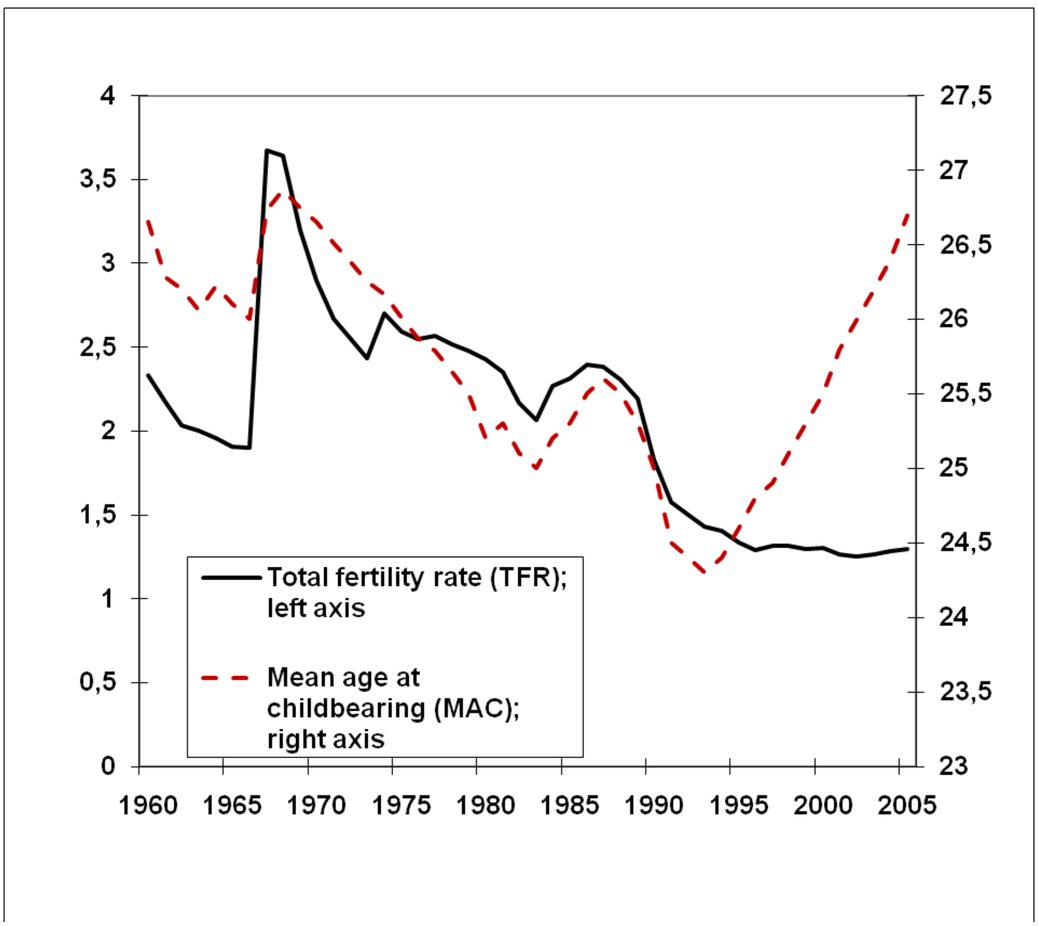
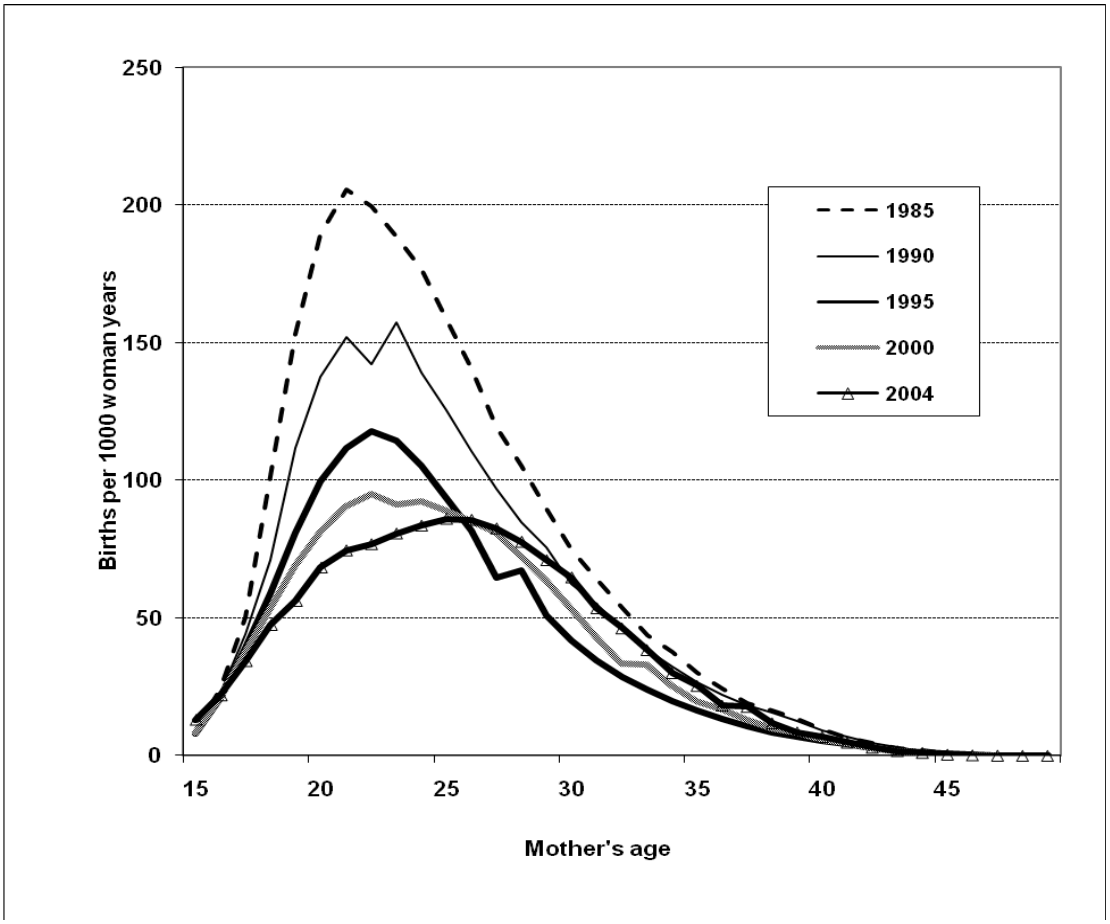


