Working Paper Series
Congressional Budget Office
Washington, D.C.

THE EFFECT OF A FIRST CHILD ON FEMALE LABOR SUPPLY: EVIDENCE FROM WOMEN SEEKING FERTILITY SERVICES

Julian P. Cristia
Congressional Budget Office
Washington, D.C.
E-mail: julian.cristia@cbo.gov

October 2006
2006-11

Abstract

Estimating the causal effect of a first child on female labor supply is complicated by the endogeneity of the fertility decision. That is, factors that trigger the decision to have a first child could also affect baseline labor supply; empirical approaches that do not account for this difficulty will yield biased estimates. This paper addresses this problem by focusing on a sample of women from the National Survey of Family Growth (NSFG) who sought help to get pregnant. After a certain period, only some of these women gave birth to a child. In this instance, fertility appears to be exogenous to labor supply in that women's employment during months prior to seeking help becoming pregnant is uncorrelated with subsequent fertility. Results using this strategy show that having a first child younger than one year old reduces female employment by 26.3 percentage points. Unlike previous studies, which found smaller effects when dealing with the endogeneity problem of fertility, estimates in this paper are close to ordinary least squares (OLS) estimates obtained using census data and to OLS and fixed-effects estimates from NSFG data.

1. Introduction[1]

Estimating the effect of fertility on female labor supply has been a longstanding problem in economics. Knowing how families optimize their labor supply decisions in response to the arrival of a child is important for several reasons. First, it is interesting to know how much of the increase in female labor supply since the World War II can be explained by delayed childbearing and reduced fertility (Goldin, 1990). Second, some researchers believe that the interruption of work attributable to childbearing is responsible for a significant proportion of the female-male wage gap (Goldin and Polachek, 1987; Gronau, 1988; Fuchs, 1989; Korenman and Neumark, 1992), and the size of the impact of childbearing on female labor supply is an important variable in that calculation. Third, if declines in labor supply after childbearing correspond to increases in child care time, then knowing the effect of childbearing on female labor supply will provide information about time inputs invested in the child (Stafford, 1987; Blau and Grossberg, 1992). Finally, and above all, economists have been interested in this question from a basic desire to know the quantitative importance of various determinants of female labor supply.

Hundreds of published studies have examined the relationship between fertility and female labor supply. However, as Browning (1992) notes in his literature review on this topic, "Although we have a number of robust correlations, there are very few credible inferences that can be drawn from them" (p. 1435). The key problem researchers face is that the fertility decision may be endogenous; therefore, the strong negative correlations found

[1]The author wishes to thank Julio Caceres Delpiano, Paul Cullinan, Mark Duggan, William Evans, Eugenio Giolito, Amy Harris, Sandra Hofferth, Arlene Holen, Joseph Kile, BeomSoo Kim, Noah Meyerson, Reed Olsen, Caroline Polk, John Sabelhaus, Seth Sanders, Jon Schwabish, John Skeen, Michael Simpson, Julie Topoleski, and Bruce Vavrichek for invaluable comments and suggestions.

between different measures of fertility and female labor supply cannot be interpreted as evidence of causal effects.

To overcome the type of criticism highlighted by Browning (1992), two strategies have exploited exogenous changes in family size in order to estimate the effect of fertility on female labor supply. The first strategy (Rosenzweig and Wolpin, 1980; Bronars and Grogger, 1994; Jacobsen, et al., 1998) used the fact that twins in the first birth represent an exogenous change in family size in order to estimate the effect of having a second child. The second strategy (Angrist and Evans, 1998) exploited parental preferences for mixed-sex siblings in order to estimate the effect of a third or higher order child.

Still, the question of how a first child affects female labor supply has not been previously addressed with a strategy that directly tackles the problem of the endogeneity of fertility. It could be argued that the effect of having a first child is the most important one, given that it applies to a vast majority of women, whereas the effect of having a second or higher order child only applies to a smaller subset of women.[2]

This paper examines a situation in which the problem of the endogeneity of fertility is minimized. In particular, I construct a sample of childless women who sought help with achieving pregnancy. At the time of seeking help, all of the women wanted to have a child; after a certain period, some of them gave birth, and others did not. I then compare the employment rates of women in the "treatment" group (i.e., those who gave birth to a child) with the controls (i.e., those who did not give birth).[3]

The contribution of this paper is that while analysis of twins and the preference for

[2]In the 1990 census, among women aged 45 to 55, 89.0 percent had at least one child, whereas 78.3 percent had at least two and 50.4 percent had at least three children.

[3]In this paper "treatment" refers to having a child.

mixed-sex siblings strategy, under certain conditions, can be used to identify the effect of a second or higher order child, the estimation strategy pursued here is able to identify the effect of a first child on female labor supply.

The strategy used in this analysis tackles the problem of fertility being an endogenous variable because all women wanted to have a child at the time they sought help. Early success in fertility treatment, however, is not expected to be completely random. Still, I provide several pieces of evidence suggesting that this strategy consistently estimates the parameter of interest. First, following Heckman and Hotz (1989), I find that women's employment, during months prior to seeking help becoming pregnant, is uncorrelated with subsequent fertility. Second, estimates of the parameter of interest are very robust to the set of covariates added to the main regression. Third, observable characteristics of the sample of childless women that sought help achieving a first pregnancy are quite similar to those of women who have their first child after age 18.

Using the exogenous assignment of children to women that seek help achieving pregnancy, I estimate that having a first child younger than one year old reduces female employment by 26.3 percentage points. These estimates are close to ordinary least squares (OLS) and fixed-effects estimates obtained from panel data from the National Survey of Family Growth (NSFG). They are also close to OLS estimates obtained using similarly defined samples from the 1980 and 1990 censuses. This finding is important because almost all previous studies that take into account the endogeneity of the fertility decision provide much smaller estimated effects than studies that assume exogenous fertility. Finally, I provide evidence of a reduction in the estimated short-term impact of childbearing on female labor supply of 40 to 50 percent between 1980 and 1990.

2. Previous Research

Interest in the effect of fertility on female labor supply is illustrated in the long list of studies that have focused on the issue. The studies can be classified into four groups, depending on how they tackle the problem of the endogeneity of the fertility decision. The first group is illustrated by the studies of Gronau (1973), Heckman (1974), and Heckman and Willis (1977), who assumed that fertility was exogenous and established a strong negative correlation between female labor supply and fertility.

A second group of studies (Cain and Dooley, 1976; Schultz, 1978; Fleisher and Rhodes, 1979) acknowledged the endogeneity of the fertility decision and tried to deal with the problem by estimating simultaneous equations models. The studies found a much smaller estimate when treating fertility as an endogenous variable than when treating it as an exogenous variable. The disadvantage of this approach is that it is difficult to find plausible exclusion restrictions that could identify the underlying structural parameters.

A third group of studies incorporated actual fertility as a regressor but added the lagged dependent variable (i.e. labor supply) to control for unobserved heterogeneity across women. Nakamura and Nakamura (1992) recommended this approach, and it has been used by a number of authors (Even, 1987; Lehrer, 1992). Although adding the lagged dependent variable can help control for unobserved heterogeneity, it still does not address the problem of the endogeneity of the fertility decision.

Finally, a fourth group of studies tackled the endogeneity of the fertility variable by exploiting exogenous sources of variation in family size. Rosenzweig and Wolpin (1980) first used this strategy by comparing labor supply of women who had twins at their first birth with that of women who had a single child. Subsequent studies by Bronars and Grogger (1994) and Jacobsen, et al. (1998) used the same strategy but managed to obtain more precise estimates by developing an algorithm to detect twin births using census data.

In the same spirit as the twins studies mentioned above, Angrist and Evans (1998) estimated the effect of a third or higher order child on female labor supply by exploiting the fact that parents typically prefer mixed-sex siblings. For a sample of couples with at least two children, they instrumented further childbearing (i.e., having more than two children) with a dummy variable for whether the sex of the second child matched the sex of the first. Because sex mix is virtually random, this strategy allows identification of the effect of a third or higher order child.

My work is most similar to this last group of studies because it uses the fact that the biology of reproduction is intrinsically stochastic to identify exogenous changes in fertility. Still, two main differences exist between the earlier studies and this paper. First, I estimate the effect of a first child on female labor supply, whereas the other studies estimate the effect of a second or higher order child. Second, whereas the other studies instrument fertility and then compute two steps least squares (2SLS) estimates, I tackle the endogeneity of fertility by focusing on a sample of women for whom fertility is plausibly exogenous and then estimate the impact with just OLS.

This paper is related to other strands of economic literature. First, its results shed light on a number of studies that have tried to explain the postwar rise in female labor supply (Mincer, 1962; Goldin, 1990). Second, it is related to a line of research that tries to establish the effect of childbearing-related withdrawal from the labor market on women's wages and earnings (Goldin and Polachek, 1987; Gronau, 1988; Fuchs, 1989; Korenman and Neumark, 1992; Miller, 2005).[4] Last, it is linked with studies focusing on how maternal work affects children's outcomes (Stafford, 1987; Desai, et al., 1989; Blau and Grossberg, 1992).

