On-Line Appendix to The Consequences of Teenage Childbearing: Consistent Estimates When Abortion Makes Miscarriage Nonrandom

Adam Ashcraft
Federal Reserve Bank of New York

Iván Fernández-Val
Boston University

Kevin Lang
Boston University, NBER and IZA

August 30, 2012

This supplementary appendix to Ashcraft, Fernandez-Val, and Lang contains the proofs of the main results of the paper, and the formal analysis for the effects of underreporting of abortions, misreporting of abortions, underreporting of miscarriages, and multiple teen pregnancies, which are outlined in the paper.

A Proofs of Section 1

A.1 Preliminary results

The following lemmas will be used in the proofs of the results of Section 1. The first lemma links the conditional ASFs with moments of the potential outcomes in populations defined by abortion and miscarriages types.

Lemma 1 *Under Assumptions 2 and 3, for* $d \in \{0, 1\}$,

$$\mu(d, X) = \mathbb{E}[Y^*(d) \mid A^* = 0, M^* = 1 - d, X], \quad a.s \tag{24}$$

Proof of Lemma 1 By definition of $\mu(d, x)$ and Assumption 2

$$\mu(d, x) = \mathbb{E}[Y^*(d) \mid A^* = 0, X = x] = \mathbb{E}[Y^*(d) \mid A^* = 0, M^* = 1 - d, X = x]. \tag{25}$$

□

The following lemma gives a system of equations linking moments of the observed variables to moments of the structural variables.

Lemma 2 (Moment conditions) *Under Assumptions 1–3, for* $t \in \{1, \ldots, \bar{T} - 1\}$ *and* $F_A^* =$

$f \in \{0,1\}$,

$$p_A(X,t) = p^*_{fA}(X,t)\{\sum\nolimits_{s=t}^{\bar{T}} p^*_{fM}(X,s) - [1-f]p^*_{fM}(X,t)\}, \quad (26)$$

$$p_M(X,t) = p^*_{fM}(X,t)\{\sum\nolimits_{s=t}^{\bar{T}} p^*_{fA}(X,s) - fp^*_{fA}(X,t)\}, \quad (27)$$

$$S_M(X,t) + S_A(X,t) = S^*_{fM}(X,t)S^*_{fA}(X,t) + 1, \quad (28)$$

$$S_M(X,t-1) + fS_A(X,t) + (1-f)S_A(X,t-1) = \quad (29)$$

$$S^*_{fM}(X,t-1)[fS^*_{fA}(X,t) + (1-f)S^*_{fA}(X,t-1)] + 1, \quad (30)$$

$$fS_M(X,t-1) + (1-f)S_M(X,t) + S_A(X,t-1) =$$

$$[fS^*_{fM}(X,t-1) + (1-f)S^*_{fM}(X,t)]S^*_{fA}(X,t-1) + 1, \quad (31)$$

$$\mathbb{E}[Y \mid D=1, X] = \mathbb{E}[Y^*(1) \mid A^* = 0, M^* = 0, X],$$

$$\mathbb{E}[Y \mid T_A = t, X] = \mathbb{E}[Y^*(0) \mid T^*_A = t, T^*_M, X], \quad (32)$$

$$\mathbb{E}[Y \mid T_M = t, X]p_M(X,t) = p^*_{fM}(X,t)\{\sum\nolimits_{s=t}^{\bar{T}} p^*_{fA}(X,s)\mathbb{E}[Y^*(0) \mid T^*_A = s, T^*_M = t, X]$$
$$- fp^*_{fA}(X,t)\mathbb{E}[Y^*(0) \mid T^*_A = t, T^*_M = t, X]\}, \quad a.s. \quad (33)$$

Proof of Lemma 2 We drop the dependence on X and f from all the functions to simplify the notation.

The first result follows from

$$p_A(t) = \sum_{s=t+1}^{\bar{T}} \mathbb{P}[T^*_A = t, T^*_M = s] + f\mathbb{P}[T^*_A = t, T^*_M = t] = p^*_A(t)\{\sum_{s=t}^{\bar{T}} p^*_M(s) - [1-f]p^*_M(t)\}, \quad (34)$$

where the first equality uses the law of total probability, and second equality uses Assumption 2. The second result follows similarly.