Figure 2. Age-specific fertility rates. Romania 1985-2004



3. Standardizing the annual TFR

Let us remind ourselves that in the most straightforward applications of event-history analysis, each individual is followed from some starting point (process time 0) and until some event of interest occurs or until the event-history is censored. For example, process time 0 may be (nine months after) the time of the respondent's first birth, a later process time may be time elapsed since (nine months after) first birth, the event in focus may be the occurrence of a second birth, and the waiting time may last until any second birth occurs or until censoring, whatever comes first. In this standard formulation, there is only one waiting time, at the end of which a single occurrence appears, unless the wait ends with censoring, in which case there is no event occurrence for the individual. Beside the process time since first birth more clocks may be in operation, for instance (i) the respondent's age attained (or in practice time since the respondent's 12th birthday, say) and (ii) real calendar time. In multi-process formulations, there may be additional clocks, such as duration of marriage (or of a consensual union), duration since entry into the labor market, and so on. It is part of the analyst's task to choose which clock to use to represent process time and to incorporate it into the analysis along with a specification of fixed and time-varying covariates.

Event-history analysis is not restricted to a set-up with a single waiting time that possibly ends with a single occurrence, however. To approach the philosophy behind the TFR as this quantity is usually computed, we follow a respondent through life from some initial age (say 12 years) instead and record one birth after another to the extent that births appear in the respondent's life history, until censoring possibly occurs or the individual reaches the end of childbearing. At each birth, one records an occurrence with multiplicity 1 for single births, multiplicity 2 for twin births, and so on. Then the usual methods of event-history analysis can be applied with all their elements, including possibly interactions between covariates. If the software is geared to the single-waiting-time situation, it can be used for a process specification with consecutive occurrences of the event in focus over each individual-level lifetime. All one needs to do is to break up the individual life history into segments between the appearances of the event and let each segment after the first be left-censored at the beginning of the segment. Since the computation of the usual TFR never takes pregnancy duration into account, we allow ourselves the license of carrying out the analysis as if the individual is exposed to the risk of the next birth immediately after the occurrence of one event, in the manner of a Poisson process. If we include the possibility of multiple events, what will appear from such an analysis is an age-dependent birth rate, estimated over the age range defined by the observational design.

In our application to childbearing in Romania, we have specified a piecewise constant childbearing intensity, which in the simple form that we have used appears as $\varphi_{xtgjk} = a_{xt} b_g d_j e_k$, with factors *A*, *B*, *C*, *D*, and *E* and corresponding parameters a_{xt}, b_g, d_j, and e_k, respectively. [1]

Factor *A* stands for <u>age attained</u> and is indexed by *x*, as it often is in demography.

Factor *B* stands for <u>parity</u> and is indexed by *g*.

Factor *C* stands for <u>calendar time</u> and is indexed by *t*.

---

[1] Factors and their parameters can be taken to have mnemotechnical names if we think of *B* as Birth order, *C* as Calendar period, *D* as the character of the District of birth, and *E* as Educational level.

Factor *D* stands for the rural or urban character of the place of birth and is indexed by *j*. Factor *E* represents educational attainment and is indexed by *k*.

The regression parameters in the formula for φ_{xtgjk} are as follows:

The parameter a_{xt} is the effect (on the birth rate) of the AC combination, i.e., the (multiplicative) effect on the childbearing intensity of being in age group *x* in calendar period *t;* note that Factors *A* and *C* appear in interaction in our application. We have experimented with various types of age groups and periods, but in our main specification we use single-year age groups and single-year periods. To reduce the effect of random variation, we use longer periods and longer age groups occasionally, as indicated below.

The parameter b_g is the effect of having parity *g*. We use parities 0, 1, 2, 3, 4, and 5+ and let b_0 =1 to normalize the b parameters. This makes each b_g a relative risk.

The parameter d_j is the effect of being born in a birth district of type *j.* In the present specification it is the effect of coming from a rural or urban birthplace. We set d_{rural} =1.

