UNDERSTANDING TEENAGE FERTILITY DECLINE

Berna M. Torr, Jacob Alex Klerman, and Jeremey Arkes
RAND Corporation, 1700 Main St., Santa Monica, CA 90407-2138
Address correspondence to Berna_Torr@rand.org (1-310-393-0411, x6127)

From their peak in the early 1990s, births to 15–17-years-olds have dropped more than 40 percent; births to 18–19-year-olds have also dropped, in this case, more than 25 percent. In contrast, over the same period, births to 20–24-year-olds are approximately stable, and births to older women (here defined as 30 and older) are rising rapidly (20 to 64 percent).

There is a moderate-sized literature attempting to explain the fertility/pregnancy decline. Those papers include Kaufman et al. (1998), Jones et al. (1999), Saul (1999), Flanigan (2001), Terry-Huymen, Manlove, and Moore (2001), Mohn, Tingle, and Finger (2003), and Santelli et al. (2005). Most of those studies use the 1988 and 1995 NSFG, but those data only included the beginning of the teenage fertility decline. Santelli et al. (2005) used more recent data from the Youth Risk Behavior Survey (YRBS), but those data only include 15–19-year-olds in public high schools. The new 2002 NSFG data provide an opportunity to explore these issues using more current data and a unified approach. Specifically, we:

- *Disaggregate the Fertility Decline*: Aggregate fertility decline is undisputed and large, but differential fertility decline is less well understood. Limited published analysis suggests that teenage fertility decline has not been uniform across sub-groups. It clearly varies by race-ethnicity. It appears to vary by marital status. We hypothesize that it varies by social class as proxied by parental marital status and education. Even in 1992, fertility was concentrated in teenagers from single parent families and families with low parental education. Thus, it seems plausible that the declines in fertility were concentrated in this group.
- *Disaggregate Contraceptive Quasi-Failure Rates*: Corresponding to this exploration of differential fertility decline, we also explore differential contraceptive failure rates, both across time and across subgroups. The recently released 2002 NSFG allows for both the estimation of trends in failure rates and, by pooling the data, more power for the estimation of heterogeneity in those failure rates. As we discuss below, our approach implies a role for a particular (quasi-)failure rate.
- *Disaggregate Sexual Activity and Contraceptive Method Choice*: We also explore differential changes in sexual activity (including primary and secondary abstinence) and contraceptive method. Again, the recently released 2002 NSFG allows for both the estimation of trends in contraceptive method choice and, by pooling the data, more power for the estimation of heterogeneity in such method choice.
- *Simulate the Changes in Proximate Determinants on Fertility*: We combine these estimates of failure rates with estimates of sexual behavior and contraceptive use from the 1988, 1995, and 2002 waves of the NSFG to understand the proximate determinants of teenage fertility decline.

This approach is—in three ways—specifically adapted to the available NSFG data. First, like all surveys, the NSFG under-reports abortion. Unlike other studies that attempt to model declines in pregnancy, our approach attempts to model birth directly, ignoring abortion (see below). Second, the NSFG allows internal estimates of contraceptive failure rates. This allows us to align our contraceptive failure rate measures directly with the concepts used in estimating contraceptive use and fertility (again, see below). Finally, proceeding with all steps of the analysis from a single data set allows unified estimates of the precision of our estimates. It is our reading of the literature that the NSFG is only border-line big enough to estimate the models of interest. Our approach gives special attention both to estimating the parameters precisely and to generating realistic estimates of how precise are estimates truly are.

Overview of Our Approach

To formally describe the problem and our approach, we use the following notation:

- *m – Contraceptive Method*, where we consider "not sexually active," "not contracepting," and combinations of methods as contraceptive "methods." Given the published NSFG tabulations for teenagers and young adults, we propose to begin with the following grouping of contraceptive methods: (i) not sexually active, (ii) not contracepting, (iii) any modern method, (iv) oral contraceptive alone, (v) condom and oral contraceptive.

- *d – Broad Demographic Group*. We think of this as the level of the statistics published in the *Vital Statistics* Reports; e.g., race/ethnicity *x* age (perhaps grouped).
- *g – Narrow Demographic (Sub-)Group*. We think of this as the covariates available in the NSFG, but not in the birth certificates, that proxy for "disadvantage": Parents/parent-figures at age 14, parental marital status, mother's age at first birth, parental education, number of siblings (all measured in the NSFG). For now we focus on whether or not the mother of the sample member had her first birth as a teenager.
- *y – Calendar time (year/month)*.