The third result follows from the equality of events $\{T_A > t, T_M > t\} = \{T^*_A > t, T^*_M > t\}$, independence of $\{T^*_A > t\}$ and $\{T^*_M > t\}$ conditional on X, and $\{T_A \leq t, T_M \leq t\} = \emptyset$ for $t < \bar{T}$ by definition of T_M and T_A, so that

$$\mathbb{P}[T^*_A > t, T^*_M > t] = S^*_A(t)S^*_M(t), \quad (35)$$

and

$$\mathbb{P}[T_A > t, T_M > t] = 1 - \mathbb{P}[T_A \leq t \text{ or } T_M \leq t] = S_A(t) + S_M(t) - 1. \quad (36)$$

The fourth and fifth results follow similarly.

The sixth result follows from Assumption 1 since $D = (1-A^*)(1-M^*)$, so that $\{D = 1\} = \{A^* = 0, M^* = 0\}$.

The seventh result follows from

$$\mathbb{E}[Y \mid T_A = t] = \mathbb{E}[Y^*(0) \mid T_A^* = t] = \mathbb{E}[Y^*(0) \mid T_A^* = t, T_M^*], \tag{37}$$

where the first equality uses Assumption 1 and second equality uses Assumption 2.

The last result follows from

$$\begin{aligned} \mathbb{E}[Y \quad \mid \quad & T_M = t]p_M(t) = \mathbb{E}[Y^*(0) \mid T_M^* = t, T_A^* > t \text{ or } T_A^*(1-f) = t, F_A^* = f]p_M(t) \\ = \quad & \{\sum_{s=t}^{\bar{T}} p_A^*(s)\mathbb{E}[Y^*(0) \mid T_A^* = s, T_M^* = t] - f p_A^*(t)\mathbb{E}[Y^*(0) \mid T_A^* = t, T_M^* = t]\} \\ & * p_M(t)/\{\sum_{s=t}^{\bar{T}} p_A^*(s) - f p_A^*(t)\}, \end{aligned} \tag{38}$$

where the first equality uses Assumption 1, and second equality uses Assumption 2 and the law of total probability. To obtain the final expression we use the second result of the Lemma. □

A.2 Proof of Proposition 1

The proposition follows from standard results for competing risks models in discrete time with independent risks (e.g., Kalbfleisch and Prentice, 2002, chap. 8), properly adjusted to deal with within-week ties in scheduled abortions and miscarriages. □

A.3 Proof of Theorem 1

We drop the dependence on X and f from all the functions to simplify the notation.

The result for $\mu(1)$ follows from Lemmas 1 and 2 since

$$\mu(1) = \mathbb{E}[Y^*(1) \mid A^* = 0, M^* = 0] = \mathbb{E}[Y \mid D = 1]. \tag{39}$$

For $\mu(0)$, by Lemma 1 and law of total probability

$$\mu(0) = \mathbb{E}[Y^*(0) \mid A^* = 0, M^* = 1] = [1 - S_M^*(\bar{T}-1)]^{-1} \sum_{t=1}^{\bar{T}-1} \mathbb{E}[Y^*(0) \mid T_A^* = \bar{T}, T_M^* = t] p_M^*(t), \tag{40}$$

where we use that $\sum_{t=1}^{\bar{T}-1} p_M^*(t) = 1 - S_M^*(\bar{T}-1)$.

By Lemma 2, for $t \in \{1, .., \bar{T}-1\}$

$$\mathbb{E}[Y \mid T_M = t]p_M(t) = p_M^*(t)\{S_A^*(\bar{T}-1)\mathbb{E}[Y^*(0) \mid T_A^* = \bar{T}, T_M^* = t] + \sum_{s=t}^{\bar{T}-1} p_A^*(s)Y_A(s) - fp_A^*(t)Y_A(t)\}, \tag{41}$$

where we use that $p_A^*(\bar{T}) = S_A^*(\bar{T}-1)$ and $\mathbb{E}[Y^*(0) \mid T_A^* = s, T_M^* = t] = Y_A(s)$, for $s \in \{t, ..., \bar{T}-1\}$. Adding over $t \in \{1, .., \bar{T}-1\}$

$$Y_M[1 - S_M(\bar{T}-1)] = S_A^*(\bar{T}-1)[1 - S_M^*(\bar{T}-1)]\mu(0) + \sum_{t=1}^{\bar{T}-1} p_A^*(t)[\sum_{s=1}^{t} p_M^*(s) - fp_M^*(t)]Y_A(t), \tag{42}$$

where we use that $\sum_{t=1}^{\bar{T}-1} p_M(t) = 1 - S_M(\bar{T}-1)$, equation (40), and interchange the indexes of the sums in the second term of the right hand side.