3. Background: The Reproductive Process and Infertility

Healthy couples having intercourse regularly have only a 20 percent chance of conceiving during a month. This statistic implies that about 26 percent of healthy couples will not have conceived after six months of unprotected sex; this number falls to about 7 percent after 12 months. As a result, couples are recommended to start fertility testing and treatment only after 6 to 12 months of trying to conceive without success. The medical community typically defines a couple as infertile if they have not conceived after 12 months of unprotected sex.[5] The National Center for Health Statistics (NCHS) estimated that in 1995 there were 2.1 million infertile married couples in the United States and that 6.1 million women aged 15 to 44 had an impaired ability to have children (Abma, et al., 1997).

Medical researchers have identified a number of factors (besides the conditions mentioned above) that affect the fertility prognosis of a couple. The woman's age, education, smoking status, consumption of recreational drugs, and obesity, as well as sexual frequency, are important predictors of the probability of conception (Baird and Wilcox, 1985; Dunson, et al., 2004).

Given the stochastic nature of the reproduction process, physicians usually start treatment with simple and inexpensive procedures (e.g., advice and testing) and only start using more invasive and expensive procedures as the simple procedures prove unsuccessful.

[4] As in this paper, Miller (2005) exploited biological fertility shocks. I aim to estimate the impact of child-bearing on female labor supply, whereas Miller was primarily concerned about how changes in the age at first birth affect long-run earnings and future wages.

[5] For example, see Mosher (1982).

For example, physicians typically recommend in vitro fertilization methods only after all other options have been exhausted or if they strongly believe that less invasive procedures will be unsuccessful.

4. Data

This paper uses data from the National Survey of Family Growth (NSFG), a survey conducted by the National Center of Health Statistics in 6 cycles (1973, 1976, 1982, 1988, 1995, and 2002). Cycles 1 to 5 were conducted at the homes of a nationally representative sample of women aged 15 to 44. Cycle 6 also sampled men aged 15 to 44. The main purpose of the surveys was to provide reliable national data on marriage, divorce, contraception, fertility, and the health of women and infants in the United States.

Data from the NSFG Cycle 5 were chosen for this paper because they provide retrospective information about births, pregnancies, use of fertility services, demographic characteristics, and the complete work history for each individual.[6] In particular, the data provide the month in which each woman sought help for the first time to achieve pregnancy, information that is critical for the strategy pursued in this paper. Other important variables included are age, race, ethnicity, educational attainment, school enrollment, and smoking history. The survey also reports data on each full-time and part-time employment spell.

The NSFG Cycle 5 used a multistage sampling design that oversampled Hispanic and black women. It took place between January and October 1995, and the response rate was 79 percent. A total of 10,847 women were interviewed.

[6] Other cycles included all needed information except monthly employment status for each woman. I cannot run this analysis without this information because I compare employment 21 months after each woman sought help becoming pregnant.

Data on fertility and employment are collected retrospectively. Although this type of design has limitations, Teachman, et al. (1998) found the NSFG Cycle 5 data to be of high quality. They concluded that the employment information matches the Current Population Survey (CPS) data reasonably well, although the data on employment spells have not been validated using external records.

5. Empirical Strategy, Parameter of Interest, and Sample Construction

5.1 Empirical Strategy

A hypothetical social experiment aimed at estimating the causal effect of childbearing on female labor supply would recruit women who wanted to have a child and then assign a child to women in the treatment group while not assigning a child to a second group (the control group).[7] Given the stochastic nature of conception, this type of experiment can be approximated. To start, we need a group of women who want to conceive a baby. Second, some of the women should receive babies in a way that is uncorrelated with baseline employment. Third, we need to observe female labor supply for both groups of women for a certain time after they start trying to conceive.

I aim to mimic the ideal social experiment and fulfill the three aforementioned conditions by focusing on the following situation. I construct a sample of women who sought help to have a first child (called the HELP sample). Because women in this sample sought help to achieve pregnancy at different points in time, I normalize time by the month in which they sought help for the first time (denoted as Month 0). Next, I classify the women according to whether they had given birth to a child by Month 21. In this way, I obtain two

[7] To be precise, this experiment estimates the effect of having a child on female labor supply for women who wanted to have a child, not for all women.

groups of women: treatment and control. Finally, I compare employment rates of the two groups of women in Month 21 to estimate the causal effect of having a first child younger than one year old on female labor supply.

I compare employment in Month 21 instead of other months for several reasons. First, at the time of my analysis, 97 percent of babies born are currently younger than one year old, making the definition of the treatment effect more precise. Second, using a longer horizon could allow some women in the treatment group to have additional children, which would complicate the analysis.[8] Third, as time since women sought help increases, those who are unsuccessful at conceiving may adopt a child. Finally, looking at a shorter time span, it is more plausible that the women received similar types of infertility treatments (e.g., in vitro fertilization treatments typically are not considered an option in the first 12 months after seeking help achieving pregnancy).

A potential problem with this empirical strategy arises if women in the control group adopt a child or start cohabitating with or marry someone with children. In the treatment evaluation literature, this behavior is denoted as "substitution bias", and it represents a situation in which control group members receive close substitutes for the treatment in question (see Heckman and Smith (1995, pp. 22-24). In the context of this paper, treatment is defined as having a natural birth, and a close substitute is adopting a child (or acquiring a stepchild). Even though substitution bias can be a problem in certain social experiments, it is not in this case.[9] Only 2.7 percent of women in the control group adopted or acquired a

[8]By Month 21 only six women had two children: Five had given birth to twins initially; only one had given birth twice.

[9]In the case of the experimental evaluation of the Job Training Partnership Act (JTPA), Heckman and Smith (1995) noted that 32 percent of control group members self-reported receiving training from other sources over the 18 months following random assignment.

stepchild in the 21 months after they sought help to become pregnant (and only 0.5 percent in the treatment group did so).

5.2 Parameter of Interest

In this study, the parameter of interest is the average impact of having a first child younger than one year old on female labor supply for women who want to have a child. Note that the study does not provide an estimate of the effect of having a first child for women whose child is unwanted. All the same, the parameter of interest applies to a fairly large population. Henshaw (1998), using data from the NSFG Cycle 5, found that 69 percent of births were planned among women aged 15 to 44 in 1994.

Throughout this study, I focus only on the short-term effects of having a first child (i.e., the estimated effect of having a child younger than one year old). It is clear that other treatment effects are worthy of attention; however, for reasons already discussed, the strategy used in this study is best suited for estimating this treatment effect.

Finally, an estimate of the impact of having a first child younger than one year old is important for a number of reasons. First, as mentioned above, this effect applies to a much wider population than estimates that focus on the effect of a second or higher order child. Second, the consensus is that the short-term effects of childbearing on female labor supply are substantially larger than the long-term effects (Browning, 1992). Thus, knowing the short-term effects is useful because it gives an upper bound for the long-term effects. Third, Shapiro and Mott (1994) provide strong evidence that labor force status following the first birth is an important predictor of lifetime work experience. This finding implies that changes in the estimated short-term impact of having a first child on female labor supply could predict a substantial change in total lifetime work experience for women. Finally, using this empirical strategy I can compare the estimated effects obtained when tackling the endogeneity problem (i.e., using the HELP sample) with estimates from strategies that do not tackle this problem (e.g., OLS on census data).

5.3 Sample Construction

The HELP sample includes childless women who sought help with becoming pregnant when aged 19 to 38. The age restriction in the sample is due to two reasons. First, the results obtained in the HELP sample are compared to those from census and NSFG samples, and an age restriction is needed in constructing these samples in order to select women at risk of having children. Second, work information is only reported for women aged 18 and older, and I want to know women's employment status one year before seeking help with becoming pregnant. Women that sought help in the 21 months preceding the interview are dropped from the HELP sample, because it is not possible to observe their child and labor status at this time.

Women seeking help with becoming pregnant are identified as those who answer affirmatively the question "Have you or your husband ever been to a doctor or other medical provider to talk about ways to help you become pregnant?" The wording of the question allows identification of a large group of women who wanted to have children but were unsuccessful after trying for certain time. The fact that women that just talked to their medical provider about ways to help with becoming pregnant are included in the HELP sample explains why, as it will be seen later, women in the HELP sample are fairly representative of women who have their first child when aged 19 to 38.

Table 1 presents the algorithm used in order to construct the HELP sample. The table shows that only 499 observations are included in the empirical analysis, a fact that may seem to be an important limitation for this study. As shown in Section 6, however, I precisely estimate the relevant coefficient of the effect of having a first child on female labor supply.