Then, the published birth rates, for a given month, can be written as:

(1) $$\forall d, y \quad b^{d,y} = \sum_{g} \sum_{m} G_{g}^{d,y} M_{g,m}^{d,y} F_{g,m}^{d,y}$$

where:

- *b – Birth Rate*: Defined as births per woman (not per 1,000 women);
- *G– Population Characteristics*: Fraction of individuals in broad demographic group *d* that are in narrow demographic group *g*;
- *M– Contraceptive Method Used*: Fraction of individuals in broad demographic group *d* and narrow demographic group *g* that are using contraceptive method, *m*;
- *F– Contraceptive Quasi-Failure Rate*: Fraction of individuals in broad demographic group *d* and narrow demographic group *g*, using contraceptive method, *m*, who experience a quasi-contraceptive failure, where a quasi-failure is defined as a month using a method that leads to a live birth.[1]

D.3. Estimation

Our goal is to use Equation 1 to simulate the effect of changes in abstinence, contraception, and the distribution of women across demographic subgroups. If we knew each of the terms—*G, M, F*—exactly, for each narrow demographic subgroup, *g*, this would be a straightforward spreadsheet computation. Of course, we do not know them exactly. Instead, we will estimate them from the NSFG data (described in the next section). The NSFG data are clearly not large enough to do these estimates in a totally disaggregated fashion. Instead, we pool all the NSFG data across the contraceptive calendars in the three waves (1988, 1995, 2002) and estimate models on monthly outcomes (i.e., What contraceptive was used this month? Did this sexual activity in this month lead to a live birth nine months later?). This allows us to "borrow strength" across time periods and across demographic groups, by running appropriate logistic-regression models.

Our approach requires estimating four sets of logistic regressions. Here, we give a brief overview of each of those relationships and how we estimate them. We use a common approach to specifying the regressors across the regressions. We discuss that common specification immediately following our discussion of the relationships themselves. For most of these models, the NSFG reports already provide some published tabulations. By using the information in the NSFG calendars and pooling across waves, we will be able to estimate the trends more precisely and estimate how they vary across demographic subgroups (i.e., disadvantage).

The four relationships are:

- *Disaggregated Birth Rate,* b*s*: We begin by modeling trends in fertility themselves. Note that we specifically model births, not pregnancies. The outcome (a conception leading to a live birth) is one binary outcome per month covered by the calendar (except that we must drop the last six to nine months to cover pregnancies in progress at the interview). Estimation proceeds by conventional logistic regression. (We discuss the computation of standard errors below.)

---

[1] N.b., we are trying to explain births. This diverges from pregnancies, which is the standard outcome in both the contraceptive failure literature and in some of the decomposition literature (Santelli et al., 2004). As we discuss below, abortions are known to be under-reported in surveys. Thus, if we want to compute true contraceptive failure rates (i.e., leading to pregnancy, rather than a live birth), we need to "adjust" the data for abortion under-reporting. This is what Ranjit et al. (2001) do. Santelli et al. (2004) use their estimates, so they need to then adjust for abortion under-reporting.

If abortion was an important candidate for explaining the observed fertility decline, then this exercise might be worthwhile. However, abortion rates appear to be declining; thus, they are unlikely to be an important explanation of declining birth rates. Thus, this approach[what approach?] seems plausible.

- *Contraceptive Method Choice,* M: Recall that we define a "method" to include abstinence, no method, and combinations of methods (e.g., pill and condom). This outcome is multinomial. We model the multinomial choice as a nested set of binomial logistic models on each non-pregnant month of the contraceptive calendar (where, for consistency, we define pregnancy as nine months prior to a live birth). For this problem, the natural nesting is: (i) sexually active, (ii) contracepting, (iii) using any modern method; (iv) using oral contraceptives; and (v) using a condom and oral contraceptives.
- *Contraceptive Quasi-Failure Rates,* F: Recall that we define a failure as a month using the method that leads to a live birth (not the conventional definition of leading to a pregnancy). We do so both because our focus is on changes in the birth rate and also because of concern about the quality of the abortion data. Estimation proceeds by simple logistic regression, where the sample is any month in the calendar using that "method" (but dropping the last six to nine months for women pregnant at the interview).
- *Population Distribution of Demographic Subgroup,* G*:* A shift/share analysis requires estimates of the shares of the subgroups within the broad demographic groups. Again, we want to estimate these shares, and we can gain statistical precision by taking the population totals as given and only estimating the distribution of demographic subgroups within groups defined by the population totals (age, race/ethnicity, perhaps marital status). These models are estimated as nested logit models for each of the characteristics in *G*.

These are the three sets of models to be estimated. In each model, we consider three classes of covariates:

- *Main Effects:* For broad demographic group (age and race/ethnicity), narrow demographic group (our proxies for disadvantage—e.g., mother first gave birth as a teenager, family not intact at age 14, mother a high school drop-out, father a high-school drop-out), and year/month.
- *Demographic Interactions*: A main effect model implicitly assumes that rates for each of the demographic groups (broad and narrow) move together. In as much as this is true, we can "borrow strength." However we would not want to impose this assumption when it is incorrect. Thus, we will include first-order interactions between the main effects. We will then pre-test and drop insignificant terms using the Akaike Information Criteria (AIC).
- *Year Interactions*: We will also interact each of the main effects with time. These interactions are of substantive interest. They answer the question, "Is there a statistically significant time effect?" As appropriate, we will consider polynomials or splines in time (e.g., one time trend between the 1988 NSFG and the 1995 NSFG and a potentially different time trend between the 1995 NSFG and the 2002 NSFG). As appropriate, we will also consider second order-interactions.

