Identifying the Consequences of Teenage Childbearing

By Adam B. Ashcraft*

17 July 2001

JEL Codes: J13, J16, J22
Keywords: instrumental variables, miscarriage, teenage childbearing

Using miscarriage as an instrument for birth suffers from a censoring problem created by the fact that women wanting the pregnancy to end in an abortion frequently pre-empt an "assigned" miscarriage. I describe evidence that this problem is severe, creates a significant bias in IV estimates, and develop non-parametric estimators to correct for censoring. While simple IV estimates previously suggested teenage childbearing had little effect on high school completion or labor market outcomes later in life, non-parametric IV estimates indicate significant costs of teenage childbearing, particularly for high school aged girls.

* Banking Studies Function, Federal Reserve Bank of New York, 33 Liberty Street, New York, NY 10045. Send e-mail to Adam.Ashcraft@ny.frb.org or phone (212) 720-1617 The views expressed I this paper are those of the author and are not necessarily the views of the Federal Reserve Bank of New York or the Federal Reserve System. I would like to thank Josh Angrist, Jonathan Guryan, John Johnson IV, and participants of the MIT Labor Lunch for their constructive comments. Any remaining errors are my own.

Identifying the Consequences of Teenage Childbearing

By Adam B. Ashcraft
Banking Studies Department
Federal Reserve Bank of New York
33 Liberty Street
New York, NY 10045
(212) 720-1617
Adam.Ashcraft@ny.frb.org

JEL Codes: J13, J16, J22
Keywords: instrumental variables, miscarriage, teenage childbearing

Abstract

This paper continues research on female labor supply, in particular the literature on the costs of childbearing to a woman's later labor market and marital outcomes. Hotz, Mullin, and Sanders (1997) use miscarriage as an instrument for birth, concluding that delaying birth might actually worsen outcomes for pregnant teenage girls. I demonstrate that while miscarriages might be randomly assigned given a woman's smoking, drinking, and contraceptive behavior, observed miscarriages are still not a valid instrument. It turns out that more than one out of every two women wanting the pregnancy to end in an abortion will never have an assigned miscarriage, and this likely creates a bias that underestimates the true costs of teenage childbearing. I consider various estimators that correct the IV estimates for this censoring problem and employ a data set better suited to measure pregnancy outcomes – the 1995 *National Survey of Family Growth*. After implementing these new estimators, I find in contrast to the existing literature that that teenage childbearing significantly reduces the probability of high school completion, but only for girls in high school. Moreover, there is some evidence of a causal link between teenage childbearing and labor market outcomes and weak evidence of a link to marital outcomes later in life.

1. Introduction

Economists and policymakers have long been interested in teenage childbearing. Bill Clinton called it "our most serious social problem" in his 1995 State of the Union Address. Recently the Center for Disease Control (CDC) enthusiastically reported that the teen pregnancy rate had fallen 19 percent from its all time high in 1991 to reach a record low in 1997. Jeffrey Koplan of the CDC noted that "few teens are ready for the challenges of parenthood," and "when they delay this responsibility it enables them to gain the education and maturity they need to become good parents and good citizens."

Despite this interest in policymakers in teenage pregnancy, economists have struggled to establish convincing evidence of a causal link between teenage childbearing and a young girl's educational, labor market, and marital outcomes. While the average outcomes of pregnant teenagers who choose to delay childbearing through abortion are generally much better than those that bring the child to term, there are probably differences in what these latter girls would have achieved if they also chose to delay motherhood. When only exploiting variation in age at first birth created by variation in age at menarche, Ribar (1994) finds there is no causal link between teenage childbearing and high school completion in the NLSY. Hotz, McElroy, and Sanders (1997) use miscarriage as an instrument for birth in the NLSY, finding that delaying childbearing may actually reduces the probability that girls complete their high school degrees and deteriorates a number of other outcomes later in life. On the other hand, as the other two instruments affect approximately the same group of young girls – and thus are attempting to estimate the same parameter of interest -- one might be concerned about large differences in reported results between them.[i]

The absence of convincing evidence establishing a causal link between teenage childbearing and stands in contrast to the existing literature on female labor supply. Angrist and Evans (1998) use the randomness of sex mix of the first two children as an instrument for having a third child, finding a small effect on female labor supply with little effect on husband outcomes. Bronars and Grogger (1994) use the randomness of twins at first birth as a natural experiment and conclude there is evidence of what seem to be

small and short-lived effects on unwed mothers. Both of these studies seem to provide evidence that having a child creates significant demands on a woman's time, and in absence of considerable family support one might think it is reasonable for these demands to affect younger girls as well.[ii]

This paper demonstrates that there is a serious censoring problem in implementing miscarriage as an instrument for birth and develops non-parametric estimators that correct for this censoring in identifying the costs of teenage childbearing. Section 2 outlines the basic framework for analysis, while Section 3 presents the essence of the censoring problem and my solution. Section 4 describes the data employed, while Section 5 describes preliminary evidence that the censoring problem could be important. Section 6 establishes the validity of miscarriage as an instrument in absence of the censoring problem while Section 7 discusses identification in its presence. Empirical results are described in Section 8, and I conclude with Section 9.

2. Basic framework for analysis

The population of interest includes those women having their first pregnancy as a teenager. Denote by Y_1 the outcome of interest for a particular woman whose pregnancy outcome is birth, and alternatively denote by Y_0 the outcome for the same women if she chose to delay childbearing. To fix ideas, Y could represent completion of a high school degree. The observed outcome Y can be written as a simple linear function of B, an indicator variable for the first pregnancy ending in birth, and these potential outcomes: $Y = Y_1B+Y_0(1-B)$.

All of these pregnancies must end in only one of the following outcomes: a live birth, abortion, or miscarriage. Implicit in this equation is the restriction that the potential outcome of woman having an abortion is identical regardless of the method used to delay birth.[iii] I assume that each woman has an underlying preference for the pregnancy outcome in the absence of a miscarriage. In particular, assume that each woman either wants the pregnancy to end in a live birth or an abortion. A woman's preference

for the outcome is captured by the indicator function B^* which is one if the woman wants the pregnancy to end in live birth, and zero if she wants to have an abortion: $B^* = 1(\gamma > \varepsilon_B)$.

Preferences are assigned by the random variable ε_B which probably has components correlated with potential outcomes (Y_1, Y_0). In particular, women preferring birth might be less likely to complete high school than women preferring abortion, even if they were to delay birth. I will often refer to women wanting births as latent-birth type women and women wanting abortions as latent-abortion type women. Whether or not women have preferences at the start of the pregnancy or these preferences evolve over the course of the pregnancy is not important. The crucial element here is that women wanting births probably have different average potential outcomes than women wanting abortions.

The causal effect of interest is the average consequence of teenage childbearing on later outcomes for a woman having preference for a child at the time of pregnancy, represented by $\alpha_0 = E[Y_1 - Y_0 | B^* = 1]$. In the terminology of Angrist, Imbens, and Rubin (1996), this parameter is the Local Average Treatment Effect (LATE), representing the average causal effect of delaying childbearing for the sub-population of pregnant teenagers who would otherwise bring the child to term. While the first term of the causal effect is readily obtained from the labor market outcomes of formerly pregnant teenagers who had births, the latter part is never observed and is central to the issue of identification. It should be clear that the only mechanism through which these pregnant latent-birth women would actually able to deliberately delay their childbearing is through abortion. By assumption, potential outcomes in the event of abortion are no different than in the event of miscarriage. In this sense the parameter represents the average effect of an abortion on potential outcomes for women who choose not to have one.

There is a close relationship between the parameter being estimated and the actual variation in pregnancy outcome exploited. As demonstrated below, the parameter α_0 is the one defined when using miscarriage as an instrument for birth in a regression of Y on B. This follows because miscarriage only affects whether or not latent-birth women actually have births, and latent-abortion women never have births. This parameter should be contrasted to the one implicitly defined by Ribar (1995), who attempts to

identify the casual effect of delaying birth by delaying fertility through the age at menarche. This instrument clearly operates on the sub-population of sexually active young women who – if became pregnant -- would bring a child to term. Even though both of these instruments seem to affect similar girls, it should be noted that most of the variation in age at menarche is between 11 and 14 years of age. The parameter estimated in Ribar (1995) thus refers to a much younger population than that used in Hotz, McElroy, and Sanders (1997) – referred to hereafter as HMS -- where the median age is 16. Another related parameter in this literature is estimated by Levitt and Donohue (2001), who exploit state-time variation in the legalization of abortion. Using this parameter, the authors estimate the causal effect of abortion access on the children of women who wanted to have abortions but were previously unable to because of the existing legal environment. In this framework these would be latent-abortion women, and thus Levitt and Donohue (2001) refer to a much different sub-population of women and one might consequently expect quite different results. Under a reasonable theory of choice, women who delay birth might have the most to gain by doing so, so the costs of childbearing might be expected to be much smaller for latent-birth women.[iv] It thus makes sense that Ribar (1994) and HMS find relatively smaller effects. Whether these are actually zero, however, is the task of this paper.