The basic empirical strategy of this paper is based on comparing women in the HELP sample who had had a baby by Month 21 with those who did not. To identify the two groups of women, I defined a variable called *AnyChildren21* which equals one if the woman had a baby by Month 21 and zero if she did not. In this setting, women from the HELP sample for whom *AnyChildren21* equals one are in the treatment group and those for whom *AnyChildren21* equals zero are in the control group.

The plausibility of the proposed empirical strategy rests on the assumption that treatment is not correlated with baseline labor supply. However, in some scenarios this assumption will not hold. For example, if women married to high earner men have a higher probability of success (through access to better fertility treatments) and tend to have a lower attachment to the labor market (due to an income effect), then the effect of fertility on female labor supply will be underestimated. To assess the plausibility of the empirical strategy, I take two steps. First, I compare summary statistics on covariates for the treatment and control groups in order to check for evidence of selection. Clearly, it is not possible to check for selection on unobservable factors, but lack of evidence of selection on observable factors gives assurance that treatment can be taken as exogenous to baseline labor supply.[10] Second, I compare employment rates for the treatment and control groups prior to seeking help (the

[10]For example, if education levels are similar across the treatment and control groups, then the hypothesis that high income women have higher probability of success is undermined.

results are presented in Section 7).

Descriptive statistics for women in the treatment and control groups are presented in Table 2. In the NSFG Cycle 5, respondents were asked about all of their employment spells, and I use those responses to construct three employment variables. The variables *Employed21* and *Employed0* are dummy variables that equal one if the respondent was employed in Months 21 and 0, respectively. Similarly, *Employed_12* represents labor status in Month –12 (i.e., 12 months before the woman sought help for the first time).

Although employment rates in Months 0 and –12 are similar between the treatment and control groups, employment rates differ by 25.3 percentage points in Month 21. Moreover, observable characteristics in Month 0 for treatment and control women are quite similar. As shown in Table 2, differences in means of key covariates between the treatment and control groups are only statistically significant at the 5 percent significance level for the dummy variables for Hispanic and smoking.[11]

A potential caveat for the strategy pursued in this paper is that, as typically is the case in social and medical experiments, the sample involved in the experiment may not be representative of the population of interest. To gauge the potential severity of this problem, Table 3 compares descriptive statistics of women in the HELP sample with those of women in the NSFG who had at least one child. For women in the HELP sample, time-varying variables (*Age*, *Year*, *Employed_12*, *Education*, *Married*, *Smoke*) are measured at the time they first sought help achieving pregnancy, whereas for NSFG women the variables are measured at the time of first birth. The second column of Table 3 presents statistics for the set of women in the NSFG who had their first child when aged 19 to 38 because that was the age

[11] In Section 7, I explore more deeply which variables predict fertility by Month 21.

range of women in the HELP sample.

Comparing the second and third column of Table 3, we see that women in the HELP sample tend to be older and more educated and to have higher employment, marriage, and smoking rates than women from the NSFG that were aged 19 to 38 when they had their first child. A lower proportion of HELP women are Hispanic or black than NSFG women. Still, basic statistics for the HELP sample are not very different from those of their counterparts in the NSFG. The last column of Table 3 presents basic statistics for the HELP sample when observations are reweighted to match the distribution by age and year groups for 19- to 38-year-old NSFG women with children. This adjustment makes the proportion of Hispanic and black women similar across the two samples, and it brings mean education closer.

Figure 1 compares the age distribution of women in the NSFG who gave birth when aged 19 to 38 to the age distribution of women in the HELP sample. The difference in mean age across the two groups is driven primarily by the group of women aged 19 to 21. This difference can be explained by some women in the NSFG group having unplanned children and by the fact that really young women tend to delay their decision to seek help with achieving pregnancy.

6. Results

This section presents the main results of the empirical analysis. In essence, I compare employment rates in Month 21 for treatment and control women in the HELP sample. The econometric model is represented by this simple OLS equation:[12]

[12]Marginal effects results for probit and logit models are similar to those obtained using OLS.

$$Employed21_i = \alpha + \beta \, AnyChildren21_i + \gamma \, X_i + u_i \quad (1)$$

where the vector of covariates includes black and Hispanic dummy variables; an indicator for insurance coverage of infertility treatments; year in which the women sought help for the first time; and the following variables measured in Month 0: age, smoking status, and years of education.

To gauge the potential importance of the problem of not having data on certain variables that may be simultaneously correlated with conception and labor supply, I run a number of regressions including separate sets of covariates. If the results were sensitive to the set of covariates added to the regression, they would raise some doubts as to whether the identification strategy consistently estimates the parameter of interest. Table 4 presents the regression results.

In the model that includes all covariates (Column 4), I estimate that having a first child younger than one year old decreases female employment by 26.3 percentage points. The results indicate that the estimated impact is remarkably robust to the set of covariates included in the regression. In particular, the estimated effect in a model with no covariates (Column 1) is –0.253. That is, including the entire set of covariates, the estimated coefficient changes by just 1 percentage point, or 4 percent of the estimated impact.

Column 5 presents linear probability estimates when observations are reweighted to match the age-year distribution for the sample of mothers in the NSFG who gave birth to their first child when aged 19 to 38. The estimated impact is similar to the one obtained from original NSFG weights (Column 4); this finding is evidence that the obtained estimates could be generalized to the target population. Finally, in Column 6 the model is augmented to check for varying treatment effects by age of mother and year in which she sought help with achieving pregnancy. Although the treatment effect does not significantly change with age, the results suggest that the short-term effects of childbearing have decreased over time (this issue is examined more deeply in Section 8.2).

Women who have a child decide not only whether to have a job (the extensive margin) but also how many hours to work (the intensive margin). Unfortunately, the NSFG does not provide retrospective information on hours worked for women in the sample. Still, it provides information about whether a woman was working full time or part time and the availability of maternity leave. Using this information, work status is determined among four categories (full time, part time, maternity leave, and no job). Table 5 presents multinomial logit regression results for the impact on work status of having a first child. Having a child younger than one year old reduces the probability of working full time by 43.1 percentage points and it raises the probability of being in the other three categories. Interestingly, the increase in the probability of working part time is quite small (4.8 percentage points).

7. Robustness of the Empirical Strategy

This section explores the robustness of the empirical strategy pursued. First, I try to identify which covariates can predict treatment and how much of the variation in the fertility variable is explained by the other variables. Second, I test whether pretreatment differences exist in the outcome variable (i.e., employment) for the treatment and control groups. Finally, I check how robust the results are to changes in the specification of the econometric model.

To start, I explore which variables in the data set predict early fertility success in the

HELP sample. Table 6 shows that, as documented in the medical literature, women's age is one of the most important predictors of fertility. In this linear probability model, an increase in one year in the age of the woman decreases her expected probability of having a child by 1.6 percentage points. Smoking, also documented in the medical literature as having an effect on fertility, is a significant negative predictor of fertility success. Finally, Hispanic and highly educated women are more likely to be successful.

Even though several variables can predict fertility, note that the adjusted R^2 is only 4.3 percent and that much of the variation in the fertility variable remains unexplained in this model.

Next, I turn to the issue of whether the significant differences in employment between treatment and control women in Month 21 can be interpreted as the effect of treatment or as just heterogeneity in labor market attachment between groups. This test is important for the empirical strategy pursued in the paper. Before the presentation of the regression results, it is useful to look at Figure 2, which plots employment rates of the treatment and control groups for months –12 to 21. Employment rates of both groups are quite similar for months –12 to 0, but they start diverging around Month 3 and are far apart by Month 21. The continuous decline in employment rates for the treatment group corresponds to the fact that as time goes by, additional women give birth until by Month 21 all had given birth.

Table 7 presents the results of regressions of employment status in Month 0 (*Employed0*) on *AnyChildren21*. Several specifications are run, in which I control for different sets of covariates to gauge the robustness of the results. The main conclusion from this table is that no statistically significant differences in employment rates in Month 0 exist between the treatment and control groups.[13]

Finally, a number of additional regressions are run to check whether the results are robust to changes in the specification. First, I again run the regressions whose results are presented in Table 4, but I add an indicator for pregnancy in Month 21. Second, the main independent variable *AnyChildren21* is replaced with another variable that equals the total number of children born per woman by Month 21. Third, I replace *AnyChildren21* with two indicators for having one child or two children in Month 21, respectively. Fourth, instead of running linear probability models of *Employed21* on *AnyChildren21,* I run probit and logit models using the same set of variables as in Table 4. In all cases, the estimated effects are similar to those reported in Section 6.

8. Comparison to Estimates from NSFG and Census Data

In his survey of the effect of children in the household, Browning (1992) concludes that studies that treat fertility as exogenous typically find significantly larger effects of fertility on female labor supply than those that treat it as endogenous and estimate simultaneous equations models. Angrist and Evans (1998) provide further evidence about this argument because they report that their 2SLS estimates of the impact of having more than two children on female labor supply are statistically significantly smaller than their OLS estimates. This section compares estimates obtained using the HELP sample with estimates from similarly defined samples but without restricting them to women who sought help to become pregnant.