Simulating the Effect of Proximate Determinants

Given this framework and estimates for each of the components, we can simulate the effect of a change by counterfactually substituting some other value for one of the terms. Thus, for example, to explore the effect of contraceptive method mix between year y and year y', we could compute the implied birth rate (with a "~") using some alternative method mix (again with a "~"):

$$(2) \qquad \tilde{\Delta}^{\mathrm{d}} = \frac{\tilde{b}^{d,y'} - b^{d,y'}}{b^{d,y'} - b^{d,y}} \quad \text{where} \quad \tilde{b}^{d,y} = \sum_{g}\sum_{m} G_{g}^{d,y'} \tilde{M}_{g,m}^{d,y'} F_{g,m}^{d,y'}$$

Thus, the fraction of the change in fertility, b, between year y and year y', explained by a factor, Δ, is the ratio of the counter-factual change in fertility (the numerator) to the actual change in fertility (the denominator). The actual change (the denominator) is the difference between observed fertility in the later year, y', and observed fertility in the earlier year y. The counter-factual change in fertility (the numberator) is the change in *predicted* fertility between the same earlier year and later years y and y', but the fertility in the later year is predicted using some alternative set of covariates; e.g., holding the contraceptive method mix fixed at its base year values. The right expression (i.e., the double summation) emphasizes that we can compute these counter-factual number of births given our estimated models; for this example, the crucial estimation is of quasi-failure rates (i.e., births given method choice). Depending on the specific counterfactual, we might want to adjust the distribution of contraceptives within the branch of people who are sexually active.[2] Other counterfactuals of interest involve changes in abstinence and the share of the population in different demographic subgroups (i.e., G).

Generating the simulations is straightforward. For each record in the database (each person-month in the NSFG contraceptive calendar, merged over the three waves), we generate values for each of the terms G, M, and F. The generated values are predictions from logistic regression models—with the observed covariates and then with the counterfactual covariates. We can then compute the implied birth rate contribution for each counterfactual, for each record in the database. Within each record, we then do the inner summation over methods. Across all records, we do the outer summation over demographic subgroups (i.e., we compute a summation of a created variable across all the records).

Computing standard errors is more complicated. The underlying estimates of method choice and failure rates have sampling variability. We have used the same data to compute both equations, so the results are almost certainly correlated. Furthermore, the NSFG is generated from a complex sample design with weights adjusting for non-response and correlation within family units and levels of the multi-level sample design. Santelli et al. (2005) considers only the simple sampling variability from method choice. He implicitly assumes that there is no sampling variability in the contraceptive failure rates (or in the abortion adjustments).

In principle, there exists a parametric approach to this challenge of estimating standard errors. In practice, a computer-intensive approach seems more attractive. Specifically, we propose to bootstrap the entire procedure from estimation through to simulation (Efron, 1979; Efron and Stein, 1981; Egron and Tibshirani, 1986, 1993). Heuristically, we will do so by randomly selecting persons (given the NSFG complicated sampling design primary sampling areas—there are 121 in the 2002 NSFG) with replacement, with probability equal to the sum of the sample weights. Doing so generates a bootstrap sample. On this bootstrap sample, we will reestimate the model. Bootstrap estimates of the standard errors of our simulation output can be relatively easily derived using Stata's bootstrap command.

Advantages of Our Approach

Our approach has several significant advantages. First, we propose to pool the 1988, 1995, and 2002 waves of the NSFG and then to test for the presence of time trends. (Earlier waves are available, but the data appear to be less comparable across waves.) Doing so should improve the precision of our estimates. (For a similar strategy with respect to contraceptive failure rates, see Schirm et al., 1982, merging the 1973 and 1976 NSFGs; Ranjit et al. 2001, merging 1988 and 1995 NSFGs). Second, we propose to estimate the entire model—abstinence/contraceptive method choice/contraceptive failure rates—from a single data set. This allows us to consistently estimate standard errors. Third, we propose to use external information on the number of births to improve the precision of our estimates. Fourth, we propose to apply a bootstrap approach to computing the induced sampling variability in the simulation results, while accounting for the complex sample design of the NSFG. Fifth, our primary approach will define away some of the issues with estimating contraceptive failure rates. Our primary interest is in explaining the decline in birth rates (not the decline in pregnancy rates). Since abortion rates appear to be moving in the "wrong" direction (i.e., they do not explain any of the fertility decline), we simply ignore them. Thus, we define the quasi-failure rate as those leading to a live birth. We then estimate quasi-failure rates as births per month of use. Under the assumption of homogeneity (not exactly correct; see the Trussell and Kost's (1986) critique), this is exactly the right concept.

---

[2] Recall that we consider not sexually active and no method to be "methods."