3. Identification -- the big picture

Consider employing the event of miscarriage as an instrument for birth a vector of covariates X that describe a woman's behavior before and during the pregnancy that make assigned miscarriage independent of potential outcomes after conditioning on them. There is evidence in the epidemiological literature that suggests some miscarriages are caused by the seemingly random events, while others are induced by a woman's behavior, which is probably not independent of her potential outcomes.[v] In this sense miscarriage is arguably a random even given a woman's behavior before the pregnancy. Figure 1 describes the ideal experiment below.[vi] The treatment group is randomly assigned miscarriages while controls are randomly permitted to realize their latent preferences. This experiment is complicated by the fact that an assigned miscarriage can take up to six months to actually happen, and consequently women wanting the pregnancy to end in abortion may actually end up having abortions before the miscarriage occurs. As the

sub-population of interest only includes latent-birth type women and for now we are implicitly assuming the offer of treatment is observed, this does not present a problem for identification. It turns out that the differences in average labor market or pregnancy outcomes by offer of treatment will recover the causal effect of interest when normalized by the fraction of women in the population wanting the pregnancy to end in birth.

Note that the average outcome of women offered miscarriages can be broken into two parts by latent-type as follows: $E[Y|Offer] = E[Y_0|Offer,B^*=1]Pr(B^*=1|Offer)+E[Y_0|Offer,B^*=0]Pr(B^*=0|Offer)$. Similarly, the average outcome of women not offered miscarriages can be broken into two parts by latent-type.

$E[Y|no\ Offer] = E[Y_1|no\ Offer,B^*=1]Pr(B^*=1|no\ Offer)+E[Y_0|no\ Offer,B^*=0]Pr(B^*=0|no\ Offer)$. We are suppressing conditioning on X for convenience. The randomness of assigned miscarriages implies we can write the difference $E[Y|no\ Offer]-E[Y|Offer]$ as simply $E[Y_1-Y_0|B^*=1]Pr(B^*=1)$, a proof of our claim that the causal effect of interest is identified.

The problem with the above analysis is that the offer of treatment is actually not observed, and that the presence of non-compliance by latent-abortion types renders miscarriage an invalid instrument. This latter claim is true as while the offer of treatment may be randomly assigned given covariates X, the event of miscarriage is no longer plausibly independent of potential outcomes as compliance with treatment assignment is correlated with a woman's latent type. In particular, all women wanting births will comply with assigned miscarriages, but only some of women wanting abortions will do so. This is illustrated in Figure 2 below.

Simple differences in the average labor market or marital outcomes of women by the event of a miscarriage fail to identify the causal effect of interest. Moreover, it is reasonable to argue that using observed miscarriage as an instrument for birth actually underestimates the costs of teenage childbearing. In particular, if latent-abortion women generally have the better potential outcomes than latent-birth women, Figure 2 illustrates that we will be overestimating the average outcomes of women not assigned

miscarriages when using the average outcomes of women actually not having miscarriages as there are latent-abortion women offered treatment that end up not having miscarriages so E[Y|no Miscarriage] > E[Y|no Offer]. At the same time, using the average outcomes of women observed to have miscarriages will underestimate the average outcomes of women assigned miscarriages as E[Y|Offer] > E[Y|Miscarriage]. These comparisons are important as from the above it is differences in offer of treatment identify the parameter of interest. When these two claims are combined they imply that the normalized difference in average outcomes by event of miscarriage above will actually overestimate the causal effect of interest as E[Y|no Miscarriage] – E[Y|Miscarriage] > E[Y|no Offer]-E[Y|Offer]. Overestimating the benefits of teenage childbearing is conceptually equivalent to underestimating the costs. This analysis implies that employing miscarriage as an instrument for birth may lead researchers to mistakenly claim that the differences in average outcomes of pregnant teenagers having births with those who do not are attributable to differences in potential outcomes and do not have a causal interpretation.

How serious is this problem? Estimates below suggest that fewer than one out of every two miscarriages assigned to latent-abortion women will ever be realized. Given significant differences between women who have miscarriages and those who have abortions, accounting for the presence of non-compliance by latent-abortion types could be substantial. For example, Table V reveals that women having abortions are more 30 percentage points more likely to earn a high school diploma than those having miscarriages

This compliance problem is actually more severe than alluded to above as there are important differences in potential outcomes between women having early abortions and those having later abortions. This is problematic as women wanting an early abortion are more likely to avoid a miscarriage than women wanting a late abortion, and women from the former group generally have better potential outcomes than women from the latter, so the bias discussed above is magnified. For example, Table V indicates that women having abortions in the first trimester 10 percentage points more likely to graduate from high school than women having late abortions. Correcting for this second source of bias is more complicated

than the first as it requires recovery of the time distribution of planned abortions and assigned miscarriages from the data given X, which is conceptually straightforward but may be too much to ask of the data.

The solution to this censoring problem is to essentially replicate the original random assignment of treatment given X, and this is the approach implemented non-parametrically below. We note that the average outcomes of women assigned miscarriages E[Y|Offer] can be broken into two parts: the average outcomes of women having miscarriages E[Y|Miscarriage]; and the average outcome of women assigned miscarriages but having abortions E[Y|Offer,No Miscarriage]. The average outcome of women having miscarriages is observed directly from the data, so our contribution is to make plausible assumptions to identify the average outcomes of latent-abortion women assigned miscarriages but having abortions. There is a similar decomposition of women not assigned miscarriages into two parts. Our estimator below assumes that compliance by latent-abortion women is ignorable so we can proxy for the average outcome of these women E[Y|Offer,No Miscarriage] simply with the average outcome of women having abortions E[Y|Abortion]. It is argued that this estimator will be still understate the true costs of teenage childbearing as the average outcomes of non-compliers (women assigned miscarriage but having an abortion) are probably better than the average outcomes of women having abortions E[Y|Offer,No Miscarriage] > E[Y|Abortion]. This is because women preferring early abortions have larger weight in the distribution of non-compliers and the distribution of all abortions Pr(Early|Offer,No Miscarriage) > Pr(Early|Abortion) and the early types have better potential outcomes $E[Y|B^*=0,Early] > E[Y|B^*=0,Late]$. This correction seems to suggest that the bias in IV when using miscarriage as an instrument is serious, and in contrast to HMS, delaying childbearing through abortion seems to have economic and educational benefits for teenage girls, although these benefits seem to vary across background characteristics.

4. Data

The data employed in this analysis is the 1995 wave of the *National Survey of Family Growth*, a survey administered by the National Center for Health Statistics (NCHS), an agency of the Department of Health and Human Services. The purpose of the survey is to produce national estimates and an information base

on factors affecting pregnancy—including sexual activity, contraceptive use, infertility, and sources of family planning services—and the health of women and infants. A national probability sample of 10,847 civilian non-institutionalized women ages 15 to 44 were interviewed between mid-January and October 1995. The interviews were conducted in person by trained female interviewers using laptop or notebook computers. The interview, which lasted an average of 105 minutes, collected data on each pregnancy; contraceptive use by her and her partner; her ability to bear children; the use of medical services for contraception; infertility and prenatal care; her marriage, cohabitation, living situation, and work history; and a variety of demographic and economic characteristics. Additional data were collected in a short self-administered interview in which the respondent heard questions over headphones and entered her own answers into the notebook computer.

Descriptive statistics are presented in Table I for two cross-sections of the data, which reports sample means and standard deviations using sample weights. The *Teen Pregnancy* sub-sample corresponds to girls who had their first pregnancy before their 18th birthday and are currently at least 20 years old. This is conceptually the same sample HMS extract from the *National Longitudinal Survey of Youth* (NLSY) in their analysis, and will be the basis for much of our empirical work below.[vii] Note that there are significant differences in the background characteristics and outcomes of the *Teen Pregnancy* sample from the *Full Sample*. These latter girls are much more likely to have a parent absent in the household, are exposed to much less parental education, and more likely to have a mother working full time. Moreover, these girls have their first almost four years earlier, are more likely to have had a failed marriage, less likely to complete high school, and work and earn less. These distinctions are made in order to emphasize the conditioning that is implicitly made below in focusing the analysis on the *Teen Pregnancy* sub-sample.

In order to describe some evidence that there is a censoring problem when using miscarriage as an instrument, I exploit all of the pregnancies tracked in the survey to construct a panel data set for the 7,660 women in the survey who have ever been pregnant. Each observation corresponds to a single pregnancy, and sample statistics for this panel are reported in Table II.[viii]

Each of these variables employed in the analysis below is constructed in a straightforward manner except for the smoking dummy variable. In the survey, women are asked if they have smoked more than 100 cigarettes in their life, and if so when they started smoking. A smoker is defined as a woman who started smoking before the age at which the pregnancy started. There is more detailed data on smoking during the pregnancy, but these questions are not asked of women who end the pregnancy with an abortion. For this reason I choose to use the aforementioned smoking variable, which is similar to that employed by HMS. The significant difference with HMS in using this data set is the inclusion of data on the use of an IUD before conception and omission of data on drinking behavior before the pregnancy.

The main advantages of using the NSFG over the NLSY are in the larger sample size and higher quality of data. As the NLSY follows a panel of young men and women defined in 1979, there is a significant attrition over time that is plausibly non-random. In contrast, the NSFG is a cross-sectional survey designed to measure national fertility statistics, and should be much more representative of the population of interest. The underreporting of abortions in survey data is also a serious problem that must be acknowledged in any study using fertility microdata, and while between 54 and 62 percent of abortions that actually occurred were probably reported in the NSFG, it is unlikely that the NLSY does anywhere near that well.[ix] As underreporting remains significant in the data employed, I consider how random and non-random non-reporting of abortions affects the non-parametric estimators implemented below.

