The Own Children Fertility Estimation Procedure: A Reappraisal

Christopher Avery
Harvard Kennedy School of Government
Travis St. Clair
Harvard Kennedy School of Government
Michael Levin
Harvard Center for Population and Development Studies
Kenneth Hill
Harvard Center for Population and Development Studies

Many developing countries lack the accurate civil registration of births needed to monitor fertility levels and trends over time. In such settings, fertility is generally estimated from household survey data, either by asking women explicitly about their childbearing or by inferring information about fertility from the age distribution of the population. In the former category, the dominant method has become the full birth history, henceforth FBH, first widely implemented by the World Fertility Survey in the late 1970s, and more recently by the Demographic and Health Surveys (DHS) program. In this approach, each woman (in some settings limited to ever-married women) interviewed is asked for the date (usually month and year) of each of her live births plus other information such as the sex and survival status of the child. In the second category, the most informative approach is the own children method, henceforth OC, whereby mothers are linked to their own children in each household, the age of the mother and the age of the child taken together providing a basis for estimating age patterns and time trends of fertility, with level adjustments then applied to take account of the deaths of children and women. The rise of the FBH approach over the last 30 years has coincided with a decline in interest in the OC method, even though the latter can be applied at low cost to census data to provide estimates at high levels of disaggregation. The purpose of this paper is to examine the relative performance of the two methods in terms of estimates of the Total Fertility Rate.

First we examine the logical basis of the two methods. The FBH method relies on complete reporting of all births for a group of women, so it produces exact computations that include children who do not live in the same household (non-own) and children who have died. By contrast, the OC method relies on household-level data that typically do not provide any information about the mothers of non-own children or about past child mortality. Consequently, the OC method must use probabilistic rules to estimate the effects of each of these phenomena. Given these procedural differences, the two methods cannot be expected to produce identical fertility estimates.

In practice, data collection practices induce a number of additional differences between the FBH and OC computations, and we aim to identify the sources of systematic discrepancies in their fertility estimates. The broader goal of the paper is to assess the accuracy of the OC method. Since the FBH method requires highly detailed data that are expensive to collect, it is only possible to use it periodically and for relatively small samples in each country. The OC method on the other hand can be applied to any household survey that collects age of household members, and offers the additional advantage of fertility estimates for single year periods and for smaller population sub-groups since it can be applied to census data; thus an analysis of its performance is useful to determine whether it can be applied without substantial loss of accuracy in fertility estimates.

We apply the FBH and OC methods to paired data sets collected by DHS for 56 different countries. DHS collected both full birth histories (in the Individual Recode (IR) data) and the equivalent of household survey data (HS) for residents of the same households in each country, facilitating a comparison of the results for the two methods.

The linked nature of the HS and IR data collected by DHS facilitates our comparisons of results for the OC and FBH methods. We develop a procedure for comparing these methods based on three different matches of observations for each country.

(1) Matching of women in the relevant age range (15 to 49) from IR and HS data.

(2) Matching of children listed in the IR data to their mothers in the HS data.

(3) Matching of children listed in the HS data to their mothers (if living in the same household) in the FBH data.

DHS surveys conducted since 2001 provide sufficient information for all three types of matching. DHS surveys before 2001 generally provide enough information for the first two types of matching, but not the third. Even for DHS surveys prior to 2001, however, we can still follow our analytic procedure under the assumption of consistent matches of children to mothers across the IR and HS data.[1] We cannot follow our analytic procedure, however, when DHS data are insufficient to distinguish own-children and non-own children in either the IR or HS data.

We begin by defining conditions for a base case that produces identical fertility estimates for the two methods. We isolate (1) women who appear in both the IR and HS data and whose ages match across these two files; (2) children who are listed as living in the same

[1] We analyze two surveys conducted since 2001 (Chad (2004) and Peru (2004)) where it is not possible to complete the third type of matching. We treat these surveys as if they were conducted before 2001.

household as those women, who can be matched across the two files, and who are listed as the same age in both files. This ensures that FBH and Own Child computations are based on the same set of children and mothers and that their ages are consistent across the files.

Table 1 lists a series of seven cumulative steps beyond the Base Case that combine to produce TFR estimates. With the exception of Step 3, each step results in a refined estimate of TFR for both the FBH and OC methods (Step 3 only affects the OC computations). The implementation of each of these steps differentiates the fertility estimates for FBH and OC methods.

Figure 1 compares the Total Fertility Rates (TFR) based on a three year sample period of births for the FBH and OC methods for the original 56 countries.[2] Simple visual inspection indicates that the FBH TFR lies above the OC TFR in the majority of instances.

Table 2 lists the average TFR's at each step for the two methods. The FBH method produces a larger TFR (at the end of Step 7) than the OC method in 80% of the cases, with an average difference of 0.16 births. The difference in means of TFR's is statistically significant at the 5% level at every step, including step 1.

Table 3 shows mean TFRs across all surveys with 95% confidence intervals for each measure. Despite the relatively small sample of 56 countries, a null hypothesis that the FBH and OC methods produce equal TFR's on average can be rejected at any reasonable level of statistical significance.

[2] As we discuss in Section 3, the Birth History method typically uses births from the past three years to estimate the Total Fertility Ratio. However, the distinction between a one year and three year sample period of births does not affect the qualitative comparisons between the estimates produced by these two methods.