The parameter e_k is the effect of having educational level *k*. We use three levels of educational attainment (low, middle, and high) plus the level "in education", as discussed by Mureşan and Hoem (2010, Appendix 1). We set e_{low} =1.

We have estimated the parameters of the birth intensity φ_{xtgjk} from the Romanian GGS data by the maximum-likelihood method, and have computed standardized estimates $TFR(t) = \sum_x a_{xt}$ for the TFR in each year *t* and the corresponding standardized mean age $\sum_x x a_{xt} / TFR(t)$ at childbearing. Figure 3 contains a plot of *TFR(t)* and also a plot of the annual sum of age-specific birth rates, namely the "un-standardized" TFR. (This stippled curve refers to the right-hand *y*-axis in the diagram.) Note how a general decline in the "un-standardized" TFR has been modified in the standardized curve, which drops a bit less strongly than the un-standardized curve does. Evidently the standardization provides some correction. We explain why below, and also explain why the un-standardized curve is systematically lower than the standardized curve, which motivates the use of two different *y*-axes for the two curves.[2]

Figure 3. Standardized and non-standardized TFR. Romania 1966-2003

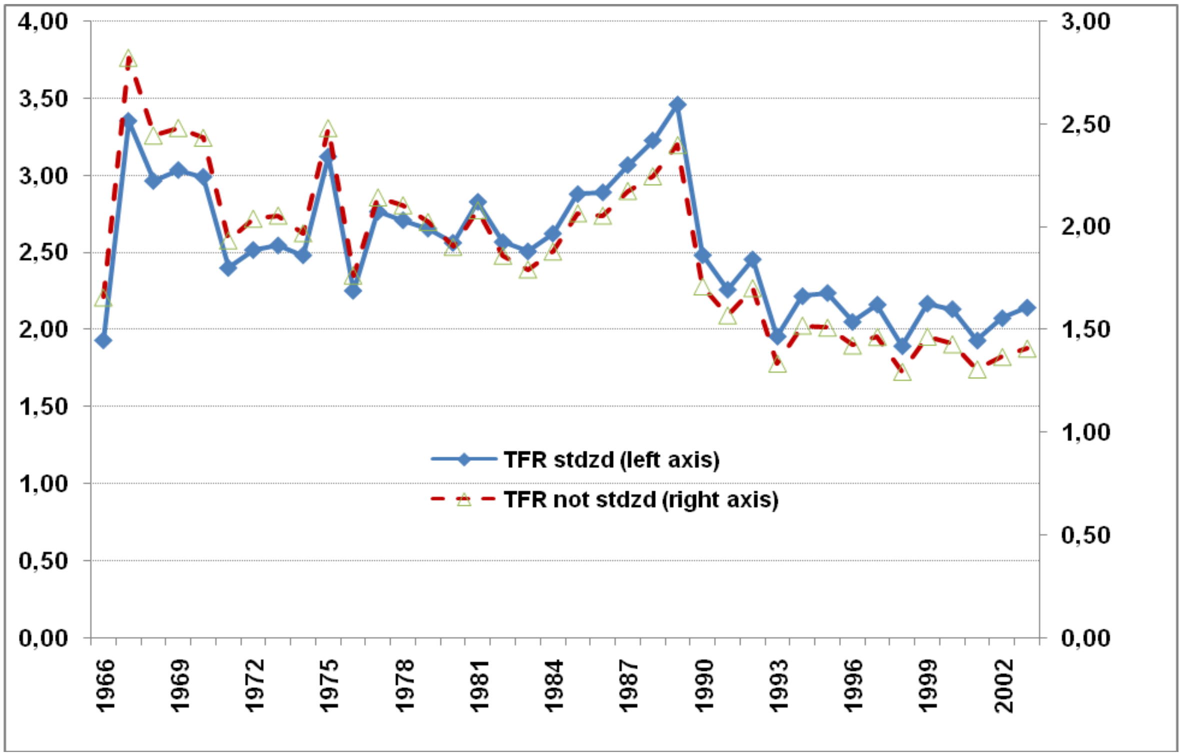
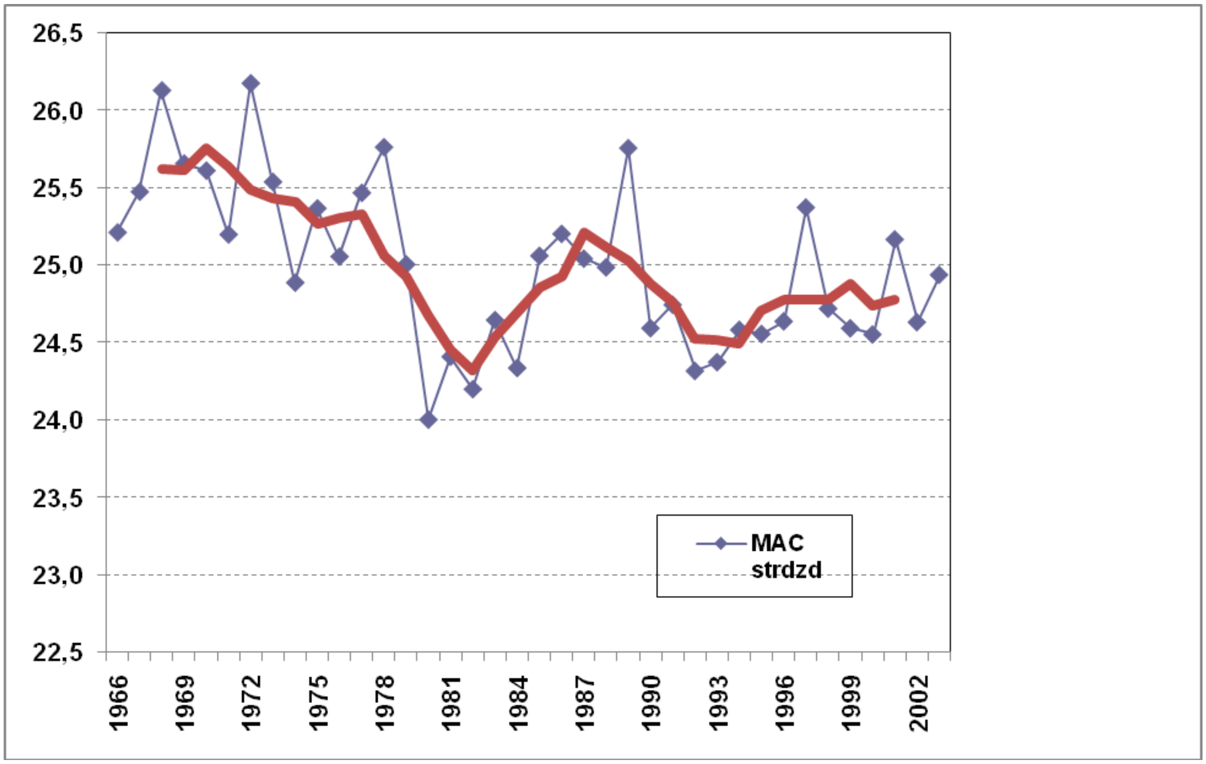