5. Preliminary evidence

Define M and A as indicators for the pregnancy outcome of miscarriage and abortion. Further define M^* and A^* as indicator functions for assigned miscarriage and latent-abortion women, respectively (where $A^* = 1-B^*$). Assume that the length of pregnancy is a discrete variable measured in weeks. This variable is used as a subscript below on pregnancy outcomes, latent-pregnancy types, and assigned miscarriages to indicate the timing of an actual or assigned event. Given the distributions of actual abortions and miscarriages over the weeks of pregnancy, it is possible to recover the distributions of assigned miscarriages and latent-abortions under some simplifying assumptions.

When both an abortion and miscarriage are assigned in the same week, I first assume that the miscarriage occurs. Under this assumption, it is possible to immediately recover the first point of each probability distribution function:

$$\Pr(M_1=1) = \Pr(M_1^*=1)$$

$$\Pr(A_1=1) = \Pr(A_1^*=1, M_1^*=0)$$
$$= \Pr(A_1^*=1)\Pr(M_1^*=0)$$

The second equality from the second line follows from the randomness of assigned miscarriages. While this is strictly true only given X, I ignore this complication now for descriptive purposes.[x] The entire distributions for M^* and A^* are defined recursively as follows:

$$\Pr(M_k = 1) = \Pr(M_k^*=1,A_1^*=0,\ldots,A_{k-1}^*=0)$$
$$= \Pr(M_k^*=1)[1-\Sigma_{j=1}^{k-1}\Pr(A_j^*=1)]$$

$$\Pr(A_k = 1) = \Pr(A_k^* = 1,M_1^*=0,\ldots,M_k^*=0)$$
$$= \Pr(A_k^* = 1)[1-\Sigma_{j=1}^{k}\Pr(M_j^*=0)]$$

Note that by integrating over the entire support of pregnancy lengths, this assumption also identifies the fraction of latent-birth women in the sample $E[B^*]$ and probability of assigned miscarriage $E[M^*]$, two key inputs into the analysis below.[xi]

These latent distributions are estimated as described above and then smoothed with a nonlinear smoother, and are illustrated in panel (a) of Figure 3.[xii] The figure illustrates that the distribution for latent-abortion generally puts much more weight on early weeks than assigned miscarriage. The data imply a latent-abortion fraction of 14.2 percent and probability of assigned miscarriage of 14.1 percent while (when

unweighted) abortions represent 13.4 percent and miscarriages 13.0 percent of all pregnancies in the sample. Writing $M = M^* C$, where C is an indicator function for compliance with assigned miscarriages, it is possible to measure the average compliance by latent-abortion types under the assumption that compliance is ignorable.

$$
\begin{aligned}
E[M] &= E[M^*C] \\
&= E[M^*C|B^*=1]Pr(B^*=1)+E[M^*C|B^*=0]Pr(B^*=0) \\
&= E[M^*]E[B^*]+E[M^*C|B^*=0](1-E[B^*) \\
&= E[M^*]E[B^*]+E[M^*]E[C|B^*=0](1-E[B^*)
\end{aligned}
$$

Rearranging these terms,

$$E[C|B^*=0] = \{E[M]/E[M^*]-E[B^*]\}/(1-E[B^*])$$

Using the data from above, these figures suggest that fewer than one out of every two miscarriages actually assigned to latent-abortion women actually occur. As is such, the bias from using actual instead of assigned miscarriages is plausibly large. For robustness, I estimate the distributions using the opposite assumption – that whenever an abortion and miscarriage are assigned in the same week an abortion occurs – but there is little change in results as is illustrated in panel (b) of Figure 3.

There is another way of making this point. If miscarriage is truly randomly assigned given the vector of covariates X, then there should exist no set of covariates Z that predict miscarriage after conditioning on X. When latent-abortion women fail to comply with assigned miscarriage, however, women that one could predict on the basis of covariates Z to terminate the pregnancy would actually be less likely to have a miscarriage. To implement this idea I use the religion in which a woman was raised and her marital status at conception in a probit model of abortion, and then predict the probability of abortion for every pregnancy in the sample. In a second-stage regression I implement a probit model of miscarriage using the risk factors known to cause miscarriage and my variable measuring predicted abortion. In a

second specification I simply regress miscarriage on abortion and risk factors, with abortion instrumented by religion and marital status at conception.

In each of these models the identifying assumption is simply that there is no omitted variable that causes miscarriage but is correlated with parents' religion or marital status at conception. The most plausible omitted variable from the model is drinking behavior, which increases the risk of a miscarriage. The main question is whether or not drinking behavior is positively or negatively correlated with latent-abortion status conditional on all the other control variables in the regression. Moreover, if drinking behavior is negatively conditionally correlated with latent-abortion status, is the correlation strong enough explain why women that are likely to have abortions are also less likely to have miscarriages. Without the data one can only speculate, but it seems plausible that latent-abortion women would be less likely to change any type of behavior that might increase the risk of miscarriage. It should also be noted that in most US states the legal drinking age is several years older than the legal smoking age and all of these girls are no older than 17. While it would be naïve to think this implies these girls didn't drink until they were of legal age, it seems unlikely they drank frequently enough to significantly affect the risk of miscarriage.[xiii]

Results are described in Table III, and generally support the hypothesis that women predicted to have abortions are significantly less likely to have miscarriages. The first and third columns indicate that being married at conception sharply reduces the probability of the pregnancy ending in abortion while the second and fourth columns indicate that women that one would predict to end the pregnancy in abortion are actually less likely to be assigned a miscarriage. In the latter case the coefficient is more than half of the probability of miscarriage, again suggesting that this censoring problem could be very important in creating a bias in the IV estimator.

6. Identification without non-compliance

I assume that all pregnancies ending in miscarriage are one of two types, either random or nonrandom. There is strong evidence that some miscarriages are caused by the abnormal development of fetal

chromosomes, and that the occurrence of this abnormal development is random. On the other hand, epidemiological studies have found that smoking and drinking during pregnancy, in addition to conception before age fifteen and use of an IUD contraceptive device, can also increase the probability of a miscarriage. As these behaviors are likely correlated with potential outcomes, these induced miscarriages are certainly not randomly assigned. An indicator for a randomly assigned miscarriage is RM, while an indicator for a miscarriage induced by a woman's behavior during the pregnancy is denoted by NR: RM = $1(\gamma_{RM}>\varepsilon_{RM})$ and NR = $1(\gamma_{NR}>\varepsilon_{NR})$.

In the analysis that follows, it is maintained that while the assignment of random miscarriages is done independently of potential outcomes $(Y_0,Y_1,\varepsilon_B) \perp\!\!\!\perp \varepsilon_{RM}$, it is probably unreasonable to assume that the assignment of nonrandom miscarriage is done independently of these outcomes $(Y_0,Y_1,\varepsilon_B) \perp\!\!\!\perp \varepsilon_{NR}$. However, it will assumed that $\varepsilon_{NR} = \delta X+\varepsilon_{nr}$, where it is maintained that $(Y_0,Y_1,\varepsilon_B) \perp\!\!\!\perp \varepsilon_{nr}$, implying $(Y_0,Y_1,\varepsilon_B) \perp\!\!\!\perp \varepsilon_{NR}|X$. This assumption is consistent with the approach of HMS and the epidemiological literature on the causes of miscarriages, which suggest that X should include smoking and drinking during the pregnancy, pregnancies before the age of 15, and the use of an IUD contraceptive device during conception.

Define an indicator for an assigned miscarriage by M^* as the union of the events NR=1 and RM=1: $M^* = NR+(1-NR)*RM$. Note both types of miscarriage could be present, but only one is required for assignment. Our assumptions above imply that while assigned miscarriages are certainly not unconditionally independent of potential outcomes, it may be reasonable to argue they are independent after conditioning on a proper set of covariates. Given the above indicator for assigned miscarriage, it is now possible to write the indicator for birth as $B = B^*(1-M^*)$.

This definition is consistent with women not wanting the pregnancy, $B^*=0$, never having births while some women wanting children, $B^*=1$, having their births censored by assigned miscarriages. The conditional independence of assigned miscarriages motivates our first result:

Theorem 1. *Using assigned miscarriages M^* as an instrument for an indicator of birth B, in a two-stage least squares regression of Y on B and covariates X, identifies LATE $E[Y_1-Y_0|B^*=1]$ given the following restrictions:*

(a) $(Y_0,Y_1,\varepsilon_B) \perp\!\!\!\perp \varepsilon_{RM}$

(b) $(Y_0,Y_1,\varepsilon_B) \perp\!\!\!\perp \varepsilon_{NR}|X$

(c) $E[M^*|X]$ *is linear in X.*

Proof.

