Constructing Fertility Measures with the Own-Children Technique: Old Methods for New Data

Michael Sobczak
Gray Swicegood

University of Illinois at Urbana-Champaign

Carl Schmertmann

Florida State University

ABSTRACT

Own-children measures can be used when direct information on fertility is not available. Although a wealth of fertility information is collected in the United States, many existing data sources are not suitable for addressing important questions because the sample sizes are too small (e.g. origin-specific immigrant subgroups) or they lack information on key covariates. Although the decennial Census no longer collects fertility information, own-children techniques applied to the 2000 PUMS files offer a potential remedy to this fundamental problem. Here we outline a protocol for constructing own-children measures for PUMS data and examine the reliability of the measures using several external comparisons. Our initial assessments suggest that the approach yields reasonable aggregate level estimates for many subgroups. We also investigate instances where the estimates appear to be less reliable. Lastly we examine the extent to which the measures can be employed to estimate individual level models of fertility behavior.

Introduction and Overview

Own-children measures are used to estimate levels of fertility in situations where direct information on births to women through vital registration systems, censuses, or surveys is not available. The techniques for constructing own child measures are most closely identified with the work of Cho, Retherford and Choe in their monograph, *The Own-Children Method of Fertility Estimation* (1986). They presciently note in their introduction that "indirect estimation techniques originally developed for situations in which vital registration was seriously deficient will remain useful even as vital registration improves." Our study focuses on the application of the own-children methodology to the contemporary United States. Given the wealth of fertility data collected in this country our focus may seem ironic, but much of the extant information precludes the investigation of substantive questions that are relevant to current population dynamics and theory testing. For example, Hispanic persons make up an increasingly greater proportion of the overall population. Overall, Hispanics appear to have higher levels of childbearing than their non-Hispanic counterparts, but fertility levels vary widely across different Hispanic subgroups. Even very large surveys such as the Current Population Survey (CPS) contain too few cases to support detailed analyses of specific Hispanic subgroups. Larger sample sizes are available from the Public Use Micro-data (PUMS) issued by the Bureau of the Census, but after 1990, the decennial U.S. Census no longer collects any direct information on fertility. One potential remedy to this problem lies in the application of the own-children approach to person and household records in the Public Use Micro-data (PUMS). In this research, we systematically explore the utility of this remedy.

We begin the study by presenting in detail the allocation scheme for linking children with their probable mothers using the 2000 5% Public Use Microdata Sample (PUMS). Based on the extensive relationship coding schema and large sample size, the allocation procedure designed for the 2000 5% PUMS provides the core framework for our own-children methodology. We next show modifications to the allocation scheme in light of the different relationship-to-householder coding schemes used in the Census 2000 Supplementary Survey.

In this paper we are concerned only with current fertility and thus consider children reported age 0 as having been born in the year prior to the Census Day 2000. Initially we compare our own-children fertility estimates based from the PUMS with published vital statistics estimates from NCHS. We report estimates for the overall population, and major racial and Hispanic subgroups. While these comparisons provide some insights into the relative reliability of the own-children estimates, it is important to note that the vital statistics standards are subject to some potential error in their own right. We attempt to pinpoint the sources of discrepancies in the different sets of estimates and draw out the implications of using own-children measures for the estimation of differential fertility in the United States.

Given that unadjusted own-children measures tend to underestimate fertility because some children can not be matched with their mothers, we also explore the utility of these measures for the study of relative fertility differentials by contrasting models of differential recent fertility based on the constructed own-children measures and the directly reported measure in the C2SS data. The C2SS provides unique analytical leverage for evaluating the PUMS own-children models because the C2SS survey included a direct fertility question on whether or not women of childbearing age had a birth in the prior year, which is not available in the 2000 PUMS.

In sum, the goals of the research are as follows:

1) To articulate a well defined algorithm for applying the own-child technique to the contemporary Census data in the United States in absence of any direct information on childbearing.
2) To assess the reliability of the own-child measures by detailed comparisons with estimates and models of recent fertility in the United States from alternative data sources.
3) To identify avenues to improve the allocation scheme underlying the construction of own-children measures.

Implementing the Own-Children Methodology:

The own-children technique is a reverse-survival method that provides a depiction of fertility levels in the years prior to a census or survey. The extent of coverage of prior years depends upon the ages of children that are under consideration. By accounting for the mother's age at the time of birth, age-specific fertility rates and summary total fertility rates (TFR's) can be estimated. The own child method relies upon establishing an accurate link between children and their mothers. Clearly, children must be residing in the same household as their mother for this link to be established. Even if co-residence is the case, many nationally based surveys do not directly ascertain the maternal-child relationship. Some Current Population Survey's have provided the link in the form of a parent line number but typically this link is directed to the child's father. Thus the matching of children in a household to their mothers must be accomplished using other information collected for household members.

Our protocol for producing own-child measures allocates children to their plausible mother based upon several factors as described below. First, "children" were defined as persons between the ages of 0 and 17, while plausible "mothers" were defined as females between the ages of 15 and 62. Therefore, a female between the age of 15 and 17 can be both a child to someone in the household and a plausible mother to someone in the household.

The foundation of our allocation procedure rests upon the "relationship to the household head" coding scheme that is included in most household surveys. This coding scheme provides information on how each person in the household is related to the household head thus allowing us to make inferences about how non-heads in a given household may be related to one another. Obviously, our first concern is with whether or not a child has a plausible mother in the household. In the simple case below, we can infer that the mother of the child is the wife of the household head.

Relationship to Head	Age	Sex
Head	45	Male
Wife	44	Female
Son of Head	15	Male

We have mapped out ways in which a woman could be a plausible mother to a child given the relationship coding in the data sources used for the study. For example, a plausible mother for a child who is coded as “the daughter of the household head” could be one of three possibilities including: (a) the household head, (b) the wife of the household head, or (c) the unmarried partner of the head. After mapping out all plausible relationships between a child with a given code and any plausible mother in the household, we assigned all children and women a grouping variable which becomes the actual basis for our matching assignments. For example, a daughter of the household head was assigned a grouping variable with a value of “1” and all women who could plausibly be her mother were assigned a grouping variable with a value of “1”. In the household shown below the grandchild of the head would be matched to the daughter of the head because they have the same grouping number (number 6), while the daughter of the head is matched to the wife of the head because both have a grouping number of “1”.