Figure 4 displays our estimate of the standardized annual mean age at childbearing (*MAC(t)*). To reduce the effect of random variation, we have smoothed this curve using a five-item un-weighted centered moving average and have plotted the outcome as the thick unmarked curve in the diagram. (The smoothed value for year *t* is computed as $\sum_{s=-2}^{2} MAC(t+s)/5$.) The MAC values rise a little over the most recent decade but nowhere near as clearly as the corresponding values in Figure 1.

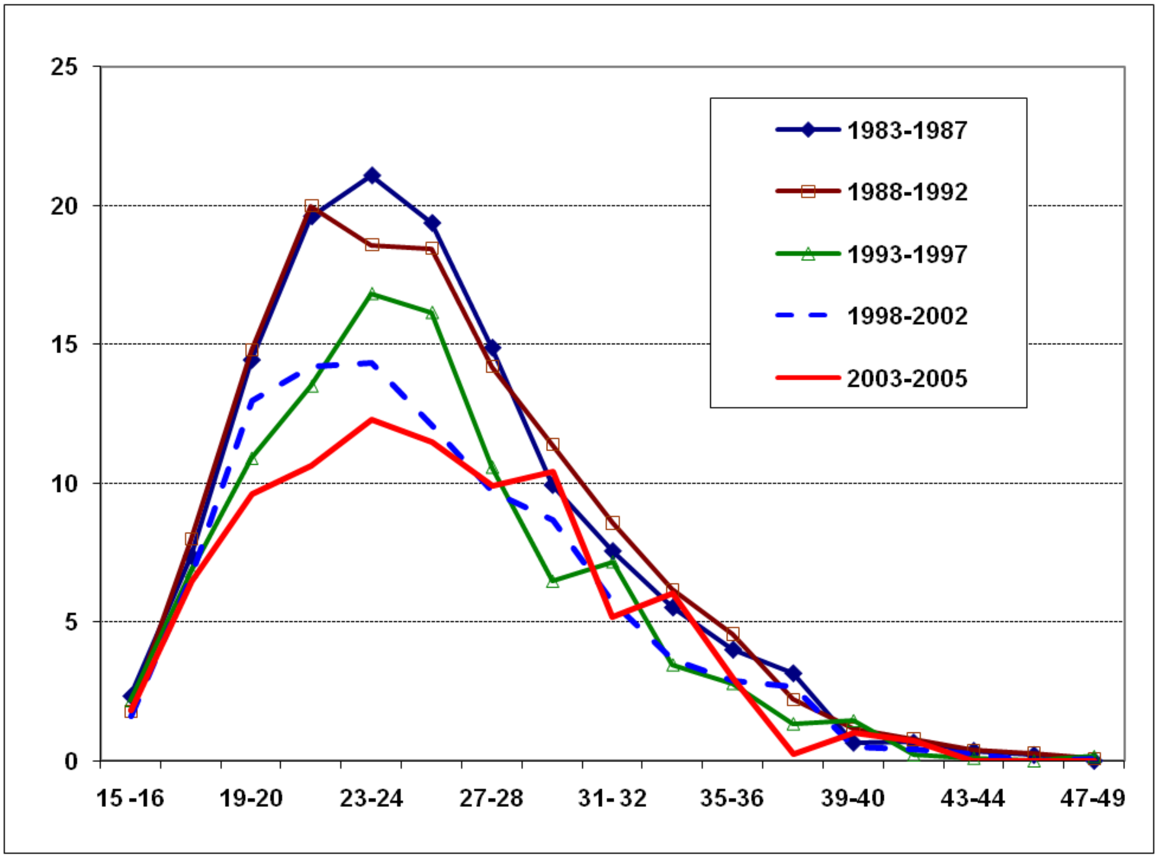
[2] The un-standardized TFR values drop from around 85 per cent of the standardized values in the late 1960s to some two-thirds of the latter in the last decade of our observations.