The instrument M^* is as good as randomly assigned given X by assumption. This implies the following.

$$E[Y|M^*=1,X] = E[Y|B^*=1,M^*=1,X]Pr(B^*=1|M^*=1,X)+ E[Y|B^*=0,M^*=1,X]Pr(B^*=0|M^*=1,X)$$

$$= E[Y_0|B^*=1,X]Pr(B^*=1|X)+ E[Y_0|B^*=0,X]Pr(B^*=0|X)$$

$$E[Y|M^*=0,X] = E[Y|B^*=1,M^*=0,X]Pr(B^*=1|M^*=0,X)+ E[Y|B^*=0,M^*=0,X]Pr(B^*=0|M^*=0,X)$$

$$= E[Y_1|B^*=1,X]Pr(B^*=1|X)+ E[Y_0|B^*=0,X]Pr(B^*=0|X)$$

$$E[B|M^*=1,X] = E[B|B^*=1,M^*=1,X]Pr(B^*=1|M^*=1,X)+ E[B|B^*=0,M^*=1,X]Pr(B^*=0|M^*=1,X)$$

$$= 0$$

$$E[B|M^*=0,X] = E[B|B^*=1,M^*=0,X]Pr(B^*=1|M^*=0,X)+ E[B|B^*=0,M^*=0,X]Pr(B^*=0|M^*=0,X)$$

$$= Pr(B^*=1|X)$$

$$\{E[Y|M^*=0,X]-E[Y|M^*=1,X]\}/\{E[B|M^*=0,X]-E[B|M^*=1,X]\} = E[Y_1-Y_0|B^*=1,X]$$

$$E[Y_1-Y_0|B^*=1] = \Sigma_{X=x} E[Y_1-Y_0|B^*=1,X]dF(X|B^*=1)$$

Assumption (c) is not restrictive when X is a saturated set of dummy variables, and is required so that two-stage least squares recovers LATE, a result found in Abadie (1999).

Q.E.D.

While the above result is certainly true, it is not necessarily useful in the sense that assigned miscarriages are not observed. It is important to understand that the assignment of a miscarriage, $M^*=1$, does not necessarily guarantee the event of miscarriage, M=1, for all women. Some women who want the

pregnancy to end in abortion may actually terminate the pregnancy before an assigned miscarriage has time to occur. In addition, the women most likely to terminate the pregnancy early, and thus most likely to avoid an assigned miscarriage, may have potential outcomes different form those women who are likely to terminate the pregnancy late. This is an important point as it implies that observed miscarriages will not be a valid instrument for birth, even after conditioning on a vector of covariates X, as compliance with assigned miscarriage is correlated with a woman's preference for the birth outcome.

That being said, the result is important for several reasons. First, the result highlights that the true challenge for identification is the absence of the original random assignment and not presence of non-compliance. Second, this result allows us to motivate discussion of bias when using observed miscarriage as an instrument, illustrated in our first corollary:

Corollary 1.1. *Using observed miscarriages M as an instrument for an indicator of birth B, in a two-stage least squares regression of Y on B and covariates X, overestimates LATE* $E[Y_1-Y_0|B^*=1]$ *when latent-abortion women have better potential outcomes than latent-birth women in the sense that* $E[Y_0|B^*=0,X]>E[Y_0|B^*=1,X]$.

Proof.

Assume latent-abortion women have better potential outcomes than latent-birth women so $E[Y_0|B^*=0,X]>E[Y_0|B^*=1,X]$. As the group of women observed to have miscarriages does not include all latent-abortion women assigned miscarriages, it follows that $E[Y_0|M=1,X]<E[Y_0|M^*=1,X]$. Similarly, the group of women observed to not have miscarriages includes some latent-abortion women assigned miscarriages, and thus $E[Y|M=0,X] >E[Y|M^*=0,X]$. These two results imply that the difference in average outcomes by pregnancy status $E[Y|M=0,X]-E[Y_0|M=1,X]>E[Y|M^*=0,X]-E[Y_0|M^*=1,X]$, and IV using miscarriage overestimates the benefits of having a teenage birth.

Q.E.D.

Finally, the framework permits discussion of a feasible strategy for identification. In particular, identification is achievable if it were possible to recover the following parts of LATE described above:

$E[Y|M^*=0,X]$, $E[Y|M^*=1,X]$, $Pr(B^*=1|X)$, and $dF(X|B^*=1)$. Doing exactly this under alternative assumptions is the purpose of the rest of this paper.

7. Identification with ignorable non-compliance

In order to accommodate the problem of compliance in feasibly estimating LATE, it is necessary to construct an indicator function for compliance, denoted by C and defined as $C = B^*+(1-B^*)1(c<\varepsilon_c)$. This compliance function is consistent with all latent-birth women complying with assigned miscarriages while a fraction of latent-abortion women failing to comply with treatment assignment. A crucial issue for identification is whether or not this compliance is ignorable in the sense that $(Y_0,Y_1,\varepsilon_B) = \varepsilon_c|X$. If non-compliance is random, then the average outcomes of women not-complying with treatment assignment can be recovered from the average outcomes of women actually having abortions. We claim below that this assumption is probably too strong, and also will lead to an overestimate of LATE. Regardless of the ignorability of compliance, it is possible to write observed miscarriages as a function of assigned miscarriages and the compliance indicator, $M = CM^*$.

The failure of latent-abortion types to fully comply with assigned miscarriages poses a serious problem for identification. While M^* would be a valid instrument for identification, assuming that conditioning on a set of covariates X is sufficient for the exclusion restriction, it not a feasible estimator as the original assignment is not observed. Moreover, since compliance is correlated with latent-type, as all latent-birth types comply with treatment, it is not reasonable to argue that observed miscarriages are a valid instrument for birth $(Y_0,Y_1,\varepsilon_B) \perp\!\!\!\perp M|X$.

Our next result states that if compliance by latent-abortion women is ignorable then LATE is identified non-parametrically. While this identifying assumption is certainly untenable given differences in potential outcomes between women wanting early and late abortions, it is a good starting point for three reasons. First, it is possible to break down the bias in the two-stage least-squares estimator into two parts: one part for recognizing latent-abortion women are misclassified when using the event of miscarriage and

another for recognizing that there are differences in potential outcomes and compliance between types of latent-abortion women. Second, this result better motivates the approach for our final result, which is our non-parametric IV estimator of ϑ_0. Finally, there are constraints in the data which may make the implementation of this estimator, which can be thought of as a lower bound on LATE, more attractive:

Theorem 2. *LATE is identified by the restrictions :*

(a) $(Y_0, Y_1, \varepsilon_B) \perp\!\!\!\perp \varepsilon_e | X$

(b) $(Y_0, Y_1, \Pi_B) \perp\!\!\!\perp \varepsilon_c | X$

(c) $(Y_0, Y_1, \varepsilon_B) \perp\!\!\!\perp \varepsilon_{RM}$

(d) $(Y_0, Y_1, \varepsilon_B) \perp\!\!\!\perp \varepsilon_{NR} | X$

Proof.

Show that we can write $E[Y|M^*=1,X]$ and $E[Y|M^*=0,X]$ as functions of observable components, and then use Result 1 to prove LATE is identified. The crucial element here is that ignorable compliance by latent-abortion types implies that the average outcome of women not complying with assigned miscarriages is no different from the average outcome of women actually having abortions $E[Y_0|M^*=1,M=0,X] = E[Y_0|A=1,X]$.

$$E[Y|M^*=1,X] = E[Y_0|M^*=1,M=1,X]Pr(M=1|M^*=1,X)+E[Y_0|M^*=1,M=0,X]Pr(M=0|M^*=1,X)$$
$$= E[Y_0|M=1,X]\gamma+E[Y_0|A=1,X](1-\gamma)$$

Note the following definition for γ as the ratio of observed miscarriages to assigned miscarriages, which is a measure of compliance with assigned miscarriages given X.

$$\gamma = Pr(M=1|X)/Pr(M^*=1|X)$$
$$E[Y|M=0,X] = E[Y|M^*=0,M=0,X]Pr(M^*=0|M=0,X)+E[Y|M^*=1,M=0,X]Pr(M^*=1|M=0,X)$$
$$E[Y|M^*=0,X] = E[Y|M=0,X]/\beta-E[Y_0|A=1,X](1-\beta)/\beta$$

Note the following definition for $(1-\beta)$ as the fraction of assigned miscarriages found among observed births and abortions.

$$(1-\beta) = \Pr(M^*=1|M=0,X)$$
$$= \Pr(M^*=1|X)/\Pr(M=0|X)$$

Given these two parts and the estimated latent probabilities $\Pr(B^*=1|X)$ and $\Pr(M^*=1|X)$, Result 1 implies we have identified $E[Y_1-Y_0|B^*=1,X]$. Finally given $\Pr(X|B^*=1)$, it is possible to integrate over the vector of covariates X in order to identify $E[Y_1-Y_0|B^*=1]$. The conditional probability distribution in the above integral is easily recovered using Bayes' rule.

$$\Pr(X|B^*=1) = \Pr(B^*=1|X)\Pr(X)/\Pr(B^*=1)$$
$$= E[B^*|X]\Pr(X)/E[B^*]$$

As all of these parts are known, our estimator is complete.

Q.E.D..

When HMS use observed miscarriage as an instrument for birth in a regression of Y on B they are implicitly assuming $\gamma=1$ and $\beta=1$ while 2SLS implicitly computes $\Pr(B^*=1|X)$ and $\Pr(X|B^*=1)$ in estimation. The theorem demonstrates it is possible to recover the parameter of interest using only parts that are known or estimable. It's premise is that miscarriages are randomly assigned given X and that compliance with assigned miscarriages is not correlated with potential outcomes for latent-abortion women.

It is likely, however, that compliance is not ignorable. In particular, women having early abortions are more likely to have their assigned miscarriage censored than women having late abortions. If there are differences in the potential outcomes of women by desired length of pregnancy, the identification strategy is potentially defeated. The following corollary documents how this bias affects the estimator described above:

Corollary 2.1: *When women having early abortions have better potential outcomes than women having late abortions, the estimator above will underestimate the costs of teenage childbearing.*

Proof.