Relationship to Head	Age	Sex	Child Group Var	Mother Group Var
Head	45	Male		
Wife	44	Female		(1)
Daughter of Head	17	Female	(1)	(6)
Grandchild	0	Male	(6)	

We also used other information to improve the accuracy of the matches. First, we implemented age constraints on whether a given match is allowed to occur. Before a potential woman is matched to a child, she has to be more than 13 years older than the child and no more than 45 years older than the child. These age constraints will have little bearing on the overall level of the measure because childbearing is relatively rare out side these ranges. Second, for complex households with one or more subfamilies, we use the subfamily variable(s) included in most household surveys that allowed us to reliably match children to plausible mothers in such households. Third, in instances where there were multiple plausible mothers for a given child, we narrow down the match to a single plausible mother using the following priorities: first, we matched the child to women with the most plausible relationship code; second, matching to ever married women instead of never married women was established; and lastly, we matched the child to the woman closest to them in the household records.

Appendix A maps out all the plausible child-mother relationships based on the relationship-to -head of household coding scheme and also shows the grouping variables in which children and potential mothers were assigned accordingly. The matching criteria and rules are also provided. This specific coding scheme pertains to the 2000 5% PUMS.

Since the own child program is largely built from the relationship to the household head "schema", it has to slightly be tailored for each major survey because the relationship to the household head schema differs from survey to survey. In the case of the C2SS, the relationship to the household head variable was less detailed than that of the 2000 PUMS data and our matching program was adapted to it. Appendix B shows the specifics of the program and the general matching criteria for the 2000 Census Supplementary Survey. Since relationship coding was less detailed for the C2SS, our allocation of children to their mothers is likely to be less accurate. For example, "step-children of the head" were identified in the PUMS coding scheme, while they were not in the C2SS. In theory this allowed us to more accurately match children to their real mother with the PUMS data.

Our own-child program though similar to that created by the Census Bureau differs in some details. For example, our age restrictions are more conservative than those implemented by the Census Bureau. While we require plausible mothers to be more than 13 years older than the child and no more than 45 years older than the child, the Census Bureau's comparable restrictions are 12 and 54 respectively. On the other hand, the Census Bureau does not allocate relatives and non-relatives to one another, while we allow for such allocations in some instances. This restriction was relaxed in our program after we identified a number of households in which non-relatives clearly appeared to be plausible mothers to children (who had been recorded as a relative to the head) in some instances. Our program resembles that of the Census Bureau in that we use marital criteria and household record position in the allocation procedure.

Reliability of Own-Child Measures:

Overall, using the own child program we allocated 91.75% of targeted children (those who were not adopted or were not foster kids) with the PUMS data (Table 1). This is a very similar figure to that published in a CPS report (Fields 2003) where 91.4% of children estimated reside with either both parents or with just their mother (see Appendix C). This initial assessment to outside published estimates provides evidence that our own child program is operating reliably.

To further assess the reliability of our own child program, we have computed Total Fertility Rates and General Fertility Rates by treating allocated children reported as age "0" as "births in the last year", the numerators for the TFR's and GFR's . The total number of women of childbearing age in the PUMS is used as denominators for these rates. These rates based from the own child program were computed for the 2000 PUMS and are compared to Vital Statistics rates for 1999 and 2000 in Tables 2a and 2b. The first two columns of these tables show the vital statistic figures for 2000 and 1999 respectively, while the third column shows the unadjusted fertility estimates derived from the own-children program. For all women of child-bearing age, the own-children TFR and GFR fertility estimates provide reasonable depictions of period fertility based on their comparability to Vital Statistics figures (Hamilton et al, 2003). The own-child TFR and GFR is about 93% of the value of the 1999 vital statistics counterparts. However the estimates based on the own-child methodology do not correspond equally well to the Vital Statistics figures across all racial/ethnic sub-groups shown in Tables 2a and 2b. For example, the GFR own-child estimate for non-Hispanic white women is over 95% of the value reported in the 1999 Vital Statistics, while the own-child estimate for non-Hispanic black women is only about 82% of the Vital Statistics figure. The ratios of the own-child estimates to Vital Statistics figures for Asian and Pacific Islanders tend to be similar to those of all women, while the ratio for Hispanic women is nearly as low as those observed for non-Hispanic blacks. These results provide some evidence of the reliability of our own-children program, but also raise questions regarding the larger discrepancies observed for some of the racial and ethnic populations.

Of course, we would not expect the TFR's and GFR's computed with the own child approach to correspond directly with the published vital statistics for several reasons. For example, we treat children age zero allocated to a plausible mother as a birth in the last year. However, such kids may be reported as children age "1" instead of age "0" due to age rounding. This phenomenon will act to reduce the size of our numerators. Second, many children reside with just their father or with a grandparent and so we are not able to match them to their mother, which further shrinks our numerators. Thirdly, children who die within their first year may not be enumerated in the census.

Infant deaths in the year prior to the census will operate to reduce the size of our numerators as well. (We can however, adjust for infant mortality across broad population groupings.) Fourth, our numerators and denominators both differ somewhat from those reported by Vital Statistics because we compute rates for women who reported a single racial category (for non-Hispanic Whites, non-Hispanic Blacks, and Asians and Pacific Islanders), despite being provided the opportunity to choose "some other race" and "two or more race" categories. On the other hand, Vital Statistics imputed those who reported "some other race" or "two or more race" categories into one of the four major racial categories (white, black, American Indian, or Asian and Pacific Islander) for their denominators. Furthermore, the racial classification schema on birth certificates was not updated to coincide with the new Census racial classification schema. As a result, mothers were not allowed to choose options such as "some other race" or "two or more race" categories. Numerators for Vital Statistics therefore differ from ours. Lastly, there may be errors in the "relationship to the head" coding scheme that could lead us to misallocate children in some instances and not properly allocate them in others. Below, the impact each of these factors has on our fertility estimates is assessed.