A problem faced in trying to replicate the HELP sample is that this data set includes

[13] Similar results are obtained when regressing employment at 12 months before women sought help to achieve pregnancy on *AnyChildren21*.

observations of fertility and labor supply for women who sought help to become pregnant at different points in time. The implication is that to replicate the results from the HELP sample, I should construct comparable data sets with observations for women at different points in time (i.e., panel data or repeated cross-sections). I therefore compare estimates from the HELP sample to estimates from panel data from the NSFG (in Section 8.1) and to estimates from census data for 1980 and 1990 (in Section 8.2).

8.1 Comparison to Estimates from NSFG Panel Data

I construct a panel data set from the NSFG Cycle 5 (i.e., NSFG panel data) following requirements similar to those used to construct the HELP sample. The unit of observation in this panel data is a woman-month. An observation is included in the NSFG panel data if the woman was aged 21 to 40 in that month, was childless or had children younger than one year old, and was cohabitating or married.[14]

Because the HELP sample corresponds to a cross-section, to use the same source of variation when estimating both models, I construct a panel data set (i.e., HELP panel data) including, for each individual in the HELP sample, observations for months –12 to 33 (remember that Month 0 corresponds to when the women first sought help with achieving pregnancy). Because the goal is to estimate the impact of having a child younger than one year old, monthly observations for a woman are dropped when her child is older than age 1. Finally, for women who did not have a baby by Month 21, monthly observations of later months are dropped if they gave birth to a child.[15]

[14]This age restriction is chosen because women in the HELP sample are aged 19 to 38 years old at Month 0 and then in Month 21 (when employment by fertility status is computed), almost all of them are 21 to 40 years old.

[15]Defining the sample in this way ensures a balanced distribution of women with respect to their children's age in months.

Table 8 presents summary statistics for the HELP panel data and the NSFG panel data. Mean values for key variables are similar and are only statistically significantly different for proportion employed and with children, calendar year, and children's age in months. Still, employment rates are not significantly different across samples once I condition for fertility status. With respect to differences in the proportion of women who have a child, this fact should be expected given that everyone in the HELP panel data did not have children for Month –12 up to (at least) Month 7.

Linear probability estimates of the impact of having at least one child (younger than one year old) on the probability of having a job are presented in Columns 1 and 3 of Table 9. In the first column, results are presented for the model estimated using the HELP panel data. The main independent variable is *AnyChildren*. The estimated impact (0.260) is similar to the estimates obtained in Section 6. In the third column, results are presented for the same model estimated using the NSFG panel data. The key finding from comparing Columns 1 and 3 is that the estimated impact using the NSFG panel data (0.259) is notably similar to the one obtained using the HELP panel data.[16]

To gauge the robustness of the results, I estimate fixed-effects models using both panel data sets. Results are presented in Columns 2 and 4 of Table 9. For the HELP panel data, the estimated impact slightly decreases in absolute value to 0.234. In the case of the NSFG panel data, the estimated impact decreases in absolute value to 0.216. This result provides some evidence that women who have children tend to have lower employment rates in the months previous to become pregnant. Still, both estimates are similar, and the *t*-value of the test of equality of coefficients is just -0.46.

[16] The t-value of the test of equality of coefficients is 0.00.

Finally, I compare the estimated impact of having a child on work status (working full time, part time, maternity leave, and no job) for the two panel data sets. As Table 10 shows, estimates of the marginal effect of having a child on the probability of being in each of the four work status categories are strikingly similar across the two data sets.

The fact that estimates from the HELP panel data are so similar to those from the NSFG panel suggests that the endogeneity problem of fertility is not severe with regard to its effects on biasing estimates of treatment effects. Another explanation is that endogeneity does create bias on estimates but the samples yield similar results because the differences in treatment effects across samples compensate for the bias (e.g., there may a positive bias in estimates on NSFG panel data, but the true treatment effect in the NSFG panel data is larger than in the HELP panel data). Given that statistics on observable characteristics across the two samples are so similar, however, the difference in treatment effects across samples should be based entirely on differences in unobservables, making the lack of endogeneity a more plausible explanation.

8.2 Comparison to Estimates from 1980 and 1990 Census Data

In the HELP sample, fertility and other covariates are observed between 1972 and 1995. On average, those variables are observed in 1986, and the 10th and 90th percentiles correspond to years 1978 and 1993, respectively. To construct samples comparable to census data, women in the HELP sample are assigned to two new samples, the EARLY and LATE HELP samples, depending on whether they sought help to become pregnant before or after 1985.[17]

[17] To construct two samples with roughly the same number of observations, the threshold year is 1985.

Using data from the 5-percent 1980 and 1990 Census Public Use Micro Samples (PUMS) I construct two samples (denoted as the 1980 and 1990 Census samples, respectively).[18] The Census samples include married women aged 21 to 40, who are childless or have children younger than one year old. To capture women who are "at risk" of having a child and make the Census samples comparable to the HELP samples, only married women are kept in the Census and HELP samples.[19]

Table 11 presents descriptive statistics for these samples EARLY HELP, LATE HELP, 1980 Census and 1990 Census. In the case of the HELP samples, the variable *Employed* equals one if the woman had a job in Month 21. For the Census samples, it equals one if the woman had a job during the week previous to the survey. The variables *AnyChildren*, *Age*, *Education*, *Hispanic* and *Black* are similarly defined in the four samples, and are all measured in Month 21 (for the HELP samples) or at the time of the survey (for the Census samples). *AnyChildren* equals one if the woman had at least one child. *Education* corresponds to the number of years of education. Finally, *Black* and *Hispanic* are dummy variables that equal one if the woman is in either group.

The results in Table 11 suggest that the 1980 and 1990 Census samples can be considered as reasonable comparison data sets for the EARLY and LATE HELP samples, respectively. Women in the 1980 Census were surveyed in April 1980, where those in the EARLY HELP sample were observed, on average, in June 1981. Similarly, women in the 1990 Census were surveyed in April 1990, while those in the LATE HELP sample were observed, on average, in January 1991. Moreover, basic statistics on education and

[18]For information about the PUMS, see Ruggles, et al., 2004.

[19]Results obtained by dropping the requirement that women in the HELP and Census samples be married are similar to those presented in this subsection.

proportion black and Hispanic are remarkably close. Conversely, the proportion of women who have a child is significantly higher in the HELP samples. This finding should be expected, given that presumably all women in the HELP samples wanted to have children. Finally, employment rates in the HELP samples, conditional on fertility status, are around 10 percent higher than in the Census samples (perhaps because employment is not defined exactly the same way in the NSFG as in the census).

Linear probability estimates of the impact of having a child younger than one year old on employment are presented in Table 12. Comparing Columns 1 and 2, we can see that the estimated impact is remarkably similar in the EARLY HELP sample and the 1980 Census sample (0.372 vs. 0.365). Similarly, the estimated impact is also quite close when comparing the LATE HELP sample and the 1990 Census sample (0.182 vs. 0.228). In both cases, *t*-tests of differences in the estimated impact cannot be rejected.

From this set of results two important conclusions can be drawn. First, the estimated effects for the HELP sample for which I can identify an exogenous change in the fertility variable are nearly identical to the estimates obtained using OLS on comparable samples from census data, for which I do not control for the endogeneity of the fertility variable. They are also close to estimates obtained using panel data from the NSFG, as described in the previous section. Second, evidence suggests a significant reduction (about 40 to 50 percent) in the short-term impact of childbearing on female labor supply in the 1980 to 1990 period.

9. Conclusions

This paper explores the issue of the causal effect of childbearing on female labor supply. This task is complicated by two factors. First, some researchers believe that women who have children at a certain age may have different baseline labor supply from women with similar observed characteristics who do not have children (Browning, 1992). This expected unobserved heterogeneity across groups suggests the existence of bias in simple cross-section comparisons. As noted by Nakamura and Nakamura (1992), we can try to deal with this problem by adding the lagged values of labor supply to regressions of current labor supply on number of children.

A second problem, however, complicates the estimation of the effect of childbearing on female labor supply, and it cannot be solved just by using longitudinal data. The problem stems from the fact that the fertility decision may be endogenous to the woman and influenced by potential labor supply. Several studies, starting with Rosenzweig and Wolpin (1980), use the fact that having twins in the first birth changes (at least temporarily) family size compared to not having twins. In order to find exogenous variation in the fertility decision, Angrist and Evans (1998) exploit the fact that parents typically prefer mixed-sex siblings. Even though these papers have made a major contribution in answering the question posed, they are only able to estimate the effect of having a second or higher order child.