LATE(X) has the following representation,

$$E[Y_1-Y_0|B^*=1,X] = [\gamma\{E[Y|M=1,X]-E[Y|M^*=1,M=0,X]\}+\{E[Y|M^*=1,M=0,X]-E[Y|M=0,X\}/\beta]/E[B^*|X]$$

Take the derivative of LATE(x) with respect to the average outcome of non-compliers:

$$\begin{aligned}\delta E[Y_1-Y_0|B^*=1,X]/\delta E[Y|M^*=1,M=0,X]\} &= \{-\gamma+1/\beta\}/E[B^*|X] \\ &= \{-Pr(M^*=1|X)/Pr(M=1|X)+Pr(M^*=0|X)/Pr(M=0|X)\}/E[B^*|X] \\ &< 0\end{aligned}$$

where the last inequality follows because $E[B^*|X] > 0$ and the term in brackets is negative if and only if:

$$\{1-Pr(M^*=1|X)\}/\{1-Pr(M=0|X)\} < Pr(M^*=1|X)/Pr(M=1|X)$$

$$Pr(M=1|X)/\{1-Pr(M=0|X)\} < Pr(M^*=1|X)/\{1-Pr(M^*=1|X)\}$$

The last inequality follows because $Pr(M^*=1|X) > Pr(M=1|X)$ and $x/(1-x)$ is always decreasing in x. It follows that if $E[Y|M^*=1,M=0,X]$ is too small (because the censoring problem is more severe among early abortion types who have better outcomes), our estimates of the benefits of delaying childbearing will be too large.

Q.E.D..

Table V below indicates that women having early abortions have better outcomes than women having late abortions. If this ordering holds for potential outcomes, it is likely that the estimator above underestimates the true costs of teenage childbearing. To accommodate this potential problem, I will replace the average outcomes of non-compliers with the average outcomes of women having early

abortions (terminate the pregnancy in the first trimester) and implement this as a second non-parametric estimator below.

As pregnancies that end in abortion are severely misreported in fertility microdata, it is important to understand how this behavior might affect our parameters of interest. It turns out that when non-reporting is ignorable, the estimator is in principle unaffected, and when non-reporting is correlated with potential outcomes, it is likely that the estimators above are lower bounds on the true costs of teenage childbearing.

Corollary 2.2: *When $E[M^*]$ is constructed using external data, the estimator in Theorem 2 is not affected if latent-abortion types randomly fail to report pregnancies that are terminated*.

Proof.

First simplify LATE(X) from the previous corollary,

$$E[Y_1-Y_0|B^*=1,X] = [\gamma\{E[Y|M=1,X]-E[Y|A=1,X]\}+\{E[Y|A=1,X]-E[Y|M=0,X]\}/\beta]/E[B^*|X]$$

$$= \{(1-E[M^*])E[M]/E[M^*]E[B]\}\{E[Y|M=1,X]-E[Y|A=1,X]\}+E[Y|A=1,X]-E[Y|B=1,X]$$

The randomness of non-reporting implies E[Y|M=1,X], E[Y|A=1,X], and E[Y|B=1,X] are all unaffected by the failure of latent-abortion women to report pregnancies. Moreover, while underreporting inflates E[M] and E[B], it does not affect the ratio. As $E[M^*]$ is taken from an external data source, there is no effect on LATE(x). Finally, our estimator of the probability distribution of X given B=1 is also unaffected. Recall the following from above,

$$Pr(X|B^*=1) = E[B^*|X]Pr(X)/E[B^*]$$

While $E[B^*]$ and $E[B^*|X]$ are both affected by non-reporting of abortions (use $E[B] = E[B^*]E[1-M^*]$), the randomness of non-reporting implies the ratio is not.

Q.E.D.

The above corollary is qualified because I actually estimate $E[M^*|X]$ from the teen pregnancy sample in the analysis below. As the fraction of actual miscarriages is overstated and actual abortions is understated, it is likely that the estimated risk of miscarriage is overestimated in the analysis above. As the ratio $\{1-E[M^*|X])\}/E[M^*|X]$ is decreasing in the probability of assigned miscarriage, this implies that our non-parametric IV estimates below are likely too small in absolute value. It is also possible that non-random reporting of pregnancies makes the estimated benefits of delaying childbearing too large. When it is women with the best potential outcomes choose to not report abortions (or are less likely to report), this implies $E[Y|A=1,X]$ is too small. This affects LATE(x) in two ways: the first was demonstrated in Corollary 2.1 in underestimating $E[Y|M^*=1,M=0,X]$ but the second comes through underestimating $E[Y,M=0,X]$ and works in the same direction (see the first line of the proof to Corollary 2.1). It follows that the non-random reporting of abortions probably makes our estimated costs of teenage childbearing too small.

8. Main Empirical Results

Table IV illustrates important sample statistics for the *Teen Pregnancy Sample*. Almost 85 percent of pregnancies occur after a girls 15th birthday, the cutoff for the first risk factor of assigned miscarriage. While over 60 percent of girls had smoked 100 cigarettes in their lifetime, only 28 percent started smoking before the pregnancy, the second risk factor. Miscarriage occurred in about 9 percent of these pregnancies, abortion in 25 percent, and birth in 65 percent.[xiv] Table V presents some statistics that are useful to have when thinking about potential biases in each of the estimators presented below. Average outcomes are broken out by pregnancy outcome, and it should be clear that latent-abortion women generally have the best outcomes. Moreover, the table illustrates that women having early abortions usually have better outcomes than those having late abortions: there is almost a 10 percentage point differential in the probability of getting a high school diploma.

Table VI presents OLS and IV estimates of the costs of teenage childbearing for the *Teen Pregnancy Sample*. The first column indicates delaying birth would increase the probability of getting a high school diploma by about 15.5 percentage points. There are significant effects on labor force participation, annual income increasing by almost $3,500 and the probability of working increasing by almost 8 percentage points. Interestingly, delaying teenage childbearing would increase spouse income by more than $3000. The second column reports IV estimates using miscarriage as an instrument, and finds results similar to HMS. The probability of getting a high school diploma actually falls by almost 10 percentage points upon delaying childbearing, and the effects on labor force participation and marital matching are much smaller, although most coefficients are not statistically significant.

The non-parametric estimator described above is implemented in Table VII under two sets of identifying assumptions. The first column identifies the average outcome of latent-birth women who do not comply with assigned miscarriage with the average outcome of women who have abortions. The second column instead uses the average outcome of women who have abortions in the first trimester to reflect the fact that women not complying with assignment are mostly having early abortions. Point estimates indicate a significant bias in the IV estimator across the full range of outcomes. Most interesting is the strong negative effect of teenage childbearing on the probability of receiving a high school diploma that is very close to OLS, and contrasts to Ribar (1994) and HMS finding no significant effect.

Of course causal inference is more complicated than point estimates. In order to perform hypothesis tests that any of these estimated coefficients are significant, I construct 90 percent confidence intervals for the non-parametric estimators via bootstrapping. To do so, I randomly sample with replacement from the full *National Survey of Family Growth* cross-section. From this random sample I construct a new *Teen Pregnancy Sample* by selecting those girls who have a pregnancy before their 18th birthday and are at least 20 years old at the time of the survey. Given this data, I construct the two non-parametric estimators and repeat for 100 bootstrap replications. The confidence intervals are simply the 10th and 90th percentiles of the distribution of point estimates.

The confidence intervals demonstrate that while there is a significant bias when using IV estimates, only in the context of labor force participation are there statistically significant costs of teenage childbearing. The coefficient on a receiving a high school diploma includes zero in its 90 percent confidence interval. While several of the OLS and IV estimates can be rejected by the confidence intervals, they remain fairly wide. It should be noted, however, that in constructing non-parametric bounds for the causal effect of interest, Hotz, Mullin, and Sanders (1997) maintained that the causal effect of teenage childbearing on hours worked is positive. It turns out, however, that is one of the few outcomes for which the analysis suggests with high probability that the opposite is true.[xv]

It is possible, of course, that there is a great deal of heterogeneity in the true costs of delaying childbearing even for the sub-population of teenage girls that would bring the child to term in absence of a miscarriage.[xvi] To investigate this possibility I break out non-parametric estimates of the causal effect across the age at conception in Table VIII using the assumption that non-compliers have the average outcomes of women having early abortions (the assumption from the second column of Table VII). The table indicates that for girls having relatively later pregnancies, there is a causal link between childbearing and the probability of completing their high school degrees.

A closer look at the data indicates it is not all that surprising the costs of bringing their first pregnancy to term are small. For girls that don't bring an early pregnancy to term, 51 percent have another pregnancy before their 18th birthday and 28 percent actually bring a pregnancy to term that started before their 18th birthday (compared to 23 percent and 15 percent, respectively, for girls having a later first pregnancy). In this sense delaying childbearing through an abortion may not have much benefit because there is no change in behavior after the pregnancy. Moreover, zero costs of childbearing for these girls is consistent with the evidence presented by Ribar (1994), whose age at menarche instrument is mostly identified from younger girls. The positive point estimate in the first row, second column of the table (although not significant) might be explained by teenage childbearing having large costs for high school girls and relatively small costs for younger girls. Girls having their first pregnancy so young might be more likely to have family support and have more time to recover from the pregnancy to complete their high

school degrees. The decision to have the child before high school then might actually be a good thing – given behavior would not change if the pregnancy was delayed -- because terminating the pregnancy simply creates another, but at what might be a more difficult time in a girls adolescence.