Assessing the Impact of External Factors on Fertility Estimation

Age Rounding

As previously noted, age rounding can significantly reduce the size of own child fertility estimates due to an inherent reduction in the numerator. To assess the presence of age rounding, we have calculated three different fertility estimates using the own child program. In addition to including children age 0 in the numerator (first fertility estimate), we have included kids age 0 and 1 (second fertility estimate), and kids age 0, 1, and 2 (third fertility estimate). If age rounding is a serious data issue we would expect to find an increase in our fertility rates across these estimates because age heaping would lead to a higher proportion of children in the numerators across the three estimates. Results reported in Appendices D and E show age rounding to have a minimal presence in PUMS data. Fertility rates barely change across the three separate fertility estimates. On the other hand, age rounding has a strong presence in C2SS data. Fertility rates clearly increase across our three fertility estimates (see Appendices D and E) suggesting age rounding is pulling down our own child C2SS fertility estimates. In fact, the own child program for the C2SS allocates children (ages 0, 1, and 2) to a plausible mother at a higher rate than the own child program for PUMS data (see Appendix F). Yet, fertility estimates for PUMS data is significantly higher than those for C2SS data, which is in large part due to age rounding. Results reported in Appendix G better illustrate this point. The percentage of women who reported a birth in the last year, but who were not allocated a child age "0" is reduced by a 1/3 to 1/2 for the various sub-groups of women after adjusting the latter category to women who were not allocated a child age "0" or age "1". Age rounding is likely the main source for this drop.

<u>Children Not Residing With Their Mothers</u>

Fertility estimates based from the own child program depend on its ability to match every child to its mother. Clearly, this is impossible because some children do not reside with their mothers. This is not an issue for the Vital Statistics because every birth is recorded along with the mother's information. Therefore, our fertility estimates based from the own child program will inherently be lower than those reported by Vital Statistics. To produce more comparable fertility estimates to those published by the Vital Statistics, we filter all children age "0" who were not matched into the numerators of women with the same racial/ethnic criteria. Results for this adjustment are reported in the fourth column of Tables 2a and 2b. By comparing this adjustment to the prior estimate (in the third column of Tables 2a and 2b) we find the adjustment to be minimal for "all women" as the TFR only increased a 1/6 of a child. Therefore, this adjustment makes little difference overall. However, the impact of this adjustment greatly varies across groups. For example, in the case of non-Hispanic white women, the adjustment of filtering unmatched kids into the numerator only raises their TFR by 1/11 of a child. Similarly the adjustment only raises the TFR for Asians and Pacific Islanders by roughly a 1/10 of a child. On the other hand, this adjustment raises the TFR's for non-Hispanic black and Hispanic women by roughly a 1/3 of a child. Clearly, the proportion of children matched for these latter groups is significantly less than those of the former groups. In sum, children of non-Hispanic black and Hispanic women either have higher rates of living without their mothers or they are more likely to reside in complex household structures, which in turn increase relationship coding errors, thereby leading them not to be matched.

Tables 5a and 5b further illustrate the variation in allocation rates across different categories of women. Using C2SS data, the logistic regression models presented here predict the likelihood of a women reporting a birth in the previous year, but not allocated a child age 0 (Table 5a) or age 0 and 1 (Table 5b). In essence, these results show the characteristics of women for whom we would like to allocate a child but could not due to factors including non-co-residence, ambiguous family relationships, or infant mortality. The relationship between age and matching mistakes takes on a curvilinear pattern with failure to allocate an infant child despite reporting a birth in the last year occurring most likely for women in their twenties. Consistent with the comparisons discussed above, the results also indicate that Hispanic, non-Hispanic black, and Asian women to be significantly more likely to have unmatched births in the prior year (see Model 2, Table 5a). For Asian women the source of this difference appears to be implicated with their nativity (see Model 3, Table 5a). Foreign-born women overall are more likely to be under-allocated. While there is a strong education effect with higher educated women more likely to have a birth in the prior year successfully allocated (see Model 4, Table 5a), the initial effect for lower educated women appears to be more a function of racial and ethnic background (see Model 5, Table 5a). The results in Table 5b modify the dependent variable such that women who were allocated a child aged 0 or 1 are considered successful matches. In effect, this modification allows for age rounding of recent births. These results are entirely consistent with those shown in the previous table and if anything the results show even sharper contrasts between the racial groups. These sharper contrasts are likely attributable to higher rates of children not residing with their mothers and from higher rates of complex family structures (which leads to greater ambiguity in relationship matching). One potential value of the results in this table involves their potential for adjusting the own-children measures by a factor commensurate with the coefficients estimated by the models. This type of adjustment could be applied to the PUMS estimates.

Infant Mortality

Vital Statistics records every live birth to every woman regardless of infant survival. However, children who die in their first year may not be enumerated in the Census. Therefore, own child fertility estimates will inherently be lower than Vital Statistics rates due to infant mortality. In order to produce comparable estimates to those reported by the Vital Statistics, we make an adjustment for infant mortality using Vital Statistics published infant mortality rates (Mathews et al, 2002). Results are reported in the last column of Tables 2a and 2b. Clearly, infant mortality does not play a major role in biasing our estimates as rates only increase slightly from the previous figure, which adjusts for unmatched children.

By making adjustments for children who do not reside with their mothers and infant mortality we produce estimates that are in theory more comparable to the Vital Statistics rates. After implementing these adjustments the ending fertility estimates using PUMS data are quite comparable to Vital Statistics for all women and the five sub-groups of women listed in Tables 2a and 2b.

Numerators and Denominators

Another source of discrepancy between fertility estimates based from the own child program and the Vital Statistics would be the differences in numerators and denominators between both data sources. Table 3 shows the numerators and denominators of PUMS and Vital Statistics (Ventura et al., 2003) for all women and various sub-groups of women. Overall, PUMS denominators are more comparable to Vital Statistics than PUMS numerators. When considering denominators, coverage for all women, non-Hispanic white, and Hispanic women is very good as the ratio of PUMS to Vital Statistics ranges from 0.961 to 0.994. Although coverage for non-Hispanic Black women is good (0.935), coverage is poor for Asian and Pacific Islander women (0.89). Furthermore, there are major discrepancies between PUMS numerators and those reported by the Vital Statistics. Although coverage for non-Hispanic white infants is somewhat good (0.918), it is particularly poor for non-Hispanic black, Asian and Pacific Islander, and Hispanic infants all with coverage ranging in the 0.70's. Much of the reason for the large discrepancy in numerators between both data sources is due to PUMS numerators including only children who were matched to a plausible mother. Therefore, not all kids age 0 are included in the figures for the numerator and there is no adjustment for infant mortality.