To deal with the problems of unobserved heterogeneity and endogeneity, I restrict my attention to a group of women who sought help to achieve pregnancy. In this sample, all the women wanted to have children, so the problem of endogeneity is minimized. Moreover, because a major proportion of the fertility variable is random, results likely will not be contaminated by unobserved heterogeneity across groups. In fact, the attractiveness of the strategy pursued here is that, by focusing on this sample of women, I mimic a hypothetical social experiment in which, for a group of women who want to have a child, some women are assigned children while others are not. Evidence favors the empirical strategy pursued: my results show that women's employment, during months prior to seeking help becoming pregnant, is uncorrelated with subsequent fertility.

Using this empirical strategy, having a first child younger than one year old reduces female labor supply by 26.3 percentage points. Interestingly, evidence strongly suggests that the estimates obtained using this strategy (which tackles the problem of the endogeneity of fertility) are similar to estimates derived from approaches that assume the exogeneity of fertility.

Given that studies that assume the exogeneity of fertility typically find larger effects of fertility on female labor supply than those that treat it as endogenous, a natural extension of this paper would be to attempt to understand why my empirical strategy reaches a different conclusion. One potential explanation is that there is not much selection bias when focusing on women aged 19 and older wanting a first child. Although Hotz, et al. (2005) found important differences in observable characteristics when comparing teen mothers to childless teenagers, for the NSFG and census samples constructed in this paper, observable characteristics of women with and without children are quite similar.

Another interesting question that this paper leaves unanswered is why fertility and baseline employment seem to be uncorrelated. Many hypotheses may predict the opposite. For example, my strategy restricted the sample to women who are homogeneous in that all wanted to have a child at certain point in time, but clearly they could differ in how much they wanted to have a child, which in turn could be correlated with baseline labor force attachment.

A potential explanation for subsequent fertility being uncorrelated with baseline labor supply could be related to the fact that women in the HELP sample typically wait a number of months before seeking help to achieve pregnancy. This “waiting” pattern could reduce the heterogeneity of women in the sample with respect to their baseline probability of having a child. Women with high probability of having a child achieve pregnancy right away and then do not seek help to become pregnant. Because individuals in the sample have similar probabilities of having a child, we approach the ideal situation of random assignment, which is characterized as one in which all individuals have *equal* probability of being treated. If evidence is found suggesting that “waiting” is a successful empirical strategy in the sense that it increases the similarity between the treatment and control groups, then the same strategy could be applied to other evaluation problems in which dynamic assignment of individuals to treatment occurs.

References

Abma, Joyce, Anjani Chandra, William Mosher, Linda Peterson, Linda Piccinino (1997). " Fertility, Family Planning, and Women's Health: New Data from the 1995 National Survey of Family Growth. National Center for Health Statistics." *Vital Health Stat,* Series 23, No. 19.

Angrist, Joshua, and Williams Evans (1998). "Children and their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." *American Economic Review,* Vol. 88, No. 3 (June), pp. 450-77.

Baird, Donna, and Alex Wilcox (1985). "Cigarette Smoking Associated with Delayed Conception." *Journal of the American Medical Association,* Vol. 253, No. 20 (September), pp. 2979-83.

Blau, Francine, and Adam Grossberg (1992). "Maternal Labor Supply and Children's Cognitive Development." *Review of Economic and Statistics,* Vol. 74, No. 3 (August), pp. 474-81.

Bronars, Stephen, and Jeff Grogger (1994). "The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment." *American Economic Review,* Vol. 84, No. 5 (December), pp. 1141-56.

Browning, Martin (1992). "Children and Household Economic Behavior." *Journal of Economic Literature,* Vol. 30, No. 3 (September), pp. 1434-75.

Cain, Glen, and Martin Dooley (1976). "Estimation of a Model of Labor Supply, Fertility and Wages of Married Women." *Journal of Political Economy,* Vol. 84, No. 4 (August), pp. S179-99.

Desai, Sonalde, Lindsay Chase-Lansdale and Robert Michael (1989). "Mother or Market? The Effects of Maternal Employment on the Intellectual Ability of 4-Year Old Children." *Demography,* Vol. 26, pp. 545-61.

Dunson, David, Donna Baird, and Bernardo Colombo (2004). "Increased Infertility with Age in Men and Women." *Obstetrics and Gynecology,* Vol. 103, No. 1 (January), pp. 51-6.

Even, William (1987). "Career Interruptions Following Childbirth." *Journal of Labor Economics,* Vol. 5, No. 2 (April), pp. 255-77.

Fleisher, Belton, and George Rhodes (1979). "Fertility, Women's Wage Rate and Labor Supply." *American Economic Review,* Vol. 69, No. 1 (March), pp. 14-24.

Fuchs, Victor (1989). "Women's Quest for Economic Equality." *Journal of Economic Perspectives,* Vol. 3, No. 1 (Winter), pp. 25-41.

Goldin, Claudia (1990). *Understanding the Gender Gap*. New York: Oxford University Press.

Goldin, Claudia, and Solomon Polachek (1987). "Residual Differences by Sex: Perspectives on the Gender Gap Earnings." *American Economic Review,* Vol. 77, No. 2 (May), pp. 143-51.

Gronau, Reuben (1973). "The Intrafamily Allocation of Time: The Value of the Housewives' Time." *American Economic Review,* Vol. 63, pp. 634-51.

Gronau, Reuben (1988). "Sex-Related Wage Differentials and Women's Interrupted Careers – the Chicken or the Egg." *Journal of Labor Economics,* Vol. 6, No. 3 (July), pp. 277-301.

Heckman, James (1974). "Shadow Prices, Market Wages and Labor Supply." *Econometrica,* Vol. 42, pp. 679-94.

Heckman, James, and V. Joseph Hotz (1989). "Choosing among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: the Case of Manpower Training." *Journal of the American Statistical Association,* Vol. 84, No. 408 (December), pp. 862-74.

Heckman, James, and Jeffrey Smith (1995). "Assessing the Case for Social Experiments." *Journal of Economic Perspectives,* Vol. 9, No. 2 (Spring), pp. 85-110.

Heckman, James, and Robert Willis (1977). "A Beta-Logistic Model for the Analysis of Sequential Labor Force Participation by Married Women." *Journal of Political Economy,* Vol. 85, No. 1 (February), pp. 27-58.

Henshaw, Stanley (1998). "Unintended Pregnancy in the United States." *Family Planning Perspectives,* Vol. 30, No. 1 (January), pp. 24-29.

Hotz, V. Joseph, Susan McElroy, and Seth Sanders (2005). "Teenage Childbearing and its Life Cycle Consequences: Exploiting a Natural Experiment." *Journal of Human Resources,* Vol. 40, No. 3 (Summer), pp. 683-715.

Jacobsen, Joyce, James Wishart Pearce III, and Joshua Rosenbloom (1999). "The Effects of Childbearing on Married Women Labor Supply and Earnings: Using Twin Births as a Natural Experiment." *Journal of Human Resources,* Vol. 34, No. 3 (Summer), pp. 449-74.

Korenman, Sanders, and David Neumark (1992). "Marriage, Motherhood and Wages." *Journal of Human Resources,* Vol. 27, No. 2 (Spring), pp. 233-55.

Lehrer, Evelyn (1992). "The Impact of Children on Married Women's Labor Supply: Black-White Differentials Revisited." *Journal of Human Resources,* Vol. 27, No. 3 (Summer), pp. 422-44.

Miller, Amalia (2005). "The Effect of Motherhood Timing on Career Path." Manuscript, Department of Economics, University of Virginia.

Mincer, Jacob (1962). "Labor Force Participation of Married Women: a Study of Labor Supply." In H. Gregg Lewis, Ed., *Aspects of Labor Economics.* Princeton, NJ: Princeton University Press, pp. 63-105.

Mosher, William (1982). "Infertility Trends Among U.S. Couples: 1965-1976." *Family Planning Perspectives,* Vol. 14, No. 1 (January), pp. 22-7.

Nakamura, Alice, and Masao Nakamura (1992). "The Econometrics of Female Labor Supply and Children." *Econometrics Review,* Vol. 11, No. 1, pp. 1-71.

Rosenzweig, Mark, and Kenneth Wolpin (1980). "Life-Cycle Labor Supply and Fertility: Causal Inferences from Households Models." *Journal of Political Economy,* Vol. 88, No. 2 (April), pp. 328-48.

Ruggles, Steven, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander (2004). "Integrated Public Use Microdata Series: Version 3.0" [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor].