The table also illustrates that the effects on labor market participation are largely driven by the younger girls, who also significantly reduce their chances of being married later in life by bringing a child to term at such a young age. The larger effects of teenage birth on spouse and family income for young girls are probably driven by this difference in marriage rates.

Finally, I break out the causal effect across the entire support of covariates in Table IX. The first column reports point estimates and confidence intervals for women who had the pregnancy after their 15th birthday and did not start to smoke before conception. The second column reports the same statistics for women who had late pregnancies but smoked before them while third column represents women who did not smoke but had early pregnancies. The final column represents women who had an early pregnancy and smoked before it occurred.[xvii] Women who did not start smoking and did not have their first pregnancy until after their 15th birthday actually have a statistically significant educational benefit from delaying childbearing, in contrast to women who smoked before a late pregnancy. On the other hand, women having early pregnancies may actually benefit from bringing the child to term, but these effects are no different from zero, and not very different across smoking before the pregnancy. While it is unlikely there is a causal link between smoking and the magnitude of costs of delaying teenage childbearing – it is likely smoking is correlated with potential outcomes -- it demonstrates that even among high school girls there is a great deal of heterogeneity in the size of these costs.

9. Conclusion

While OLS estimates of the cost of teenage childbearing on high school completion might indicate teenage childbearing is our most serious social problem, economists have struggled to find a causal link between these two. This paper demonstrates that when implemented correctly, using miscarriage as an instrument

for birth reveals clear evidence that childbearing makes graduation much tougher on high school girls. Moreover, this evidence seems to make sense in the context of the existing instrumental variables literature on female labor supply, as there is no effect of teenage birth on high school completion for very young girls, consistent with Ribar (1994). There is not clear evidence of what this implies for outcomes later in life: hours and employment seem to fall given a teenage birth, but the effect on income is imprecisely measured (although economically significant at 15 percent of its mean for high school girls). Finally, it should be emphasized that the causal effect of interest is an average and there is a great deal of variance even within the sub-population of high school girls in these costs. This implies that without radical changes in subsequent behavior (that would change potential outcomes), it is possible that for some girls bringing a child to term makes them better off all else equal. From a policy perspective this does not we should try to identify these costs on a case by case basis and push girls in one direction or another, but should work harder to change behavior that eliminates the benefits of delaying childbearing.

Banking Studies Department, Federal Reserve Bank of New York, 33 Liberty Street, New York, NY 10045;
Adam.Ashcraft@ny.frb.org

REFERENCES

ABADIE, A. (1999): "Semiparametric Instrumental Variable Methods for Causal Respond Models," Massachusetts Institute of Technology PhD Thesis.

ANGRIST, J. D. and B. EVANS (1998): "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size," *American Economic Review*, Volume 88 (3), pages 450-477.

ANGRIST, J. D., G. IMBENS, and D. B. RUBIN (1996): "Identification of Causal Effects using Instrumental Variables," *Journal of the American Statistical Association*, Volume 91 (434), pages 444-55.

BRONARS, S. G. and J. GROGGER (1994): "The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment," *American Economic Review*, Volume 84 (5), pages 1141-56.

HOTZ, V. J., S. W. MCELROY, and S. G. SANDERS (1997): "Assessing the Effects of Teenage Childbearing on Maternal Outcomes in the United States: Exploiting a Very Natural Experiment," Unpublished manuscript, University of Chicago.

HOTZ, V. J., MULLIN, C. H., and S. G. SANDERS (1997): "Bounding Causal Effects Using Data from a Contaminated Natural Experiment: Analyzing the Effects of Teenage Childbearing," *Review of Economic Studies*, Volume 64 (4), pages 575-603.

KLINE, J. and Z. STEIN (1987): "Epiidemiology of Chromosomal Abnormalities in Spontaneous Abortion," in Porter, I. and Hook, E. (eds.) *Human Embryonic and Fetal Death*, (New York: Academic Press).

KLINE, J. and Z. STEIN (1989): "Conception to Birth: Epidemiology of Prenatal Development," (New York: Oxford University Press).

LEVITT, S. and J. J. DONOHUE III (2001): "The Impact of Legalized Abortion on Crime," *Quarterly Journal of Economics*.

RIBAR, D. C. (1994): "Teenage Fertility and High School Completion," *Review of Economics and Statistics*, Volume 76 (3), pages 413-24.

	Treatment Offer	
	Miscarriage	No Miscarriage
Latent-Type Abortion	**Abortion Miscarriage**	**Abortion**
Birth	**Miscarriage**	**Birth**

Figure 1. Pregnancy outcome by a woman's latent-type and offer of treatment.

	Pregnancy Outcome	
	Miscarriage	No Miscarriage
Latent-Type Abortion	**Miscarriage**	**No Miscarriage** **Miscarriage**
Birth	**Miscarriage**	**No Miscarriage**

Figure 2. Offer of treatment by a woman's latent-type and pregnancy outcome.

Figure 3: Panel (a)

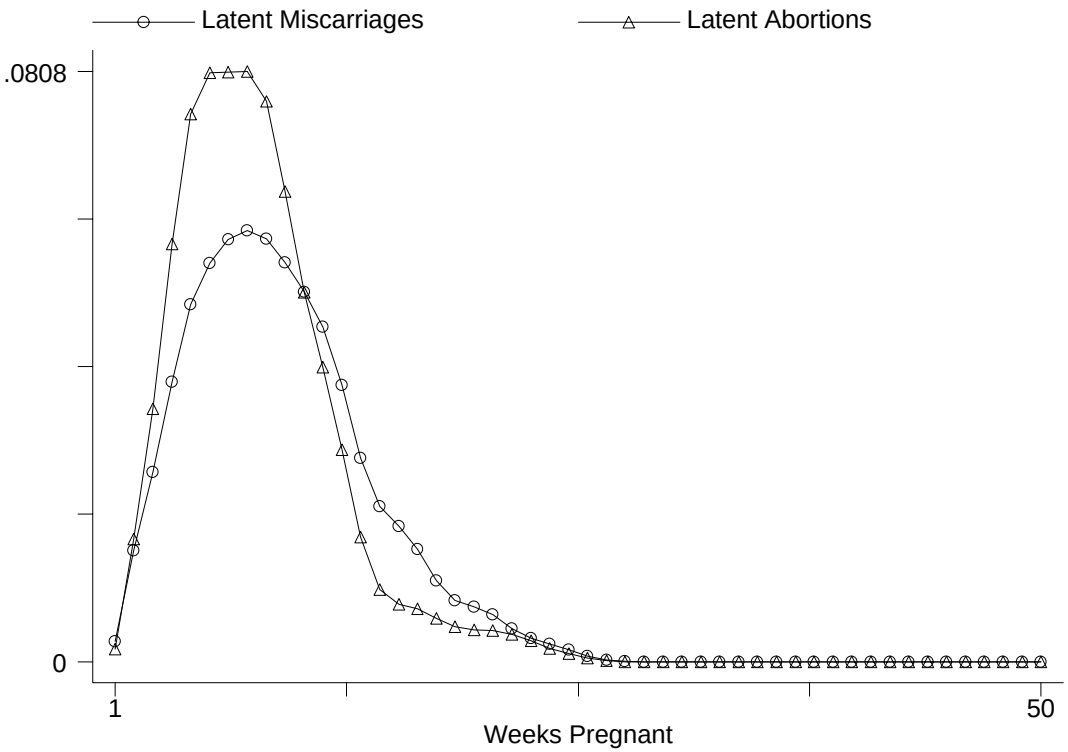

Figure 3: Panel (b)