Figure 4. Standardized mean age at childbearing. Romania 1966-2003



To get age profiles similar to those of Figure 2, we need to counteract strong random variation in the survey data and have combined the single calendar years into longer periods and single-year ages into two-year age groups; see Figure 5. With some good will one can recognize patterns similar to those of our Figure 2 (in particular the drop in fertility is clearly evident in Figure 5 as well), but we doubt that anyone would be induced to suggest a postponement of childbearing after the fall of state socialism on the basis of these new diagrams alone. The shift to the right of the later age profiles in Figure 2 is too subtle a development to be picked up by a blunt tool like the standardized age profiles of Figure 5.

Figure 5. Age profiles of fertility. Romania 1983-2005. Grouped calendar periods and age years. Standardized rates per 1000 person-months



Our method of analysis also provides estimates of the relative risks for covariates *B, D*, and *E* along with the standardized values of the annual TFRs etc., as in Table 1, whose values we can see as summary values of the effects of parity, educational attainment, and character of birthplace in the data population. To us the most remarkable feature of this table is the pattern of relative risks for the parity variable. We see that its relative risk *increases* with parity after Parity 2, and that it becomes very large for Parities 5 and above. There may be an element of selectivity in the latter feature, in that only highly fertile women can be expected to go beyond Parity 4; note that the group with Parity 5 or more includes women who during the pronatalist communist period 1954-1989 were rewarded with honors and favorable lifetime pensions for their childbearing feats (Kligman 1988).[3] The Romanian 2005 GGS sample shows that women with 5 or more children more often than others belong to neo-protestant religions (12.1% as compared to 3.3% of the whole population) or to the Roma ethnic group (10.7% as compared to only 1.5%). Andersson (2007-08) has found similar effects of selectivity in Swedish data. Since there must have been rather few women in this group, even in Romania, it is also likely that random variation has influenced the outcome.

Table 1. Relative risks for our control variables. Romania 1966-2003.

Parity		Education		Birthplace	
0	1	Low	1	Urban	1
1	0.86	Mid	0.72	Rural	1.19
2	0.45	High	0.73		
3	0.74	In ed.	0.32		
4	0.78				
5+	1.65				

This table of relative risks provides the key to the reason why the curve for the annual non-standardized values lies systematically below that of the standardized TFR in our Figure 3. (Remember that the two curves refer to different ordinal axes. Cf. Footnote 3.) The annual *standardized* TFR is the total fertility rate for the group of women who have the relative risk of 1 on all control variables. In the current application this is the group of nulliparous Romanian women with a low educational attainment and a rural birthplace. Its *non-standardized* counterpart is the total fertility rate of an average Romanian woman, for whom no account is taken of parity, education, and character of birthplace. The relative risks that are not taken into account will therefore be for a woman with parity above 0 and with some non-elementary

[3] Women who had at least five children were awarded a medal, women with at least seven children were called "glorious mothers", and women with ten or more children became "heroine mothers". After about the year 2000 the laws were abrogated.

educational attainment.[4] Because of the pattern of Table 1, the non-standardized TFR will therefore be at a lower level than the corresponding standardized TFR in each year.[5]

The smaller gradient of the standardized curve in Figure 3 (as compared with the un-standardized curve) is a compositional effect caused by the fact that the distribution over educational categories changes between the early and the more recent years of the diagram as more and more women spent time in education and improved their educational attainment. In principle there could be a similar effect of a changing distribution over parities, but that does not seem to have been an important feature, as we indicate in Section 5 below.

The theory that we have sketched provides a solution to some issues that appear in applications of a measure like the TFR and quantities derived from it. For instance our intensity formula $\varphi_{xtgjk} = a_{xt} b_g d_j e_k$ provides an annual standardized and *parity-specific* TFR given as $b_g \sum_x a_{xt}$, for g = 0, 1, 2, and so on, and this would provide a sensible alternative to the order-specific TFRs computed from fertility rates of the second kind, used by Mureşan et al. (2008, Figures 3a, 5b, 5c, and 5d) and by many other authors. Unfortunately, this neat theory has the problem that unless one can impose some hierarchy over the levels of the factors *B*, *D*, and *E*, there is no natural a-priori choice (g_o, j_0, k_0) of baseline group for all covariates. A change of baseline levels for the control variables would scale the annual standardized TFRs up or down correspondingly; in principle each choice (g_o, j_0, k_0) of baseline categories for *B*, *D*, and *E* will give a numerically different annual set of standardized TFRs. If for instance we had selected women at parity 2 with a middle level of education (and born in an urban district) as our baseline, then for each year the corresponding standardized TFR would have been about one third as large as in Figure 3 (since 0.45*0.72 = 0.324). Altogether we could get 48 (=6*4*2) different curves of annual standardized TFRs, corresponding to the levels on the factors in Table 1, and essentially none of them would be ruled out or preferred on a-priori grounds. Each set could be compared to the un-standardized annual TFRs in a diagram like Figure 3, simply by

[4] The birthplace differential is too small to have much influence in this connection.