Furthermore, some of the discrepancies in coverage are due to differences in questions on race between both data sources. Vital Statistics numerators allow mothers to report one of four major racial groups only (white, black, Asian and Pacific Islander, Native American and Alaskan Native), while Census classification allows reporting of the same four major racial groups along with "some other race" and "two or more race" categories. These differences in racial classification suggest that our numerators would inherently be smaller than those of the Vital Statistics. Differences in denominators between both data sources also reflect these same differences in racial classification. An important question that arises is: do differences in racial classification between both data sources explain differing fertility estimates between the PUMS and Vital Statistics? To answer this important question we compute fertility rates for "some other race" and "two or more race" groups separately and by Hispanic ethnicity (see Tables 4a and 4b).

Results show little difference between non-Hispanics reporting "some other race" or "two or more race" categories and the fertility estimates for non-Hispanic groups in Tables 2a and 2b. Therefore, if we were to filter non-Hispanic women of "some other race" and "two or more race" categories into one of the four major non-Hispanic racial categories (non-Hispanic white, non-Hispanic black, Asian and Pacific Islander, and Native American and Alaskan Native) reported in Table 2, such a procedure would have very little impact on altering our fertility estimates. In sum, we find that the difference in racial classification between both data sources counts little toward explaining differences in fertility estimates between the two data sources.

Fertility of Hispanic Sub-Groups

Vital Statistics does not produce fertility estimates for many important sub-groups of women such as various national origin groups of women or foreign-born women. However, in the case of Hispanic women, separate fertility rates are published for women of Mexican, Cuban, and Puerto Rican backgrounds as well as a residual "other Hispanic" category. Since we are interested in evaluating the reliability of the own-child program for numerically smaller categories of women we compare fertility estimates based from the own-child program to rates reported by the Vital Statistics. Comparisons for Hispanic women overall and women in the four Hispanic subgroups are shown in Tables 6a and 6b. Overall, for Hispanic women, the PUMS TFR estimates are about one-third of a child lower than the 1999 Vital Statistics figure. Most of this discrepancy is attributable to the own-child underestimation of Mexican and "other-Hispanic" fertility. The own-child estimates for Puerto Rican women are quite comparable to the Vital Statistics and are actually higher for the Cuban women. When we return unmatched children to the numerators for all Hispanic women and for the various Hispanic subgroups, the PUMS estimates closely correspond to the Vital Statistics reports. Thus, lower values of the own-child estimates are more a function of our inability to match all children with a mother, than the absence of Hispanic children from the household records collected by the Census Bureau. Of course some of the unmatched children we move to the PUMS numerators may in fact be the offspring of non-Hispanic women and Hispanic fathers. This might for example account for the relatively high own-child estimate of Cuban fertility. In effect we must impute the mother's ethnicity from that reported for the child, when we adjust for unallocated kids. Comparisons of GFR's are shown in Table 6b. These results are largely consistent with those discussed for the TFR's with the exception that the unadjusted own-child estimate for Puerto Rican women exceeds the reported Vital Statistics GFR for these women.

Modeling Fertility Outcomes Using the Own-Child Measures

The value of the own-children measures extends beyond the provision of reasonable estimates of aggregate fertility levels for subgroup comparisons. These aggregate measures have individual-level counterparts, which can then be used to model fertility behavior at the individual or household level. (See for example, Rindfuss and Sweet, 1977; Bean and Swicegood 1985.) By applying the own-children technique to the PUMS 5% samples, we can potentially model fertility with a richer set of covariates for smaller population groups, for example, Mexican origin women in Iowa. This type of investigation would not be feasible on a national level with any existing vital statistics data or fertility survey data source of which we are aware. In this final analytical section, we examine the reliability of individual level models of current fertility using the own-child measure as the dependent variable. The comparisons in this instance are logistic regression models estimated with the C2SS data and the PUMS. The dependent variable in the case of the C2SS is whether or not the woman had a birth in the last year, a direct fertility measure.

The results of these comparisons are shown in Table 7. The models are not designed to test any particular theory. We use them to assess the reliability of the own-child measures in terms of the similarity of patterns of coefficients for basic socio-demographic variables when compared to models estimated with the direct question on recent fertility in the C2SS. We present only two models for illustrative purposes. The first includes age and race/ethnicity as predictor variables. The second adds education and a dummy variable for nativity. There is a strong correspondence between the PUMS and C2SS results regarding the logistic coefficients for age. The most obvious exception is the case of 30-34 year old women where the PUMS coefficient is -.103 while the C2SS coefficient is only -.036 (see model 2, Table 7). When we examine the coefficients for race and ethnicity the results are somewhat less similar across data sources. Relative to the C2SS non-Hispanic black fertility appears to be underestimated in the PUMS. It is important to note here, the while the sign for these coefficients in the PUMS models is negative that the actual magnitude of the effect is trivial. As we would expect given the aggregate comparison discussed earlier in the paper, the "effects" for Asian, Hispanic and other variables are underestimated with the own-child measure when the C2SS models are the point of reference. In the case of Hispanics and "Others", the gap between the two models is fairly substantial. Nevertheless, both models do capture the substantially higher fertility of this group. Differences between foreign-born and native-born women are also indicated in both data sets, with the estimated differential being higher in the case of the C2SS model. Estimated effects of education s are quite minor in both data sources.

The differences that we report between the models for the PUMS own-child measure and the C2SS direct question are largely expected. Hispanics, non-Hispanic blacks, and the foreign born all have more complex household structures and are more likely to have children living with non-relatives. Thus they are social groups for whom the own-child methodology is less likely to match children with their mothers. This leads to an underestimate of the group differentials when the comparison group is non-Hispanic whites. Differential age reporting across groups could exacerbate that tendency. Still however, there is much congruence between the sets of models, and major differentials are captured in the own-child models.

Conclusion

The primary objective of the research that we have presented is to implement and evaluate an own-children methodology of fertility estimation for the 2000 PUMS data. The value of this exercise resides largely in the fact that these data are the best possible source for national level analysis of fertility behavior for relatively small subpopulations. No longer does the Census or the CPS collect direct information on fertility. While the SIPP, NSFG, and American Community Surveys may partially address the data requirements for the analysis of American childbearing, only the PUMS contains the sample sizes necessary for fine-grained analyses and spatial comparison.