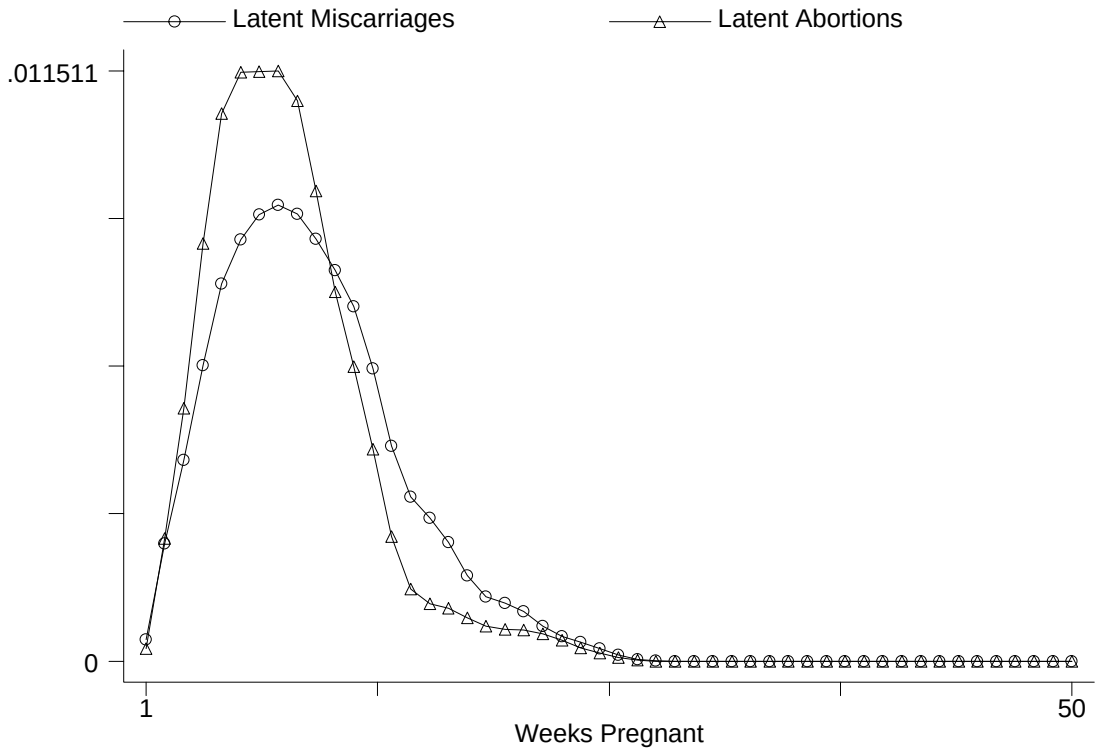

Table I
Descriptive Statistics

	Full Sample	Teen Pregnancy
Age at Interview	30.1	32.3
	(8.5)	(6.9)
Age 20 plus	0.849	1.00
	(0.358)	(0.000)
Teen Pregnancy	0.169	1.00
	(0.375)	(0.000)
Teen Birth	0.119	0.710
	(0.323)	(0.454)
White	0.706	0.593
	(0.455)	(0.491)
Hispanic	0.111	0.137
	(0.315)	(0.344)
Black	0.136	0.246
	(0.343)	(0.429)
Protestant	0.544	0.590
	(0.498)	(0.492)
Roman Catholic	0.343	0.304
	(0.475)	(0.460)
No Mother	0.043	0.092
	(0.204)	(0.289)
No Father	0.109	0.210
	(0.312)	(0.407)
Mother's Education	11.3	9.7
	(4.11)	(4.6)
Father's Education	10.6	8.2
	(5.25)	(5.6)
Working Mother	0.446	0.496
	(0.497)	(0.500)
Age at First Birth	21.7	17.2
	(4.9)	(2.95)
Currently Married	0.493	0.527
	(0.500)	(0.499)
Currently Divorced or Separated	0.123	0.241
	(0.329)	(0.428)
Currently Working	0.680	0.626
	(0.467)	(0.484)
Current Income	13,660	11,050
	(16,686)	(12,912)
Current Hours	25.9	24.6
	(20.5)	(21.3)
Current Spouse Income	18.270	17,214
	(25,002)	(21,173)
Current Family Income	38,144	31,193
	(32,484)	(26,722)
Has a High School Diploma?	0.791	0.685
	(0.407)	(0.465)
Observations	10,847	1913

Notes: Table reports sample means and standard deviations and employs sample weights. The *Full Sample* corresponds to all of the women surveyed in the *National Survey of Family Growth* while the *Teen Pregnancy Sample* represents women having a pregnancy before their 18th birthday who were at least 20 years old when surveyed in 1995.

Table II
Descriptive Statistics for the Pregnancy Panel

Age at Interview	34.4
	(6.4)
Hispanic	0.133
	(0.340)
White	0.666
	(0.472)
Black	0.160
	(0.366)
Roman Catholic	0.357
	(0.479)
Protestant	0.542
	(0.498)
Jewish	0.013
	(0.124)
No Mother	0.064
	(0.245)
No Father	0.141
	(0.348)
Mother's Education	10.3
	(4.4)
Father's Education	9.5
	(5.4)
Mother Worked Full Time?	0.414
	(0.493)
Age of Mother's First Birth?	20.9
	(4.3)
Age at Conception	23.6
	(5.4)
Married at Conception?	0.656
	(0.475)
Smoker	0.434
	(0.500)
Used IUD before Pregnancy	0.024
	(0.153)
Birth	0.709
	(0.454)
Abortion	0.136
	(0.343)
Miscarriage	0.135
	(0.341)
Stillbirth	0.008
	(0.089)
Ectopic Pregnancy	0.012
	(0.110)
Observations	20,904

Notes: The table reports sample means and standard and employs sample weights. Each observation corresponds to a pregnancy, representing 7660 women from the *Full Sample*.

Table IV
Some Evidence of a Censoring Problem

	Abortion	Miscarriage	Abortion	Miscarriage
Catholic	-0.264	NA	-0.023	NA
	(0.117)		(0.024)	
Agnostic	-0.0723	NA	0.014	NA
	(0.123)		(0.025)	
Protestant	-0.379	NA	-0.053	NA
	(0.116)		(0.024)	
Jewish	0.104)	NA	0.043	NA
	(0.156)		(0.031)	
Other	-0.172	NA	-0.000	NA
	(0.155)		(0.031)	
Married?	-0.858	NA	-0.231	NA
	(0.023)		(0.006)	
Pr(Abort)	NA	-0.076	NA	-0.044
		(0.034)		(0.021)
Age	NA	-0.009	0.002	-0.002
		(0.003)	(0.001)	(0.001)
Smoker?	NA	0.055	0.025	0.013
		(0.021)	(0.005)	(0.005)
White	NA	0.090	-0.005	0.018
		(0.056)	(0.013)	(0.010)
Hispanic	NA	-0.034	0.015	-0.011
		(0.043)	(0.014)	(0.008)
Black	NA	-0.084	-0.030	-0.020
		(0.049)	(0.014)	(0.007)
No Mother	NA	0.043	0.025	0.011
		(0.054)	(0.013)	(0.012)
No Father	NA	0.052	0.062	0.014
		(0.045)	(0.010)	(0.005)
Father's Education	NA	0.001	0.006	0.001
		(0.005)	(0.001)	(0.001)
Mother's Education	NA	0.009	0.007	0.002
		(0.003)	(0.001)	(0.001)
Mother Working Full Time	NA	0.001	0.014	-0.001
		(0.021)	(0.005)	(0.006)
Pregnancy Order	NA	0.070	0.014	0.019
		(0.009)	(0.002)	(0.003)
Observations	20,904	20,904	20,904	20,904

Notes: The first column reports coefficients and standard errors from a probit model of abortion as a pregnancy outcome on plausibly exogenous background characteristics that predict abortion. The second column is a probit model of miscarriage on factors known to cause miscarriage, other covariates as controls – including dummies for age at conception which are suppressed – and the probability of abortion as predicted from the first column. Standard errors in each of these columns are corrected for heteroskedasticity, clustered at the person level, and neither model uses sample weights. The third column is an OLS regression of abortion on the instruments used in the first column plus the full set of covariates. The final column is an IV regression of miscarriage on abortion instrumented by parent's religion and marital status at conception with the same set of covariates. Standard errors are again corrected for heteroskedasticity and clustered at the person level, but here each model employs sample weights.

Table IV
Teen Pregnancy Outcomes

Conception at Age 11	0.003
	(0.056)
Conception at Age 12	0.009
	(0.094)
Conception at Age 13	0.035
	(0.185)
Conception at Age 14	0.091
	(0.287)
Conception at Age 15	0.176
	(0.381)
Conception at Age 16	0.305
	(0.461)
Conception at Age 17	0.380
	(0.486)
Have smoked 100 cigarettes in lifetime?	0.621
	(0.486)
Do You Smoke now?	0.474
	(0.499)
Started Smoking Before Pregnancy?	0.282
	(0.450)
Married at conception?	0.124
	(0.330)
Birth	0.648
	(0.478)
Abortion	0.250
	(0.433)
Miscarriage	0.089
	(0.284)
Stillbirth	0.008
	(0.090)
Ectopic Pregnancy	0.005
	(0.070)
Had another pregnancy?	0.886
	(0.318)
Age at next pregnancy	19.6
	(3.5)
Observations	1913

Notes: The table reports sample means and standard deviations and employs sample weights. The sample corresponds to the *Teen Pregnancy Sample*.

Table V
Simple Consequences by Pregnancy Outcome

	Birth	Abortion	Early Abortion	Late Abortion	Miscarriage	Other
High School Diploma	0.414	0.726	0.746	0.654	0.458	0.409
	(0.493)	(0.447)	(0.436)	(0.478)	(0.500)	(0.503)
Annual Income	9,918	13,948	14,164	13,178	10,949	12,568
	(11,852)	(14,721)	(14,925)	(14,020)	(13,402)	(15,036)
Currently Working?	0.595	0.685	0.691	0.662	0.656	0.515
	(0.502)	(0.465)	(0.462)	(0.475)	(0.476)	(0.511)
Currently Married	0.530	0.511	0.504	0.536	0.553	0.512
	(0.499)	(0.500)	(0.501)	(0.501)	(0.499)	(0.511)
Divorced or Separated	0.251	0.222	0.243	0.146	0.242	0.084
	(0.434)	(0.416)	(0.430)	(0.355)	(0.430)	(0284)
Age at Next Pregnancy	19.7	20.0	19.9	20.2	18.2	19.4
	(3.4)	(3.8)	(3.7)	(4.3)	(2.7)	(3.1)
Age at First Birth	16.0	20.8	20.8	21.1	18.9	19.3
	(1.1)	(3.9)	(3.8)	(4.2)	(3.0)	(3.1)
Current Hours	23.4	26.6	26.5	26.9	27.0	27.8
	(21.1)	(20.9)	(20.9)	(21.1)	(22.2)	(27.0)
Spouse Annual Income	15,623	21,634	22,052	20,145	16,729	15,013
	(19,265)	(26,059)	(26,590)	(24,139)	(17,868)	(17,158)
Family Annual Income	28,634	38,489	39,217	35,894	29,654	29,231
	(25,119)	(30,001)	(30,280)	(28,986)	(24,942)	(25,222)
Observations	1313	414	311	103	163	23

Note: Statistics represent the sample mean and standard deviation and correspond to the *Teen Pregnancy Sample*. All statistics employ sample weights. An early abortion is defined by an abortion in the first 3 months of the pregnancy, and late is defined as any abortion thereafter. Other pregnancy outcomes include an ectopic pregnancy or stillbirth.