[5] It so happens that the general relation between the two annual versions of the TFR is as 4 to 3 in the present case (as is seen by the relationship between the two ordinal axes), but this must be accidental and we cannot see that it has any particular significance.

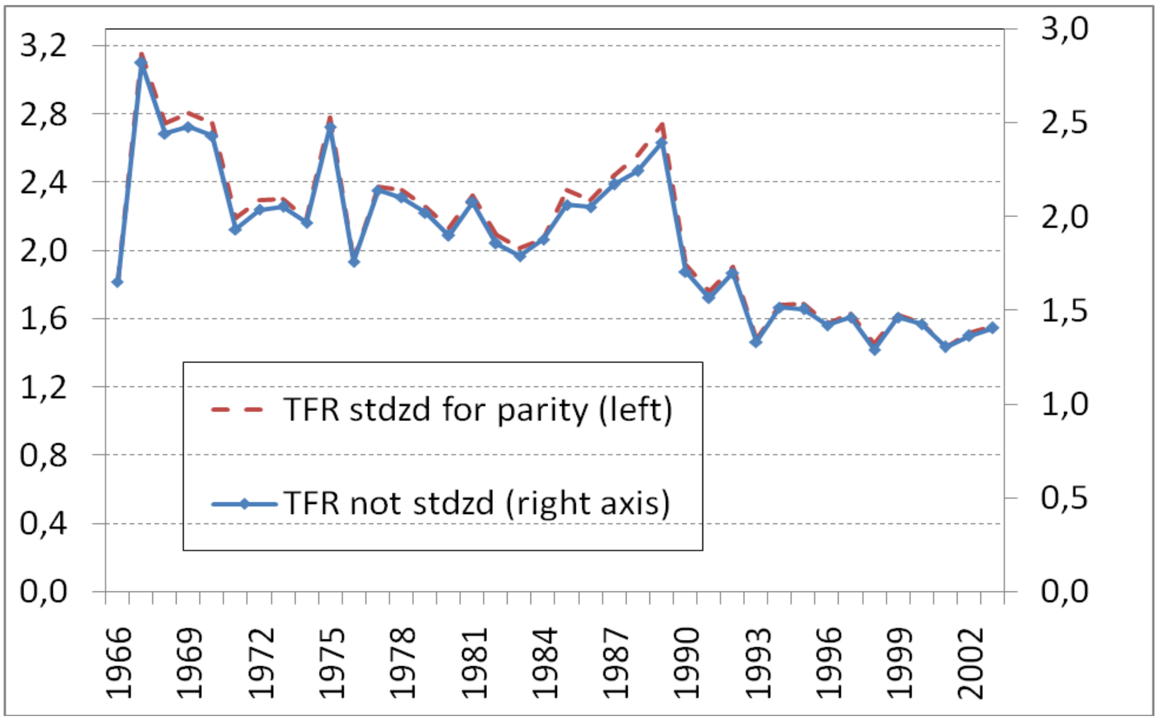
rescaling the right-hand *y*-axis, but it would not be very satisfactory to have to consider so many possibilities as competitors to the simple choice of the classical TFR. (One may possibly reduce this proliferation of standardizing possibilities by anchoring the parity factor at Parity 0 and leaving out other parities as useful baselines, but there would still be eight baseline possibilities due to the choices of baselines for educational attainment and region of birth.)

Another type of worry is that our use of the formula $\varphi_{xtgjk} = a_{xt} b_g d_j e_k$ for the childbearing intensity induces some limitations to how realistically our model can represent real-life childbearing processes, quite apart from the fact that we have disregarded a real loss of exposure to the "risk" of childbearing during periods of pregnancy (which is a natural feature of any TFR). For instance we have proceeded as if fertility changes across parities as specified by the parameters $\{b_g\}$. Thus on the arrival of the *g*-th birth, according to the model the childbearing intensity changes by a factor b_g / b_{g-1} at all ages (and also at all educational levels and both types of locations). Our computations are made *as if* the age profile of the intensity is the same (and is given by the *x*-pattern of the $\{a_{xt}\}$) for all parities in each year *t*. More realistically, the different parities should have different age patterns, however. Fortunately the rigid assumptions inherent in our formula for φ can be loosened up a bit by the use of interactions. For instance an interaction between age and parity would allow for a parity-specific age profile of fertility. To include this interaction, we drop the parameters b_g and replace the $\{a_{xt}\}$ with another set of parameters, say $\{a_{xtg}\}$. For each parity *g*, adding up over *x* gives a parity-specific standardized total fertility rate $TFR_g(t) = \sum_x a_{xtg}$, which (when sub-specification of empirical exposures by parity is possible) again provides an alternative to the usual quantities based on rates of the second kind.

Apart from the problem of riches produced by the indeterminacy in the choice of baseline levels for Factors *D*, *E*, and possibly *B* in the intensity model, we also need to give some consideration to the possibility of making a different choice of control variables. While parity (Factor *B*) is a natural ingredient in any *description* of fertility, Factors *D* and *E* are features that would naturally be included in an *explanation* of fertility patterns and trends instead. When we use educational attainment and type of region of birth as control variables, we therefore produce a hybrid TFR that incorporates both explananda and explanans of childbearing behavior.