However, the PUMS data can serve us well only if we are able to construct reliable fertility measures from the variables that are collected by the Census. Here we have detailed an allocation schema using data on age, relationships to household head, marital status, subfamily codes, and record location within the household. In the first instance our approach is a conservative one. We have tried to ensure that the allocation of children to mothers is accurate rather than to maximize the proportion of children allocated. Relative to prior Censuses, the C2SS, and prior CPS's, our matching scheme benefits from the more detailed relationship codes implemented in the 2000 Census.

After the implementation of the own-children methodology, we have gauged the reliability or the own-children technique by comparing the own-child estimates with external standards, primarily the Vital Statistics and the C2SS reports on births in the prior year. We know that the allocation schema is working well overall, to the extent that a high proportion of all children who are in households with a plausible mother are matched. When we compare aggregate own-child estimates from the PUMS with Vital Statistics reports not surprisingly, we find that aggregate estimates by race/ethnicity are quite good for some groups and less so for others. In particular, the rates for Hispanics appear to be considerable underestimated by the own-child measure. Subsequently we pinpoint the source of the discrepancies by considering the impact of age rounding, non-co-residence of mother and child, infant mortality, and count differences in numerators and denominators for the rates across data sources. To the extent that we are able to make adjustments for these differences, we produce estimates for all larger groups that are very comparable to the Vital Statistics figures. A similar conclusion holds for the smaller Hispanic sub-groups that we investigate later in the paper. In further analyses, we used the C2SS data to directly model the discrepancy between own-child estimates and births in the last year. These results quantify in a more precise way the extent of differential error that may be introduced in the construction of own children measures. One avenue for further research involves using those results to produce a hot deck procedure for allocating children to individual mothers in cases where we were not able to match them with our basic schema.

In a final set of analyses, we address the question of the reliability of the own child measure for estimating individual level models of fertility. We compare models estimated from C2SS data with those from the PUMS using a basic set of socio-demographic covariates of fertility. These results reaffirm to some extent group differentials that we observed in the earlier analyses, but they also show that the own-child models can capture differentials that are the result of real behavior differences as opposed to differential measurement error across groups. In some ways this outcome is as good as we can hope for. We know that differences in living arrangements and other factors will produce inconsistencies between own-child and direct measures across groups. These sources of error cannot be totally abated by adjustments at the individual level even with very strong data requirements and assumptions. The extent of measurement error in the own-child measure that can be tolerated is an important question that must be addressed in the context of specific substantive concerns of the researcher and weighed against alternatives. In our view, the own child technique can provide sufficiently reliable estimates as to afford greater insights into contemporary American fertility regime than might otherwise be possible. Given the increasing diversity of the country it is important to gauge fertility levels for many different social categories of women such as immigrants, multiracial persons etc. The own child technique when applied to our largest, most comprehensive data source (PUMS) offers such an opportunity.

Table 1: Number of Children Under Age 18: 2000 U.S. Census Public Use Microdata Sample, 5 percent sample[1]

	Number of children under age 18	
	Targeted children	**All Children**
Children who are matched with their mothers	3,240,357 *91.75%*	3,240,357 *89.35%*
Children who are not matched with their mothers	291,251 *8.25%*	291,251 *8.03%*
Adopted children	–	79,689 *2.20%*
Foster Children	–	15,429 *0.43%*
Total	3,531,608	3,626,726

Source: 2000 U. S. Census Public Use Microdata Sample, 5-percent sample (U. S. Bureau of Census 2002)

[1] These figures exclude all people under age 18 who are living in group quarters, householders, spouses of householders, subfamily reference people, and spouses of subfamily reference people (this is done to allow an unbiased comparison to the published CPS figures reported in Appendix C).

Table 2a: $TFR_{15\text{-}44}$ OWN-CHILD MEASURES (2000 5% PUMS) COMPARED With VITAL STATISTICS, 1999 & 2000

	2000 Vital Statistics	1999 Vital Statistics	OC Program Age 0 (number of women)	OC Program Age 0 with Unmatched Kids Added to Numerator	Adjusted for Infant Mortality
All Women	**2.056**	**2.008**	**1.875** (N=2,934,741)	**2.032**	**2.046**
White	**2.051**	**2.008**	**1.852** (N=2,155,335)	**1.961**	**1.972**
Non-Hispanic White	**1.866**	**1.839**	**1.811** (N=1,977,344)	**1.903**	**1.914**
Black	**2.129**	**2.083**	**1.796** (N=361,759)	**2.098**	**2.127**
Non-Hispanic Black	**2.179**	**2.134**	**1.791** (N=355,137)	**2.090**	**2.118**
Asian Alone or P.I.	**1.892**	**1.755**	**1.594** (N=122,439)	**1.703**	**1.712**
Hispanic	**2.730**	**2.649**	**2.272** (N=397,829)	**2.575**	**2.589**
Native American[a] And/or Alaskan Native	**1.773**	**1.784**	**2.030** (N=32,576)	**2.375**	

[a] Age specific infant mortality rates were complete for this group. Therefore we were unable to complete the infant mortality adjustment.

Table 2b: GFR_{15-44} OWN-CHILD MEASURES (2000 5% PUMS) COMPARED With VITAL STATISTICS, 1999 & 2000

	2000 Vital Statistics	1999 Vital Statistics	OC Program Age 0 (number of women)	OC Program Age 0 with Unmatched Kids Added to Numerator	Adjusted for Infant Mortality
All Women	**65.9**	**64.4**	**59.6** (N=2,934,741)	**64.5**	**65.0**
White	**65.3**	**64.0**	**58.0** (N=2,155,335)	**61.4**	**61.8**
Non-Hispanic White	**58.5**	**57.7**	**56.3** (N=1,977,344)	**59.2**	**59.5**
Black	**70.0**	**68.5**	**57.6** (N=361,759)	**67.3**	**68.2**
Non-Hispanic Black	**71.4**	**69.9**	**57.4** (N=355,137)	**66.9**	**67.9**
Asian Alone or P.I.	**65.8**	**60.9**	**56.1** (N=122,439)	**60.0**	**60.3**
Hispanic	**95.9**	**93.0**	**78.5** (N=397,829)	**89.0**	**89.5**
Native American And/or Alaskan Native	**58.7**	**59.0**	**65.0** (N=32,576)	**76.0**	**76.7**