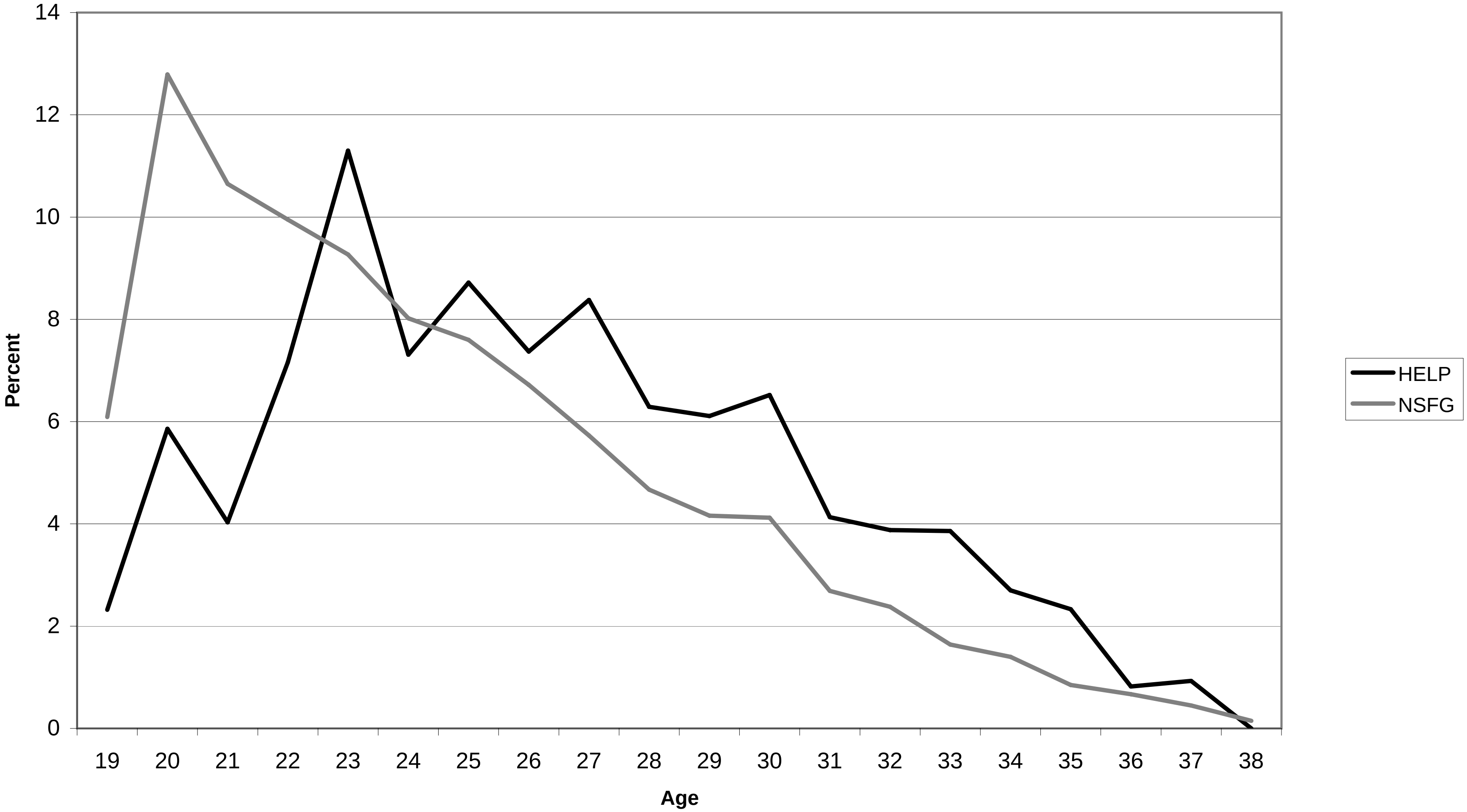
Schultz, T. Paul (1978). "The Influence of Fertility on Labor Supply of Married Women: Simultaneous Equation Estimates." In Ronald Ehrenberg, Ed., *Aspects of Labor Economics.* Greenwich, CT: JAI Press, pp. 273-351.

Shapiro, David, and Frank Mott (1994). "Long-Term Employment and Earnings of Women in Relation to Employment Behavior Surrounding the First Birth." *Journal of Human Resources,* Vol. 29, No. 2 (Spring), pp. 248-275.

Stafford, Frank (1987). "Women's Work, Sibling Competition and Children's School Performance." *American Economic Review,* Vol. 77, No. 5 (December), pp. 972-80.

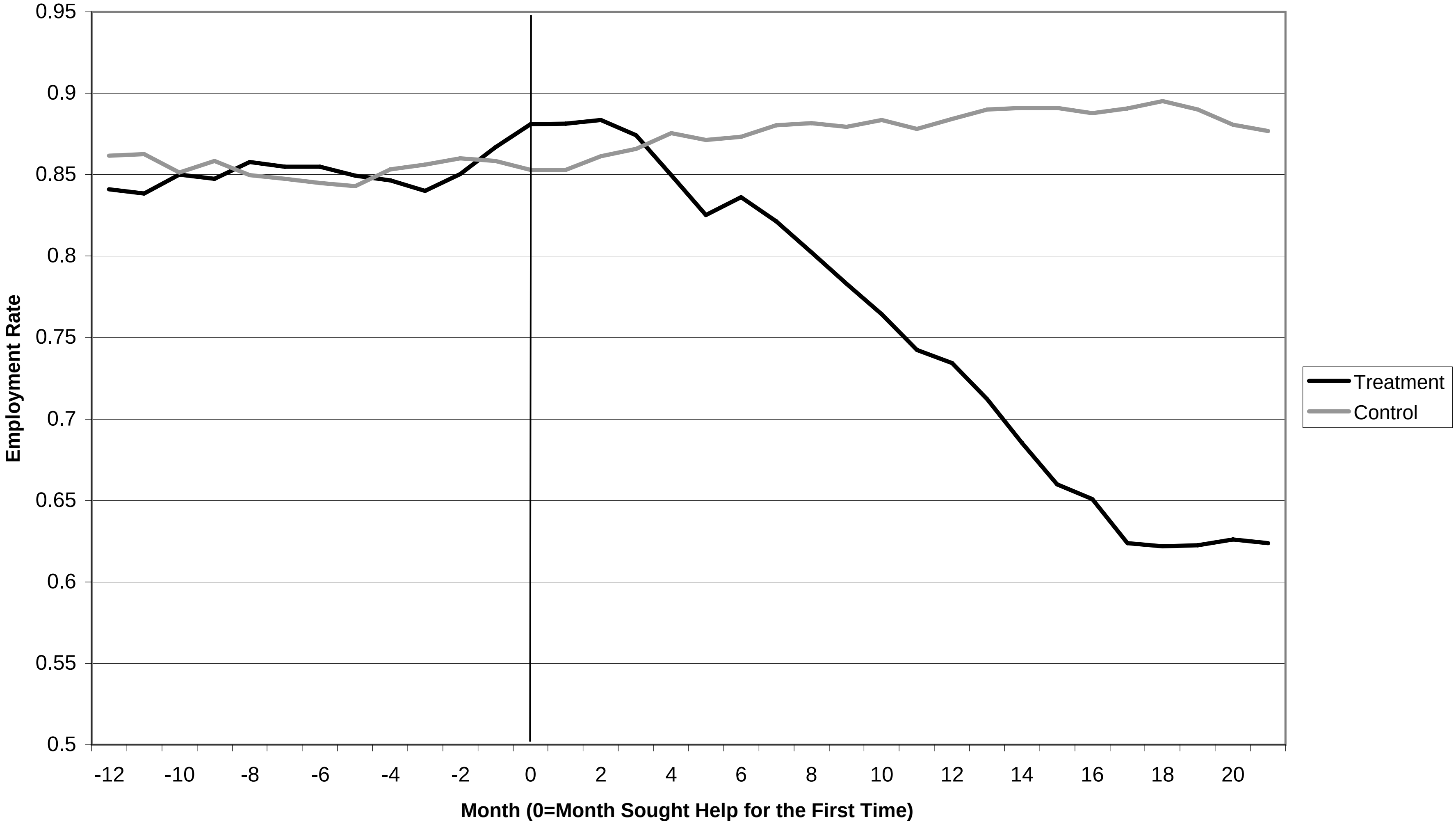
Teachman, Jay, Lucky Tedrow and Kyle Crowder. (1998) "Evaluation of the 1995 NSFG." Manuscript, Sociology Department, Western Washington University.

Figure 1: Age Distribution of Women in the HELP Sample and Childbearing Women in the NSFG Sample



Note: HELP=age at which first sought fertility services, women in HELP sample; NSFG=age at first birth, NSFG women having first child when aged 19 to 38. NSFG: National Survey of Family Growth.

Figure 2: Employment Rates for Treatment and Control Groups by Month



Note: TREATMENT= women who had a child by Month 21; CONTROL=women how did not have a child by Month 21.

Table 1: Algorithm for Constructing the HELP Sample

Step	Number of remaining observations
1. Start with the whole NSFG sample.	10,847
2. Drop women who did not seek help to get pregnant.	895
3. Drop women who sought help for the first time less than 21 months prior to the interview.	788
4. Drop women who were younger than age 19 or older than age 38 when seeking help for the first time.	745
5. Drop women who had already a child when seeking help for the first time.	553
6. Drop women who had adopted or step children when seeking help for the first time.	536
7. Drop women who were pregnant at some point of the month in which they sought help for the first time.[a]	500
8. Drop a woman with missing information in the insurance coverage variable.	499

[a]This group could include women who became pregnant right after seeking help for the first time (which occurred in the same month), or who were pregnant at the time when they sought help but did not know it. In fact, 23 of the 36 women reported as being pregnant in the same month they first sought help became pregnant in that month or in the previous one.