The result follows by solving for $\mu(0)$. □

A.4 Proof of Theorem 2

We drop the dependence on X and f from all the functions to simplify the notation.

The expression of γ^* follows by the law of total probability.

Let $Y_D = \mathbb{E}[Y \mid D = 1]$ and $Y_A = \mathbb{E}[Y \mid A = 1]$. Then, from Theorem 1

$$\Delta = \mu(1) - \mu(0) = Y_D - Y_M + \frac{\gamma^*[1 - S_A^*(\bar{T}-1)]}{S_A^*(\bar{T}-1)}[Y_A - Y_M], \tag{43}$$

where we use that $Y_A(t) = \mathbb{E}[Y^*(0) \mid T_A^* = t] = \mathbb{E}[Y^*(0) \mid A^* = 1, M^* = 0] = Y_A$ by Lemma 2, Assumption 4, and Assumption 2; $\sum_{t=1}^{\bar{T}-1} p_A^*(t)[\sum_{s=1}^{t} p_M^*(s) - fp_M^*(t)] = [1 - S_M^*(\bar{T}-1)][1 - S_A^*(\bar{T}-1)]\gamma^*$ by the expression of γ^*; and $1 - S_M(\bar{T}-1) = [1 - S_M^*(\bar{T}-1)]\{S_A^*(\bar{T}-1) + \gamma^*[1 - S_A^*(\bar{T}-1)]\}$ by the law of total probability and the expression of γ^*.

The OLS and IV estimands can be expressed as

$$\Delta_{ols} = \mathbb{E}[Y \mid D = 1] - \mathbb{E}[Y \mid M = 1] = Y_D - Y_M, \tag{44}$$

and

$$\Delta_{iv} = Y_D - Y_M + \frac{[1 - S_A^*(\bar{T}-1)]\{1 - \gamma^* + \gamma^* S_M^*(\bar{T}-1)\}}{S_A^*(\bar{T}-1)S_M^*(\bar{T}-1)}[Y_A - Y_M], \tag{45}$$

where we use that $\mathbb{E}[D \mid M = 0] = S_A^*(\bar{T}-1)S_M^*(\bar{T}-1)/S_M(\bar{T}-1)$, $\mathbb{E}[D \mid M = 1] = 0$, $\mathbb{E}[Y \mid M = 0] = Y_A + S_A^*(\bar{T}-1)S_M^*(\bar{T}-1)[Y_D - Y_A]/S_M(\bar{T}-1)$, and $S_M(\bar{T}-1) = [1 - S_A^*(\bar{T} -$

$1)]\{1-\gamma^*+\gamma^* S_M^*(\bar{T}-1)\} + S_A^*(\bar{T}-1)S_M^*(\bar{T}-1)$. Then,

$$\frac{[1-\gamma^*]\Delta_{ols} + \gamma^* S_M^*(\bar{T}-1)\Delta_{iv}}{1-\gamma^*+\gamma^* S_M^*(\bar{T}-1)} = Y_D - Y_M + \frac{\gamma^*[1-S_A^*(\bar{T}-1)]}{S_A^*(\bar{T}-1)}[Y_A - Y_M] = \Delta. \tag{46}$$

To show that $\Delta_{ols} \leq \Delta \leq \Delta_{iv}$, it is sufficient to show that $\Delta_{ols} \leq \Delta_{iv}$, because Δ is a convex linear combination of Δ_{ols} and Δ_{iv}. A sufficient condition for this inequality is that $Y_A \geq Y_M$. This result follows from Assumptions 2 and 5 because $Y_A = \mathbb{E}[Y^*(0) \mid A^* = 1]$ and

$$Y_M = \frac{[1-S_M^*(\bar{T}-1)]\{S_A^*(\bar{T}-1)\mathbb{E}[Y^*(0) \mid A^*=0] + \gamma^*[1-S_A^*(\bar{T}-1)]\mathbb{E}[Y^*(0) \mid A^*=1]\}}{1-S_M(\bar{T}-1)} \leq Y_A. \tag{47}$$

We add back the dependence on f to show the inequality $\Delta_0(x) \leq \Delta_1(x)$. By the previous result, we need to show that

$$\frac{\gamma_0^*[1-S_{0A}^*(\bar{T}-1)]}{S_{0A}^*(\bar{T}-1)} \leq \frac{\gamma_1^*[1-S_{1A}^*(\bar{T}-1)]}{S_{1A}^*(\bar{T}-1)}. \tag{48}$$