This problem and most of those mentioned in this discussion will disappear if we limit the variables on which we standardize to the parity factor (*B*) and settle for Parity 0 as its baseline level. When we then fit a birth-intensity model $\varphi_{xtg} = a_{xt} b_g$ to the Romanian GGS data, we get the time series of parity-standardized TFRs that appear in Figure 6 along with a plot of the un-standardized (classical) TFRs that are estimated from the same data.[6] The two curves essentially coincide, which tells us that there are no compositional effects of the kind we displayed in Figure 3. Thus the distribution of the population across parities does not change appreciably over the years included in the diagram.[7]

Figure 6. TFR, non-standardized and standardized with respect to parity only. Romania 1966-2003



[6] The left-hand *y*-axis in Figure 6 runs from 0 at the bottom to 3.4 at the top of the axis.

[7] Plotting the parity-standardized mean age at childbearing (not shown here) does not produce any more convincing evidence of a postponement of childbearing than Figure 4 does.

We could pursue this reasoning further and discuss more features of the intensity model φ and of Table 1, but it is more productive to turn to a more realistic representation of childbearing behavior, as follows.

4. Separate analysis by parity

So far we have mostly used a Poisson-process type of event-history analysis to study childbearing behavior, using age attained as our process time and letting a woman's parity appear merely as a "control" variable. An alternative and actually more common type of procedure would be to study each birth order separately and to combine the results to get an overview of fertility developments across the various parities. Demographers use this approach all the time, indeed it has been used for our current data by Mureşan and Hoem (2010). As we now proceed to show, it turns out that some of their findings more explicitly establish the systematic postponement of childbearing in Romania since the fall of state socialism, a feature which is essential for an understanding of the adequacy of the usual period TFR as a measure of population fertility. Our present account of their findings will concentrate on fertility postponement.

Translating their symbols so they fit better with our account, Mureşan and Hoem (2010) specify the following model for the logarithm of the first-birth intensity $h_{i1}(x)$ for individual *i* at age *x*:

$$\ln h_{i1}(x) = y_1(x) + \sum_k \beta_{k1} w_{ik1}(x) + z_{c1}(c_{i0} + x - 1950) .$$

The log-baseline intensity $y_1(x)$ is a linear spline function of *x* and the $\{w_{ik1}(x) : k = 1, 2, \ldots\}$ are possibly age-dependent determinants of first births. Mureşan and Hoem (2010) concentrate on the effect of educational attainment and let $w_{ik1}(x)$ be an indicator of whether a female respondent has reached educational level *k* at age *x* (including whether she is deemed to be under education at that process time). Then $\exp(\beta_{k1})$ is the relative risk of occurrence of a first birth for educational category *k* (with β_{k1} =0 for *k*=1, say). Finally, a second linear spline z_{c1} is supposed to catch the effect of calendar time (counted since the beginning of the year 1950) on first-birth "risks". The argument c_{i0} is the calendar time at which we take individual *i* to start to be exposed to the risk of a first birth, i.e., the calendar month in which she turns 12. Thus $c_{i0} + x - 1950$ is the calendar month in which the respondent is 12+*x* years old, counted from the beginning of 1950. The authors could easily have added further covariates as determinants of the birth intensity.

For second births they use a specification

$$\ln h_{i2}(s) = y_2(s) + \sum_k \beta_{k2} w_{ik2}(s) + z_{c2}(c_{i1} + s - 1950) + z_{a2}(x_{i1} - 18),$$

where z_{a2} is another linear spline used to pick up the effect of the age x_{i1} at first birth (counted from age 18) and the other items are quite similar to those of the intensity of a first birth. Process time *s* is now months since first birth, and c_{i1} is the calendar month of the first birth. For third births their intensity specification is quite similar to that for second births. Note that this includes a much more explicit specification (through the spline functions {$z_{c\ell}$; ℓ =1,2,3}) of the effect of calendar time than what we had in Section 3 above. The authors selected the nodes of the three splines {$z_{c\ell}$} so as to pick up major changes in family policies in Romania.

Figure 7 contains estimates of the three order-specific calendar splines. We see that the first-birth rate has declined rather steadily since 1990, and that second- and third-birth rates have declined even more and consistently since mid-1985.[8] This analysis establishes without any doubt that Romanian fertility has been declining since the fall of state-socialism, or even earlier.

[8] For the definition of this node, see the end of Section 3 in Mureşan and Hoem (2010).

Figure 7. First-, second-, and third-birth intensities. Duration splines by calendar year