Table 2: Descriptive Statistics
HELP Sample

Variable	Mean (Standard deviation) All women	Treatment *AnyChildren21*=1[a]	Control *AnyChildren21*=0[a]
Employed21 (=1 if employed in Month 21)	0.798 (0.402)	0.624 ** (0.484)	0.877 (0.329)
Employed0 (=1 if employed in Month 0)	0.862 (0.345)	0.881 (0.324)	0.853 (0.354)
Employed_12 (=1 if employed in Month –12)	0.855 (0.352)	0.841 (0.366)	0.862 (0.345)
OwnChildren21 (number of own children in Month 21)[b]	0.323 (0.491)	1.036 ** (0.185)	0.000 (0.000)
AnyOtherChildren21 (=1 if had adopted or stepchildren in Month 21)	0.020 (0.141)	0.005 (0.073)	0.027 (0.162)
Age0 (age in Month 0)	26.3 (4.3)	25.9 (4.7)	26.5 (4.1)
Year0 (year in Month 0 normalized as 1970=0)	14.7 (5.7)	15.0 (6.1)	14.5 (5.5)
Education0 (years of education in Month 0)	13.6 (2.5)	13.8 (2.6)	13.5 (2.4)
Hispanic (=1 if Hispanic)	0.069 (0.254)	0.113 * (0.317)	0.050 (0.217)
Black (=1 if black)	0.087 (0.281)	0.078 (0.267)	0.091 (0.287)
Married0 (=1 if married in Month 0)	0.884 (0.320)	0.884 (0.320)	0.884 (0.321)
Smoke0 (=1 if smoked in Month 0)	0.370 (0.483)	0.286 * (0.452)	0.408 (0.492)
InsuranceCovered (=1 if insurance covered infertility treatments)	0.789 (0.408)	0.792 (0.406)	0.787 (0.409)
N	499	164	335

* Significantly different from the control group at the 5% level.
** Significantly different from the control group at the 1% level.
[a]*AnyChildren21*=1 if the woman had given birth to at least one child by Month 21.
[b]Six women had two children. Five had given birth to twins, and one had given birth twice.

Table 3: Comparison of HELP Sample with Childbearing Women in the NSFG

Variables [a]	Mean (Standard deviation)			
	NSFG — All Mothers [b]	NSFG — Mothers with first birth when aged 19 to 38 [c]	HELP sample	HELP sample reweighted [d]
Age	22.9 ** (4.9)	24.5 ** (4.2)	26.3 (4.3)	24.5 (4.2)
Year	14.0 * (7.0)	15.0 (6.4)	14.7 (5.7)	14.7 (6.1)
Employed_12 [e]	N/A	0.787 ** (0.409)	0.855 (0.352)	0.835 (0.371)
Education	12.3 ** (2.6)	12.8 ** (2.5)	13.6 (2.5)	13.1 (2.4)
Hispanic	0.125 ** (0.331)	0.112 ** (0.316)	0.069 (0.254)	0.098 (0.298)
Black	0.150 ** (0.357)	0.110 (0.312)	0.087 (0.281)	0.112 (0.315)
Married	0.702 ** (0.457)	0.782 ** (0.413)	0.884 (0.320)	0.857 (0.350)
Smoke	0.336 (0.472)	0.329 (0.470)	0.370 (0.483)	0.420 (0.494)
N	6,911	5,150	499	499

*Significantly different from the HELP sample at the 5% level.
**Significantly different from the HELP sample at the 1% level.
Note: NSFG: National Survey of Family Growth.
[a]Variables in Columns 2 and 3 are measured at the month in which the women gave birth to their first child (except for *Employed_12*). Variables for women in the HELP sample (Column 4) are measured in the month in which they first sought help to get pregnant (except for *Employed_12*).
[b]This sample is constructed by selecting from the NSFG sample all women who had at least one child.
[c]Includes all women in the NSFG sample who gave birth their first child when aged 19 to 38.
[d]Observations are reweighted to match the age and year distribution in the sample of NSFG women whose first birth occurred when aged 19 to 38.
[e]*Employed_12* equals 1 if the woman was employed 12 months before her first birth (third column) or 12 months before she first sought help to get pregnant (fourth column). In the case of the NSFG — All mothers sample (second column) this variable cannot be computed as work status is asked in the survey only for months after the woman reaches 18 years old.

Table 4: Linear Probability Estimates. Impact of a First Child on Employment HELP sample

	Dependent variable: *Employed21* Coefficient (Standard Error)					
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)
AnyChildren21	–0.253 (0.045)	–0.254 (0.044)	–0.261 (0.043)	–0.263 (0.043)	–0.283 (0.047)	–0.812 (0.265)
AnyChildren21 Age0*	—	—	—	—	—	0.011 (0.011)
AnyChildren21 Year0*	—	—	—	—	—	0.017 (0.008)
Age0	—	0.007 (0.005)	0.000 (0.005)	0.000 (0.005)	–0.005 (0.006)	–0.004 (0.006)
Year0	—	0.010 (0.004)	0.010 (0.004)	0.011 (0.004)	0.014 (0.004)	0.004 (0.004)
Smoke0	—	—	–0.045 (0.042)	–0.046 (0.041)	–0.025 (0.047)	–0.044 (0.040)
Education0	—	—	0.021 (0.007)	0.020 (0.007)	0.030 (0.008)	0.019 (0.007)
Hispanic	—	—	–0.131 (0.071)	–0.138 (0.069)	–0.087 (0.075)	–0.150 (0.070)
Black	—	—	0.014 (0.050)	–0.016 (0.051)	–0.149 (0.076)	–0.020 (0.051)
Married0	—	—	—	–0.089 (0.038)	–0.148 (0.046)	–0.104 (0.037)
InsuranceCovered	—	—	—	0.109 (0.049)	0.177 (0.055)	0.099 (0.048)
Constant	0.877 (0.019)	0.563 (0.126)	0.459 (0.156)	0.449 (0.149)	0.395 (0.162)	0.675 (0.169)
Adjusted R^2	0.0854	0.1190	0.1467	0.1666	0.2151	0.1904
N	499	499	499	499	499	499

Note: Observations in regressions 1 through 4 and 6 are weighted using weights from the NSFG. Regression 5 observations are reweighted to match the age and year distribution in the sample of NSFG women whose first birth occurred when aged 19 to 38. The mean of *Employed21* using NSFG weights is 0.798. For the reweighted sample, the mean of *Employed21* is 0.771. NSFG: National Survey of Family Growth.

Table 5: Multinomial Logit Estimates. Impact of a First Child on Work Status HELP sample

	Marginal effects of changing *AnyChildren21* from 0 to 1 (Standard error)	
	HELP sample	HELP sample reweighted [a]
No Job	0.291 (0.047)	0.314 (0.052)
Maternity leave	0.092 (0.027)	0.083 (0.025)
Part time	0.048 (0.027)	0.054 (0.036)
Full time	–0.431 (0.050)	–0.450 (0.055)
N	499	499
Log pseudo-likelihood value	–374.30	–381.25
Pseudo R^2	0.1738	0.2000

Note: The dependent variable has four categories: no job, maternity leave, part time, and full time. Covariates: *Age0, Year0, Smoke0, Education0, Hispanic, Black, Married0, InsuranceCovered*. NSFG: National Survey of Family Growth.

[a]Observations are reweighted to match the age and year distribution in the sample of NSFG women whose first birth occurred when aged 19 to 38.

Table 6: Linear Probability Estimates. Predicting Fertility using Selected Covariates HELP sample

	Dependent variable: *AnyChildren21* Coefficient (Standard error)
Independent variable	
Age0	–0.016 (0.006)
Year0	0.007 (0.005)
Smoke0	–0.102 (0.046)
Education0	0.013 (0.010)
Hispanic	0.198 (0.082)
Black	–0.052 (0.066)
Married0	–0.012 (0.051)
InsuranceCovered	0.022 (0.054)
Constant	0.472 (0.184)
Adjusted R^2	0.0427
P-value of *F*-test of joint significance	0.0020
N	499

Note: The mean of *AnyChildren21* is 0.312.