This result follows because by Proposition 1, $S_{0A}^*(\bar{T}-1) \geq S_{1A}^*(\bar{T}-1)$ and $S_{0M}^*(\bar{T}-1) \leq S_{1M}^*(\bar{T}-1)$, so that using the expression of γ_f^*,

$$\begin{aligned}\gamma_0^*[1-S_{0A}^*(\bar{T}-1)] &= \frac{1-S_M(\bar{T}-1)-S_{0A}^*(\bar{T}-1)}{1-S_{0M}^*(\bar{T}-1)} \\ &\leq \frac{1-S_M(\bar{T}-1)-S_{1A}^*(\bar{T}-1)}{1-S_{1M}^*(\bar{T}-1)} = \gamma_1^*[1-S_{1A}^*(\bar{T}-1)]. \end{aligned} \tag{49}$$

□

A.5 Proof of Theorem 3

By definition of the ATE with $F_A^* = f$

$$\Delta_f = \sum_{x\in\mathcal{X}} \mathbb{P}[X = x \mid A^* = 0, F_A^* = f]\Delta_f(x). \tag{50}$$

Then, by the Bayes rule

$$\mathbb{P}[X = x \mid A^* = 0, F_A^* = f] = \frac{S_{fA}^*(x,\bar{T}-1)\mathbb{P}[X=x]}{\sum_{x\in\mathcal{X}} S_{fA}^*(x,\bar{T}-1)\mathbb{P}[X=x]}. \tag{51}$$

□

B Underreporting and Misreporting

B.1 Underreporting of abortions

We show that underreporting abortions affects the identification of the effects of interest, even when underreporting is random conditional on characteristics. Let R_A^* denote the indicator for reported abortion. We assume that reporting is independent of potential outcomes and scheduled weeks conditional on the characteristics X.

Assumption 7 (Conditionally random abortion underreporting)

$$R_A^* \perp\!\!\!\perp [Y^*(0), Y^*(1), T_M^*, T_A^*] \mid X \quad a.s.$$

We denote with tilde all the variables, probabilities and survival functions in the population with underreporting; for example, $\tilde{p}_A(x,t) := \mathbb{P}(\tilde{T}_A = t \mid \tilde{X} = x) = \mathbb{P}(T_A = t \mid X = x, R_A^* = 1)$. We assume that women that do not report abortions are not observed in the population of pregnant teenagers. An implication of this assumption is that there is no misclassification of types and treatment in the population with underreporting. The source of the bias is that underreporting has an asymmetric effect depending on whether miscarriage is scheduled before or after abortion.

The following result shows that the weighted average of the OLS and IV estimands in Theorem 2 does not estimate the conditional ATE with underreporting of abortions. Under Assumption 5, the bias is negative towards the OLS estimand.

Theorem 4 (Conditional ATE with underreporting of abortions) *Under Assumptions 1–5, and 7, if* $F_A^* = f \in \{0,1\}$ *and* $\mathbb{P}(R_A^* = 1 \mid X) > 0$ *a.s.,*

$$\frac{[1-\tilde{\gamma}_f^*(X)]\tilde{\Delta}_{ols}(X) + \tilde{\gamma}_f^*(X)\tilde{S}_{fM}^*(X,\bar{T}-1)\tilde{\Delta}_{iv}(X)}{1-\tilde{\gamma}_f^*(X) + \tilde{\gamma}_f^*(X)\tilde{S}_{fM}^*(X,\bar{T}-1)} \le \Delta_f(X), \quad a.s. \tag{52}$$

Proof To simplify the notation, we drop the dependence on X and f from all the functions and assume that there are no ties, that is $\mathbb{P}(T_M^* = T_A^* \mid X) = 0$ a.s.

Note that $\tilde{\mu}(1) = \mu(1)$ because underreporting does not affect women of birth type, $A^* = M^* = 0$. However, underreporting might affect $\tilde{\mu}(0)$ because it alters some of the relationships between the observed and structural variables in Assumption 1.