Table 3: NUMERATORS AND DENOMINATORS: PUMS VERSUS 2000 VITAL STATISTICS

	PUMS Numerators Weighted	Vital Stats Numerators (PUMS/Vital)	PUMS Denominators Weighted	Vital Stats Denominators (PUMS/Vital)
All Women	**3,483,664**	**4,058,814** (0.858)	**60,189,581**	**61,576,997** (0.977)
Non-Hispanic White	**2,163,708**	**2,358,050** (0.918)	**39,430,702**	**41,049,946** (0.9606)
Non-Hispanic Black	**439,774**	**600,155** (0.733)	**7,958,023**	**8,509,294** (0.9352)
Asian Alone or P.I.	**145,667**	**200,055** (0.728)	**2,714,499**	**3,049,759** (0.8901)
Hispanic	**638,798**	**812,661** (0.786)	**8,458,466**	**8,509,703** (0.9940)

Table 4a: $GFR_{15\text{-}44}$ OWN CHILD MEASUES (2000 5% PUMS) FOR SOME OTHER RACE AND TWO OR MORE RACE RESPONDENTS BY ETHNICITY

	OC Program Age 0 (number of women)	OC Program Age 0 with Unmatched Kids Added to Numerator	PUMS Denominators Weighted
Some Other Race Alone	**2.291** (N=187,678 women) (N=15,114 kids)	**2.593** (1,991 kids added)	**3,889,060**
Some Other Race Alone ***(Hispanic) – 97.6%***	**2.303** (N=183,150 women) (N=14,847 kids)	**2.605** (1,945 kids added)	**3,791,965**
Some Other Race Alone ***(non-Hispanic) – 2.4%***	**1.765** (N=4,528 women) (N=267 kids)	**2.076** (46 kids added)	**97,095**
Two or More Races	**1.931** (N=74,954 women) (N=4,808 kids)	**2.308** (941 kids added)	**1,581,732**
Two or More Races ***(Hispanic) – 32.4%***	**2.169** (N=24,280 women) (N=1,796 kids)	**2.554** (318 kids added)	**508,250**
Two or More Races ***(non-Hispanic) – 67.6%***	**1.813** (N=50,674 women) (N=3,012 kids)	**2.189** (623 kids added)	**1,073,482**

Table 4b: GFR_{15-44} OWN-CHILD MEASURES (2000 5% PUMS) FOR SOME OTHER RACE AND TWO OR MORE RACE RESPONDENTS BY ETHNICITY

	OC Program Age 0 (number of women)	OC Program Age 0 with Unmatched Kids Added to Numerator	PUMS Denominators Weighted
Some Other Race Alone	**80.5** (N=187,678 women) (N=15,114 kids)	**91.1** (1,991 kids added)	**3,889,060**
Some Other Race Alone *(Hispanic) – 97.6%*	**81.1** (N=183,150 women) (N=14,847 kids)	**91.7** (1,945 kids added)	**3,791,965**
Some Other Race Alone *(non-Hispanic) – 2.4%*	**59.0** (N=4,528 women) (N=267 kids)	**69.1** (46 kids added)	**97,095**
Two or More Races	**64.1** (N=74,954 women) (N=4,808 kids)	**76.7** (941 kids added)	**1,581,732**
Two or More Races *(Hispanic) – 32.4%*	**74.0** (N=24,280 women) (N=1,796 kids)	**87.1** (318 kids added)	**508,250**
Two or More Races *(non-Hispanic) – 67.6%*	**59.4** (N=50,674 women) (N=3,012 kids)	**71.7** (623 kids added)	**1,073,482**

Table 5a: Logistic Regression Results Predicting Women Who Reported a Birth in Last Year But Who Were Not Allocated a Child Age 0: Women Ages 15-44 (C2SS)

	Model 1	Model 2	Model 3	Model 4	Model 5
Age					
15-19	-0.706**	-0.712**	-0.692**	-1.061**	-0.918**
20-24[b]	—	—	—	—	—
25-29	-0.068	-0.076	-0.095	-0.031	-0.065
30-34	-0.135	-0.119	-0.141	-0.112	-0.125
35-39	-0.766**	-0.733**	-0.754**	-0.756**	-0.748**
40-44	-1.402**	-1.347**	-1.365**	-1.395**	-1.366**
Race and Ethnicity					
Non-Hispanic White[b]		—	—		—
Non-Hispanic Black		0.602**	0.578**		0.531**
Non-Hispanic Asian		0.400**	0.104		0.175
Hispanic		0.807**	0.612**		0.514**
All Others		0.633**	0.605**		0.579**
Nativity					
Foreign-born			0.399**		0.340**
Native-born[b]			—		—
Education					
0-8th grade				0.510**	0.179
Less than high school				0.211**	0.076
High school[b]				—	—
Some college				-0.407**	-0.375**
College graduates				-0.438**	-0.367**
Sample size	78,608	78,608	78,608	78,608	78,608
Model chi-square	300.4**	495.8**	527.5**	429.5**	594.2**

Source: Census 2000 Supplementary Survey (C2SS)

Note: [b] Reference category. * p< .05, ** p< .01

Table 5b: Logistic Regression Results Predicting Women Who Reported a Birth in Last Year But Who Were Not Allocated a Child Age 0 or Age 1: Women Ages 15-44 (C2SS)

	Model 1	Model 2	Model 3	Model 4	Model 5
Age					
15-19	-0.225*	-0.233*	-0.214	-0.726**	-0.560**
20-24[b]	—	—	—	—	—
25-29	-0.240*	-0.251*	-0.269*	-0.160	-0.199
30-34	-0.249*	-0.229*	-0.250*	-0.191	-0.202
35-39	-0.607**	-0.562**	-0.583**	-0.572**	-0.554**
40-44	-0.844**	-0.771**	-0.787**	-0.816**	-0.768**
Race and Ethnicity					
Non-Hispanic White[b]		—	—		—
Non-Hispanic Black		0.871**	0.851**		0.766**
Non-Hispanic Asian		0.379*	0.126		0.252
Hispanic		1.028**	0.865**		0.711**
All Others		0.607**	0.583**		0.534**
Nativity					
Foreign-born			0.346**		0.258*
Native-born[b]			—		—
Education					
0-8th grade				0.580**	0.227
Less than high school				0.302**	0.145
High school[b]				—	—
Some college				-0.576**	-0.537**
College graduates				-0.823**	-0.720**
Sample size	78,608	78,608	78,608	78,608	78,608
Model chi-square	63.1**	253.2**	266.4**	213.7**	351.4**

Source: Census 2000 Supplementary Survey (C2SS)
Note: [b] Reference category. * p< .05, ** p< .01