Table 4 computes the "Difference in Difference" in results from each step to the next to isolate the contribution of each step of analysis towards the overall difference of TFR's at the end of Step 7. Steps 1 through 4 and also Step 7 contribute statistically significant differences in estimates for the two methods (at the 5% significance level).

The remainder of the paper attempts to assess why the two methods diverge to this degree. We discuss the results for each of Steps 1 through 7, but emphasize Steps 1 through 4 in the remainder of the paper.

Conclusion

We find that the estimates of Total Fertility Rates are systematically higher for the FBH method than for the OC method, and that the magnitude of these differences is surprisingly large – an average difference of about 4% in estimates. Further, we find strong evidence that FBH tends to overestimate fertility as a result of biased sampling; this effect explains about half of the difference in the fertility estimates produced by the two methods.

Table 1: Seven Distinct Steps in the Computation of Total Fertility Rates

Case	Definition
Base Case	Include women listed with the same age in both IR and HS data. Include children living in the same households as those mothers and who are listed with the same ages in both files.
Step 1	Use ages from IR data for FBH computations. Use ages from HS data for OC computations.
Step 2	Divide past years into months.
Step 3	Include women of age 15 to 49 who appear in the HS data but who do not appear in the IR data.
Step 4	Include "Non-Own" children in computations. Conduct "reverse survival" for mothers in HS data.
Step 5	Account for child mortality.
Step 6	Weight the observations based on reported sample weights.
Step 7	Compute results based on 5-year age groups of mothers.

Figure 1: Birth History and Own-Child TFR's for 56 Countries

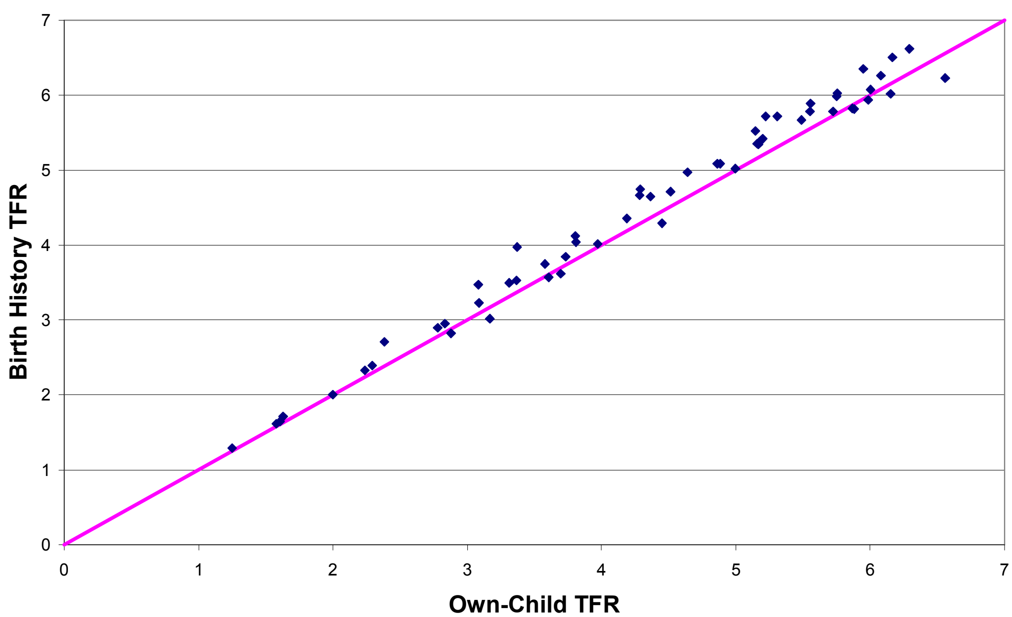

Table 2: Total Fertility Rates for Steps 1 through 7
Full Sample of 56 Countries

	Full Birth History Mean TFR	**Own Child Mean TFR**	**Difference Of Means**	**Number of Times with TFR(BH) > TFR (OC)**
Step 1	3.998	3.958	+ 0.040	33 of 56
Step 2	4.043	3.954	+ 0.089	52 of 56
Step 3	4.043	3.820	+ 0.223	56 of 56
Step 4	4.121	3.934	+ 0.187	52 of 56
Step 5	4.425	4.259	+ 0.166	48 of 56
Step 6	4.432	4.274	+ 0.158	46 of 56
Step 7	4.455	4.288	+ 0.167	46 of 56

Table 3: Descriptive Statistics for FBH and OC TFR

Variable	Mean (Standard Deviation)	95% Confidence Interval for Mean
Own Child (OC) TFR	4.288 (1.432)	(3.905, 4.672)
Full Birth History (FBH) TFR	4.455 (1.467)	(4,062, 4.848)
FBH TFR – OC TFR	0.167 (0.185)	(0.117, 0.216)
FBH TFR / OC TFR	1.041 (0.043)	(1.029, 1.052)
NUMBER OF COUNTRIES	56	

Table 4: Difference in Differences of Total Fertility Ratios

	Diff. in Diff. of TFRs from Prior Step	T-Statistic for Hypothesis Test	Number of Times with Diff in Diff > 0
Step 1	+0.040	3.89	33 of 56
Step 2	+0.049	11.71	52 of 56
Step 3	+0.134	11.34	45 of 56
Step 4	-0.037	-2.84	20 of 56
Step 5	-0.021	-1.28	28 of 56
Step 6	-0.008	-0.87	29 of 56
Step 7	0.009	2.25	27 of 56