Table VI
Simple Estimates of the Costs of Teenage Childbearing

	OLS	IV
High School Diploma	-0.155	0.092
	(0.028)	(0.066)
Annual Income	-3,414	-438
	(747)	(1,783)
Currently Working	-0.077	-0.054
	(0.028)	(0.068)
Currently Married	0.018	-0.013
	(0.028)	(0.068)
Currently Divorced or Separated	0.018	-0.042
	(0.024)	(0.061)
Weekly Hours	-3.77	-4.75
	(1.26)	(3.14)
Spouse Annual Income	-3,147	2,164
	(1,317)	(2,461)
Family Annual Income	-5,223	3,476
	(1,650)	(3,602)

Notes: The table reports coefficients and standard errors from regressions of each dependent variable on a dummy for birth. Regressions use sample weights on the *Teen Pregnancy Sample*, and are corrected for heteroskedasticity. Covariates include age at interview, dummies for age at conception, age at menarche, use of an IUD or smoking before pregnancy, all of the family background variables listed in Table 1, dummies for race, and a dummy for survey conducted by phone interview. The first column reports OLS coefficients while the second column reports IV estimates when using miscarriage as an instrument for birth.

Table VII
Non-Parametric IV Estimates

	E[Y| M*=1,M=0] = E[Y| A=1]	E[Y| M*=1,M=0] = E[Y| A=1,t=E]
High School Diploma	-0.060	-0.069
	[-0.148,0.019]	[-0.150,0.013]
Annual Income	-1,312	-1368
	[-3370,575]	[-3195,515]
Currently Working	-0.079	-0.080
	[-0.163,-0.016]	[-0.162,-0.018]
Currently Married	-0.044	-0.040
	[-0.117,0.034]	[-0.117,0.041]
Currently Divorced or Separated	0.018	0.001
	[-0.059,0.098]	[-0.069,0.098]
Weekly Hours	-4.68	-4.57
	[-8.12,-1.78]	[-7.94,-1.60]
Spouse Annual Income	-1,281	-1456
	[-4169,600]	[-4248,284]
Family Annual Income	-1,459	-1,773
	[-4343,1883]	[-4897,1595]

Notes: The table reports estimates of the causal effect of teenage childbearing and 90 percent confidence intervals. The first column reports non-parametric IV estimates described in the text using the average outcome of women having abortions as the average outcome of non-compliers. The second column reports non-parametric estimates using the average outcome of women having abortions in the first trimester as the average outcome of non-compliers.

Table VIII
Non-Parametric IV Estimates by Age at Conception

	18 < Age <= 15	Age < 15
High School Diploma	-0.100 [-0.176,-0.0165]	0.140 [-0.016,0.297]
Annual Income	-1,655 [-3777,90]	390 [-2204,3581]
Currently Working	-0.055 [-0.149,0.020]	-0.255 [-0.416,0.003]
Currently Married	-0.015 [-0.101,0.074]	-0.224 [-0.421,-0.021]
Currently Divorced or Separated	0.018 [-0.054,0.085]	-0.054 [-0.230,0.154]
Weekly Hours	-3.93 [-7.40,-0.33]	-9.16 [-14.79,-0.72]
Spouse Annual Income	-698 [-3548,1582]	-6454 [-14110,1243]
Family Annual Income	579 [-3005,3446]	-18,045 [-28394,-6627]

Notes: The table reports estimates of the causal effect of teenage childbearing and 90 percent confidence intervals for the *Teen Pregnancy Sample*. Each column reports non-parametric IV estimates described in the text using the average outcome of women having early abortions as the average outcome of non-compliers. The first column represents girls having their first pregnancy between the ages of 15 and 18 while the second column represents girls having this pregnancy before their 15th birthday.

Table IX
Non-Parametric IV Estimates by Age at Conception and Smoking Before Pregnancy

	Not Smoker and 18 > Age >=15	Smoker and 18 > Age >= 15	Not Smoker and Age < 15	Smoker and Age < 15
High School Diploma	-0.134 [-0.224,-0.021]	-0.014 [-0.208,0.120]	0.150 [-0.099,0.320]	0.116 [-0.238,0.543]
Annual Income	-673 [-2385,1511]	-4,056 [-10951,1511]	-506 [-3554,3036]	6449 [86,16276]
Currently Working	-0.059 [-0.143,0.022]	-0.046 [-0.251,0.022]	-0.312 [-0.519,-0.098]	0.050 [-0.499,0.791]
Currently Married	0.019 [-0.065,0.111]	-0.090 [-0.267,0.111]	-0.319 [-0.552,0.119]	0.344 [-0.213,0.786]
Currently Divorced or Separated	-0.010 [-0.084,0.068]	0.086 [-0.055,0.68]	-0.011 [-0.222,0.212]	-0.349 [-0.895,0.418]
Weekly Hours	-2.82 [-7.02,1.67]	-6.66 [-16.7,-0.74]	-11.87 [-18.6,-2.63]	6.33 [-18.6,43.46]
Spouse Annual Income	-1,244 [-4567,1400]	636 [-8968,5767]	-6666 [-15196,1586]	-6769 [-323211,20134]
Family Annual Income	-1792 [-5936,2400]	6377 [-4011,2400]	-21669 [-34002,6983]	1092 [-20411,22120]

Notes: The table reports estimates of the causal effect of teenage childbearing and 90 percent confidence intervals for each point in the support of covariates X. Each column reports non-parametric IV estimates described in the text using the average outcome of women having early abortions as the average outcome of non-compliers.

[i] While both instruments affect women that would bring a child to term, Ribar (1994) implicitly focuses on much younger women than HMS (1997) as most of the variation in age at menarche occurs between the ages of 12 and 14 compared to a median age of 16 in the HMS sample.

[ii] It seems unlikely that such support exists as girls having a teenage pregnancy are much more likely to be missing a parent, are exposed to much less parental education, and are more likely to have a mother working full-time. See Table I below for more details.

[iii] This implies that miscarriages and abortions affect potential outcomes (Y_0, Y_1) in exactly the same way. The presumption here is that the first-order differences in potential outcomes across the three pregnancy outcomes are between births and non-births.

[iv] Unfortunately there is no way to compare the estimated parameters across instrumental variables strategies because there are no common outcomes of interest between them. Levitt and Donahue (2001) focus on criminal activity of the woman's children while Ribar (1994) and HMS (1997) focus on labor market and educational outcomes of the woman. It is difficult to imagine effects on children without a driving effect on educational, labor market, or marital outcomes of mothers.

[v] See, for example, studies by Kline and Stein (1987) and Kline, Stein, and Susser (1989).

[vi] There are actually several experiments. Discretize the support of X into a finite number of points. Conceptually there is one experiment for each point in this support, and the causal effect of interest is an average over each of these experiments.

[vii] The NLSY is an annual survey originating in 1979 of a nationally representative sample of youths who were between 14 and 21 years of age in 1979. A retrospective pregnancy history was administered in 1984 and repeated every 2 years thereafter. A self-administered questionnaire was implemented in 1984 to record all abortions that had not been recorded in open NLSY interviews prior to that date. HMS exploited a sample of 980 women who reported a pregnancy prior to their 18th birthday, including those women in oversamples of blacks and Hispanics. HMS do not report using weights in their estimation, but break out all statistics and regressions by race.

[viii] Note that 36 percent of women have only had one pregnancy, 28 percent two pregnancies, 18 percent three pregnancies, and 18 percent more than three pregnancies.

[ix] These statistics are computed by comparing the total number of abortions implied by NSFG data with institutional sources, and are reported in user's documentation. It is difficult to compare or interpret the statistics by HMS with those I the NSFG because HMS do not use sample weights and the NLSY suffers from severe attrition problems.

[x] In principle the assumption requires conditioning on X. I abstract away from this here in constructing descriptive evidence, but take a much more rigorous approach below in estimating the probability distribution function separately for each point in the support of X.

[xi] An alternative strategy would be to use medical data on the probability of assigned miscarriage, which pins down the fraction of latent-birth women (see below) and can then be used to estimate these latent distributions. As there is at best a rough mapping between available data on the risk of miscarriage and the covariates X used in the analysis below, I prefer this approach.

[xii] Another important issue here is that it is inappropriate to use sample weights in constructing these probability density functions (as they are appropriate only when using a cross-section of data). Instead of constructing synthetic panel weights, I choose not to use weights at all here. In the analysis below when I will be working with a single cross-section of pregnancies they will be used.

[xiii] HMS assume girls who reported drinking before the pregnancy had 1-2 alcoholic drinks per day during the pregnancy, which likely doubles the risk of miscarriage.

[xiv] I simply ignore other outcomes instead of calling them births or miscarriages, which are infrequent.

[xv] The authors also conclude there is a positive effect on income, and the results above suggest there is a sizeable bias in this outcome as well, but the confidence interval is too large to say anything too precise.

[xvi] HMS actually break out all of their results by race. I presume part of this is to deal with the oversampling of minorities in the NLSY, but the authors do find results to vary across race.

[xvii] In the original *Teen Pregnancy Sample* there were 1213, 415, 245, and 40 girls in each of these bins, respectively.