Table 6a: TFR_{15-44} OWN-CHILD MEASURES (2000 5% PUMS) COMPARED With VITAL STATISTICS, 1999 & 2000

	2000 Vital Statistics	1999 Vital Statistics	OC Program Age 0 (number of women)	OC Program Age 0 with Unmatched Kids Added to Numerator	Adjusted for Infant Mortality
Hispanic	**2.730**	**2.649**	**2.272** (N=397,829)	**2.575**	**2.589**
Mexican	**2.907**	**2.823**	**2.437** (N=234,772)	**2.779**	**2.794**
Puerto Rican[a]	**2.179**	**2.105**	**1.934** (N=37,184)	**2.178**	
Cuban[a]	**1.528**	**1.389**	**1.709** (N=11,352)	**1.836**	
Other Hispanic[a]	**2.564**	**2.517**	**2.070** (N=114,521)	**2.325**	

[a] Age specific infant mortality rates were complete for this group. Therefore we were unable to complete the infant mortality adjustment.

Table 6b: $GFR_{15\text{-}44}$ OWN-CHILD MEASURES (2000 5% PUMS) COMPARED With VITAL STATISTICS, 1999 & 2000

	2000 Vital Statistics	1999 Vital Statistics	OC Program Age 0 (number of women)	OC Program Age 0 with Unmatched Kids Added to Numerator	Adjusted for Infant Mortality
Hispanic	**95.9**	**93.0**	**78.5** (N=397,829)	**89.0**	**89.5**
Mexican	**105.1**	**101.5**	**85.9** (N=234,772)	**97.9**	**98.5**
Puerto Rican	**73.5**	**71.1**	**65.0** (N=37,184)	**73.3**	**73.9**
Cuban	**49.3**	**47.0**	**56.6** (N=11,352)	**60.9**	**61.1**
Other Hispanic	**85.1** (N=114,521)	**84.8**	**69.9**	**78.5**	**78.9**

Table 7: Logistic Regression Results Predicting Women Who Were Allocated a Child Age 0: Women Ages 15-44 (2000 PUMS; C2SS)

	PUMS		C2SS	
	Model 1	Model 2	Model 1	Model 2
Age				
15-19	-1.196**	-1.196**	-1.212**	-1.281**
20-24[b]	—	—	—	—
25-29	0.145**	0.110**	0.104*	0.078
30-34	-0.064**	-0.103**	-0.004	-0.036
35-39	-0.846**	-0.879**	-0.877**	-0.907**
40-44	-2.305**	-2.339**	-2.000**	-2.026**
Race and Ethnicity				
Non-Hispanic White[b]	—	—	—	—
Non-Hispanic Black	-0.030**	-0.020*	0.113*	0.097*
Non-Hispanic Asian	-0.094**	-0.256**	0.047	-0.142
Hispanic	0.239**	0.163**	0.389**	0.252**
All Others	0.038*	0.021	0.223*	0.209*
Nativity				
Foreign-born		0.173**		0.238**
Native-born[b]		—		—
Education				
0-8th grade		-0.009		0.008
Less than high school		-0.013		0.014
High school[b]		—		—
Some college		-0.107**		-0.204**
College graduates		0.135**		0.022
Sample size	2,934,741	2,934,741	78,608	78,608
Model chi-square	55281.6**	56914.9**	1753.4**	1820.4**

Source: 2000 5% PUMS; Census 2000 Supplementary Survey
Note: [b] Reference category. * p< .05, ** p< .01

APPENDIX A: Matching Criteria for 2000 5% PUMS

Group	Child Group	Mother Group 1	Mother Group 2
1	Parent-child subfamily Subfamily 1	Parent in subfamily 1	
1.2	Parent-child subfamily Subfamily 2	Parent in subfamily 2	
1.3	Parent-child subfamily Subfamily 3	Parent in subfamily 3	
2	Natural Born Son/Daughter	Householder Wife[1] Unmarried Partner[2]	
3	Stepkid	Wife Unmarried Partner[2]	
4	Nephew/Niece	Sister Sister-In-Law	
5	Householder Brother/Sister	Parent	
7	Husband/Wife Brother/Sister-In-Law	Parent-In-Law	
8	Grandchild	Natural daughter Adopted daughter Stepdaughter Daughter-In-Law	
9	Father/Mother	Grandparent	
10	Cousin	Aunt	
11	Roomer/Boarder	Roomer/Boarder	
12	Housemate/Roommate	Housemate/Roommate	
13	Other Non-Relative	Other Non-Relative	Unmarried Partner

[1] Only included when there is no female householder.
[2] Only included when there is no female Householder or Wife of householder.
Women with the relationship codes in the highlighted boxes are included in the analysis only when there are no women in the HH with the same relationship codes as those in the boxes to the left of the highlighted boxes.

In cases where a child is matched to more than one mother the following occurs:

(1) If a child is matched to both a ever married female(s) (married, separated, divorced, or widowed) and a never married female(s), the allocation to the never married female(s) will be released while the allocation to the ever married female(s) will remain intact.
(2) If a child is still matched to more than one mother (e.j. two or more females are ever married or two or more females are never married) then we match the child to the female closest to the child in the list of individuals in the household.

APPENDIX B: Matching Criteria for 2000 Census Supplementary Survey

Group	Child Group	Mother Group 1	Mother Group 2
1	Parent-child subfamily Subfamily 1	Parent in subfamily 1	
1.2	Parent-child subfamily Subfamily 2	Parent in subfamily 2	
1.3	Parent-child subfamily Subfamily 3	Parent in subfamily 3	
2	Son/Daughter	Householder Wife[1] Unmarried Partner[2] Other Non-Relative[3]	
3	Other Relative	Other Relative	Wife Brother/Sister Grandchild Unmarried Partner
4	Householder Brother/Sister	Father/Mother	
5	Other Non-Relative	Other Non-Relative	Housemate Roommate Unmarried Partner
6	Grandchild	Son/Daughter In-law	Other Relative Other Non-Relative
7	Husband/Wife In-law	In-law	
8	Roomer/Boarder	Roomer/Boarder	
9	Housemate/Roommate	Housemate/Roommate	

[1] Only included when there is no female householder.
[2] Only included when there is no female Householder or Wife of householder.
[3] Only included when there is no female Householder, Wife of householder, or Unmarried partner of householder.