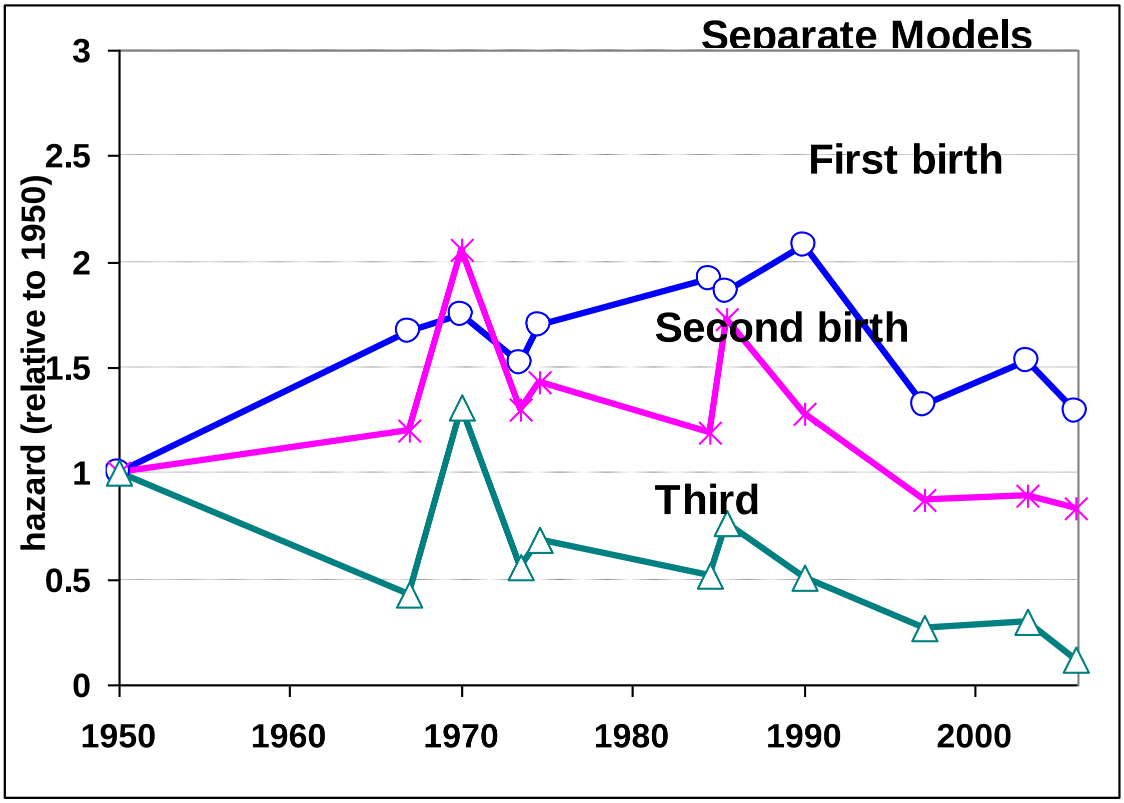


Figure footnote: Source: Mureşan and Hoem (2010, Appendix 2b).

As we have seen above, the model used for this analysis of childbearing behavior separately for each parity represents a more accurate specification than the piecewise-constant model in Section 3, and the ensuing results are sharper in the details they cover than the more conventional results of Mureşan et al. (2008) are. On the other hand, the latter analysis gives a better impression of general features of the postponement, like the change in age profile and the mean age at childbearing for all parities taken together. It is much more difficult to produce such features from the parity-specific analysis in the present section, which is geared toward different issues. The analysis made possible by our account of the standardized TFR in Section 3 has the same potential as the one based on official statistics by Mureşan et al. (2008), but a use of that potential seems to require data sets considerably bigger than that of the female respondents in a normal Gender and Generations Survey.

5. Discussion

The role of a TFR is to act as a measure of a population's fertility level in a given period and to allow for a study of trends in that level over time. The general purpose of standardization is to reveal the existence or absence of compositional effects and to neutralize them if present. Our discussion after Table 1 showed that the standardization of the TFR with respect to a subdivision of the population according to criteria other than age provides summary indicators of fertility effects according to the subdividing features, which may be convenient for some purposes. The standardization can also serve its usual purpose (Figures 3 and 6). With the possible exception of a subdivision according to current parity, standardization does not immediately provide an alternative to the TFR, however. The computations involved also produce quantities that may be moderately useful for an improved understanding of some features of fertility, such as a standardized mean age at childbearing and standardized period age profiles of childbearing rates (compare our Figures 4 and 5), but unless the data set is larger than in the first round of the current Gender and Generations surveys, the tools are too blunt to pick up subtler trends and patterns. If providing summary measures has lower priority than accurate description and substantive explanation, then other tools are more effective, as is seen in our Figures 1 and 2 and in Section 4.

Acknowledgement

We are grateful for perceptive comments from participants in a colloquium at the Stockholm University Demography Unit in February, 2010.

References

Andersson, Gunnar (2007-08). Selectivity of higher-order childbearing in Sweden. *Finnish Yearbook of Population Research* 43, 33-40.

Bongaarts, John and Griffith Feeney (1998). On the quantum and tempo of fertility. *Population and Development Review* 24 (2), 271-291.

Goldstein, Joshua R., Tomáš Sobotka, and Aiva Jasilioniene (2009). The end of 'lowest-low' fertility? Max Planck Institute for Demographic Research, Working Paper 2009-029. Also more briefly in *Population and Development Review* 35 (4), 663-699.

Kligman, Gail (1988). The Politics of Duplicity: Controlling Reproduction in Ceauşescu's Romania. Berkeley: University of California Press

Mureşan, Cornelia; Paul-Teodor Hărăguş, Mihaela Hărăguş, and Christin Schröder (2008). Romania: Childbearing metamorphosis within a changing context. *Demographic Research* 19 (23), 855-905.

Mureşan, Cornelia and Jan M. Hoem (2010). The negative educational gradients in Romanian fertility. *Demographic Research* 22 (4), 95-114.

Sobotka, Tomáš (2004). *Postponement of Childbearing and Low Fertility in Europe*. Dutch University Press, Amsterdam.