Table 7: Linear Probability Estimates. Explaining Employment in Month 0 using Fertility Status in Month 21 — HELP Sample

	Dependent variable: *Employed0* Coefficient (Standard Error)				
Independent variable	(1)	(2)	(3)	(4)	(5)
AnyChildren21	0.028 (0.035)	0.027 (0.035)	0.025 (0.034)	0.022 (0.034)	0.006 (0.041)
Age0	—	0.005 (0.004)	0.002 (0.005)	0.002 (0.004)	–0.003 (0.005)
Year0	—	0.007 (0.004)	0.007 (0.004)	0.008 (0.004)	0.011 (0.004)
Smoke0	—	—	–0.019 (0.041)	–0.020 (0.040)	0.011 (0.045)
Education0	—	—	0.010 (0.006)	0.009 (0.006)	0.021 (0.009)
Hispanic	—	—	–0.073 (0.067)	–0.075 (0.070)	–0.013 (0.067)
Black	—	—	0.003 (0.047)	–0.024 (0.048)	–0.107 (0.068)
Married0	—	—	—	–0.071 (0.037)	–0.093 (0.045)
InsuranceCovered	—	—	—	0.119 (0.052)	0.123 (0.056)
Constant	0.853 (0.022)	0.629 (0.112)	0.580 (0.142)	0.553 (0.146)	0.466 (0.164)
Adjusted R^2	0.0014	0.0242	0.0333	0.0592	0.0837
N	499	499	499	499	499

Note: Observations in regressions 1 through 4 are weighted using weights from the NSFG. Regression 5 observations are reweighted to match the age and year distribution in the sample of NSFG women whose first birth occurred when aged 19 to 38. The mean of *Employed0* using NSFG weights is 0.862. For the reweighted sample, the mean of *Employed0* is 0.847. NSFG: National Survey of Family Growth.

Table 8: Descriptive Statistics
HELP Panel Data and NSFG Panel Data

	Mean (Standard deviation)	
Data	NSFG — Cycle 5 (1995)	NSFG — Cycle 5 (1995)
Sample	HELP panel data	NSFG panel data
Unit of observation	Woman-month	Woman-month
Employed	0.841 * (0.365)	0.808 (0.394)
AnyChildren	0.087 ** (0.281)	0.171 (0.376)
Age	27.0 (4.4)	27.0 (4.4)
Education	14.0 (2.5)	14.0 (2.6)
Married	0.873 (0.333)	0.891 (0.311)
Smoke	0.361 (0.480)	0.327 (0.469)
Year (1970=0)	14.8 ** (5.5)	15.6 (5.8)
Hispanic	0.059 (0.236)	0.066 (0.248)
Black	0.076 (0.264)	0.056 (0.230)
Baby age in months (for women with babies)	5.5 ** (3.5)	6.1 (3.7)
Number of observations	19,743	237,751
Number of women	467 [a]	4,786

*Significantly different from the mean of the NSFG panel data at the 5% level.
**Significantly different from the mean of the NSFG panel data at the 1% level.
Note: NSFG: National Survey of Family Growth.
[a]Thirty two women included in the HELP sample answered the NSFG less than 33 months after seeking help becoming pregnant. These women are not included in the panel data set because it includes monthly observations for each woman in the 33 months after seeking help to get pregnant.

Table 9: Impact of a First Child on Employment
HELP Panel Data and NSFG Panel Data

	Dependent variable: *Employed* Coefficient (Standard error)			
Data	NSFG — Cycle 5 (1995)		NSFG — Cycle 5 (1995)	
Sample	HELP panel data		NSFG panel data	
Unit of observation	Woman-month		Woman-month	
Regression model	OLS	Fixed effects	OLS	Fixed effects
AnyChildren	–0.260 (0.036)	–0.234 (0.034)	–0.259 (0.010)	–0.216 (0.010)
Pregnant	–0.092 (0.020)	–0.065 (0.017)	–0.074 (0.008)	–0.050 (0.007)
Age	0.003 (0.004)	0.004 (0.007)	0.000 (0.002)	–0.003 (0.003)
Education	0.011 (0.005)	0.032 (0.023)	0.005 (0.001)	0.029 (0.011)
Married	–0.037 (0.033)	–0.055 (0.025)	0.033 (0.011)	–0.020 (0.010)
Smoke	–0.026 (0.037)	0.091 (0.060)	0.016 (0.003)	0.007 (0.022)
Year (1970=0)	0.008 (0.003)	—	–0.073 (0.019)	—
Hispanic	–0.103 (0.066)	—	0.024 (0.018)	—
Black	–0.035 (0.048)	—	–0.020 (0.010)	—
Constant	0.568 (0.143)	0.321 (0.287)	0.576 (0.045)	0.616 (0.158)
Adjusted R^2	0.0813	0.6666	0.0880	0.5761
N	19,743	19,743	237,751	237,751

Note: Fixed-effects model for the HELP panel data includes dummy variables for individuals and months relative to the first time they sought help to become pregnant. Fixed-effects model for the NSFG panel data includes dummy variables for individuals and calendar years. Observations are clustered by individual. NSFG: National Survey of Family Growth.

Table 10: Multinomial Logit Estimates. Impact of a First Child on Work Status HELP Panel Data and NSFG Panel Data

	Marginal effects of changing *AnyChildren* from 0 to 1 (Standard error)	
Data	NSFG — Cycle 5 (1995)	NSFG — Cycle 5 (1995)
Sample	HELP panel data	NSFG panel data
Unit of observation	Woman-month	Woman-month
No Job	0.253 (0.038)	0.246 (0.009)
Maternity leave	0.115 (0.015)	0.116 (0.004)
Part time	0.010 (0.021)	0.005 (0.006)
Full time	–0.378 (0.036)	–0.368 (0.009)
N	19,743	237,751
Log pseudo-likelihood value	–13887.78	–195,128.98
Pseudo R^2	0.0973	0.0904

Note: The dependent variable has four categories: no job, maternity leave, part time and full time. Covariates: *Age*, *Year*, *Smoke*, *Education*, *Hispanic*, *Black*, *Married*. Observations are clustered by individual. NSFG: National Survey of Family Growth.

Table 11: Descriptive Statistics
HELP and Census Samples

	Mean (Standard deviation)			
Sample	EARLY HELP	1980 Census	LATE HELP	1990 Census
Sample description	Married women in HELP sample who sought help before 1985	Married women aged 21 to 40 childless or with children younger than 1 year old	Married women in HELP sample who sought help on or after 1985	Married women aged 21 to 40 childless or with children younger than 1 year old
Time point	21 months after seeking help for the first time	1980	21 months after seeking help for the first time	1990
	(1)	(2)	(3)	(4)
Observation year	1981.5 ** (3.5)	1980.3 (0.0)	1991.0 ** (2.6)	1990.3 (0.0)
Employed	0.731 (0.443)	0.726 (0.446)	0.854 * (0.353)	0.796 (0.403)
AnyChildren	0.289 ** (0.453)	0.158 (0.364)	0.358 ** (0.479)	0.128 (0.334)
Age	26.1 ** (3.1)	27.2 (4.9)	30.2 ** (4.1)	29.3 (5.3)
Education	13.4 (2.4)	13.4 (2.6)	14.1 (2.6)	13.9 (2.5)
Hispanic	0.050 (0.218)	0.053 (0.223)	0.081 (0.272)	0.077 (0.266)
Black	0.065 (0.246)	0.061 (0.239)	0.069 (0.253)	0.061 (0.240)
N	216	287,292	224	301,371

*Significantly different from the mean of the Census comparable samples at the 5% level.
**Significantly different from the mean of the Census comparable samples at the 1% level.
Note: The EARLY HELP sample is compared to 1980 Census data and LATE HELP to 1990 Census data.

Table 12: Linear Probability Estimates. Impact of a First Child on Employment HELP and Census Samples

	Dependent variable: *Employed* Coefficient (Standard error)			
Sample	EARLY HELP	1980 Census	LATE HELP	1990 Census
Sample description	Married women in HELP sample who sought help before 1985	Married women aged 21 to 40 childless or with children younger than 1 year old	Married women in HELP sample who sought help on or after 1985	Married women aged 21 to 40 childless or with children younger than 1 year old
Time point	21 months after seeking help for the first time	1980	21 months after seeking help for the first time	1990
Mean of dependent variable	0.731	0.726	0.854	0.796
Independent variable	(1)	(2)	(3)	(4)
AnyChildren	–0.372 (0.072)	–0.365 (0.002)	–0.182 (0.055)	–0.228 (0.003)
Age	0.007 (0.012)	–0.004 (0.000)	–0.011 (0.007)	–0.001 (0.000)
Education	0.024 (0.013)	0.030 (0.000)	0.021 (0.008)	0.031 (0.000)
Hispanic	0.042 (0.096)	–0.047 (0.004)	–0.259 (0.099)	–0.087 (0.004)
Black	–0.024 (0.109)	–0.017 (0.003)	0.032 (0.062)	–0.033 (0.004)
Year0	0.014 (0.011)	—	–0.002 (0.008)	—
Constant	0.196 (0.270)	0.489 (0.006)	1.011 (0.272)	0.417 (0.007)
Adjusted R^2	0.2145	0.1222	0.1253	0.0814
N	216	287,292	224	301,371