Women with the relationship codes in the highlighted boxes are included in the analysis only when there are no women in the HH with the same relationship codes as those in the boxes to the left of the highlighted boxes.

In cases where a child is matched to more than one mother the following occurs:

(1) If a child is matched to both an ever married female(s) (married, separated, divorced, or widowed) and a never married female(s), the allocation to the never married female(s) will be released while the allocation to the ever married female(s) will remain intact.
(2) If a child is still matched to more than one mother (e.j. two or more females are ever married or two or more females are never married) then we match the child to the female closest to the child in the list of individuals in the household.

APPENDIX C: Comparison of Aggregate Matching

Number of Children Under Age 18 by Living Arrangement: March 2002 Current Population Survey[2]

	Two parents	**Mother only**	**Father only**	**Neither parent**	**All children**
Children with native parents	38,255 *77.0 %*	14,192 *86.2%*	2,817 *85.4%*	–	–
Children with at least one foreign-born parents	11,411 *23.0%*	2,281 *13.8%*	480 *14.6%*	–	–
Total	49,666 *68.7%*	16,473 *22.8%*	3,297 *4.6%*	2,885 *4.0%*	72,321

Source: Current Population Reports (Fields 2003)

[2] In the Current Population Report (Fields 2003) all people under age 18 who are living in group quarters, householders, spouses of householders, subfamily reference people, and spouses of subfamily reference people are excluded from the figures.

APPENDIX D: $TFR_{15\text{-}44}$ Own Child Measures for Different Combinations of Ages Using the PUMS and C2SS

	PUMS			C2SS		
	OC Program Age 0	OC Program Age 0,1	OC Program Age 0,1,2	OC Program Age 0	OC Program Age 0,1	OC Program Age 0,1,2
All Women	**1.875**	**1.865**	**1.845**	**1.575**	**1.695**	**1.710**
Non-Hispanic White	**1.811**	**1.804**	**1.779**	**1.550**	**1.670**	**1.682**
Non-Hispanic Black	**1.791**	**1.789**	**1.785**	**1.475**	**1.614**	**1.684**
Asian Alone or P.I.	**1.594**	**1.554**	**1.543**	**1.358**	**1.425**	**1.471**
Hispanic	**2.272**	**2.252**	**2.234**	**1.879**	**1.943**	**1.947**
Mexican-Origin	**2.437**	**2.421**	**2.401**	**2.036**	**2.143**	**2.111**
Mexican-Born	**2.778**	**2.710**	**2.691**	**2.347**	**2.382**	**2.369**

APPENDIX E: $GFR_{15\text{-}44}$ Own Child Measures for Different Combinations of Ages Using the PUMS and C2SS

	PUMS			C2SS		
	OC Program Age 0	OC Program Age 0,1	OC Program Age 0,1,2	OC Program Age 0	OC Program Age 0,1	OC Program Age 0,1,2
All Women	**59.6**	**59.5**	**59.2**	**49.6**	**53.9**	**54.8**
Non-Hispanic White	**56.3**	**56.5**	**56.2**	**48.2**	**52.6**	**53.5**
Non-Hispanic Black	**57.4**	**57.3**	**57.2**	**45.1**	**50.0**	**52.4**
Asian Alone or P.I.	**56.1**	**54.9**	**54.6**	**47.9**	**50.2**	**52.2**
Hispanic	**78.5**	**77.7**	**76.9**	**63.9**	**66.2**	**66.4**
Mexican-Origin	**85.9**	**85.0**	**84.0**	**71.2**	**74.5**	**73.4**
Mexican-Born	**99.6**	**97.8**	**97.7**	**82.2**	**85.1**	**85.8**

APPENDIX F: Percentage of Targeted Kids Matched by Survey

	Age 0	Age 1	Age 2
C2SS	94.92	95.10	94.50
PUMS	93.61	93.71	93.25

APPENDIX G: Percentage Change in the Number of Women Who Reported a Birth In the Last Year But Who Were Not Allocated a Child, C2SS

	Percent Reporting Birth in Last Year, But Not Allocated a Child Age 0	Percent Reporting Birth in Last Year, But Not Allocated a Child Age 0 or 1
All Women	2.15	1.16
Non-Hispanic White	1.68	0.83
Non-Hispanic Black	3.09	2.00
Asian or Pacific Islander	2.61	1.24
Hispanic	3.96	2.40
Mexican-Origin	4.23	2.57

APPENDIX H: Comparison of Vital Statistics to Birth in Last Year Estimates, $TFR_{15\text{-}44}$ **(C2SS)**

	2000 Vital Stats	1999 Vital Stats	C2SS – Birth Last Year
Non-Hispanic White	1.866	1.839	1.973
Non-Hispanic Black	2.179	2.134	2.257
Asian Along or P.I.	1.892	1.755	1.968
Hispanic	2.730	2.649	2.823
Mexican-Origin	2.907	2.823	3.022

References

Cho, Lee-Jay, Robert D. Retherford and Minja Kim Choe. 1986. The Own Children Method of Fertility Estimation. Honolulu: East-West Center

Fields, Jason. 2003. "Children's Living Arrangements and Characteristics: March 2002." Current Population Report P20-547. Retreived February 22, 2004 http://www.census.gov/prod/2003pubs/p20-547.pdf.

Hamilton, B.E., P.D. Sutton, & S.J. Ventura. 2003. Revised birth and fertility rates for the 1990s and new rates for hispanic populations, 2000 and 2001: United States. National Vital Statistics Reports vol 51, no. 12. Hyattsville, Maryland: National Center for Health Statistics.

Mathews, M., F. Menacker, & M.F. MacDorman. 2002. Infant mortality statistics from the 2000 period linked birth/infant death data set. National Vital Statistics Reports vol 50, no. 12. Hyattsville, Maryland: National Center for Health Statistics.

Rindfuss, Ronald R. and James Sweet. 1977. Postwar Fertility Trends and Differentials in the United States. New York: Academic Press.

U.S. Bureau of Census. 2002. Census of Population and Housing [United States]: Public Use Microdata Sample: 5-percent sample [computer file]. Washington: U.S. Bureau of Census.

Ventura, S.J., B.E. Hamilton, P.D. Sutton. 2003. Revised birth and fertility rates for the United States, 2000 and 2001. National Vital Statistics Reports vol 51, no. 4. Hyattsville, Maryland: National Center for Health Statistics.