By a similar argument to the proof of Proposition 1, but without using the results of Lemma 2

$$\begin{aligned}\mathbb{E}[\tilde{Y}^*(0) \mid & \tilde{A}^* = 0, \tilde{M}^* = 1] \qquad (53)\\ &= \frac{[1-\tilde{S}_M(\bar{T}-1)]\tilde{Y}_M - \tilde{\gamma}^*[1-\tilde{S}_M^*(\bar{T}-1)][1-\tilde{S}_A^*(\bar{T}-1)]\mathbb{E}[\tilde{Y}^*(0) \mid \tilde{A}^*=1, \tilde{T}_M^* < \tilde{T}_A^*]}{\tilde{S}_A^*(\bar{T}-1)[1-\tilde{S}_M^*(\bar{T}-1)]}.\end{aligned}$$

Let

$$\tilde{\mu}(0) = \tilde{\mu}(1) - \frac{[1-\tilde{\gamma}^*]\tilde{\Delta}_{ols} + \tilde{\gamma}^*\tilde{S}_M^*(\bar{T}-1)\tilde{\Delta}_{iv}}{1-\tilde{\gamma}^* + \tilde{\gamma}^*\tilde{S}_M^*(\bar{T}-1)}. \qquad (54)$$

By a similar argument to the proof of Theorem 2

$$\begin{aligned}\tilde{\mu}(0) &= \frac{[1-\tilde{S}_M(\bar{T}-1)]\tilde{Y}_M - \tilde{\gamma}^*[1-\tilde{S}_M^*(\bar{T}-1)][1-\tilde{S}_A^*(\bar{T}-1)]\tilde{Y}_A}{\tilde{S}_A^*(\bar{T}-1)[1-\tilde{S}_M^*(\bar{T}-1)]} \qquad (55)\\ &= \frac{\tilde{\gamma}^*[1-\tilde{S}_A^*(\bar{T}-1)]}{\tilde{S}_A^*(\bar{T}-1)}\{\mathbb{E}[\tilde{Y}^*(0) \mid \tilde{A}^*=1, \tilde{T}_M^* < \tilde{T}_A^*] - \mathbb{E}[\tilde{Y}^*(0) \mid \tilde{A}^*=1, \tilde{T}_M^* > T_A^*]\}\\ +\mathbb{E}[\tilde{Y}^*(0) \mid & \tilde{A}^*=0, \tilde{M}^*=1] \geq \mathbb{E}[Y^*(0) \mid A^*=0, M^*=1] = \mu(0),\end{aligned}$$

where in the first equality we use that $1-\tilde{S}_M(\bar{T}-1) = [1-\tilde{S}_M^*(\bar{T}-1)]\{\tilde{S}_A^*(\bar{T}-1) + \tilde{\gamma}^*[1-\tilde{S}_A^*(\bar{T}-1)]\}$. The second equality follows from $\tilde{Y}_A = \mathbb{E}[\tilde{Y}^*(0) \mid \tilde{A}^*=1, \tilde{T}_M^* > \tilde{T}_A^*]$ and the expression of $\mathbb{E}[\tilde{Y}^*(0) \mid \tilde{A}^*=0, \tilde{M}^*=1]$. The third inequality follows from the equalities of events: $\{\tilde{A}^*=1, \tilde{T}_M^* > \tilde{T}_A^*\} = \{A^*=1, T_M^* > T_A^*, R_A^*=1\}$ and $\{\tilde{A}^*=1, \tilde{T}_M^* < \tilde{T}_A^*\} = \{A^*=1, T_M^* < T_A^*, R_A^*=1\}$ so that

$$\mathbb{E}[\tilde{Y}^*(0) \mid \tilde{A}^*=1, \tilde{T}_M^* < \tilde{T}_A^*] - \mathbb{E}[\tilde{Y}^*(0) \mid \tilde{A}^*=1, \tilde{T}_M^* > T_A^*] = 0, \qquad (56)$$

by Assumptions 2 and 7; and $\{\tilde{A}^*=0, \tilde{M}^*=1\} = \{A^*=0, M^*=1\} \cup \{A^*=1, M^*=1, T_M^* < T_A^*, R_A^*=0\}$ so that

$$\mathbb{E}[\tilde{Y}^*(0) \mid \tilde{A}^*=0, \tilde{M}^*=1] \geq \mathbb{E}[Y^*(0) \mid A^*=0, M^*=1], \qquad (57)$$

by Assumptions 5 and 7. The last equality follows from definition of $\mu(0)$.

The conclusion of the theorem then follows from

$$\frac{[1-\tilde{\gamma}^*]\tilde{\Delta}_{ols} + \tilde{\gamma}^*\tilde{S}_M^*(\bar{T}-1)\tilde{\Delta}_{iv}}{1-\tilde{\gamma}^* + \tilde{\gamma}^*\tilde{S}_M^*(\bar{T}-1)} = \tilde{\mu}(1) - \tilde{\mu}(0) \leq \mu(1) - \mu(0) = \Delta. \qquad (58)$$

□

The effect of underreporting abortions on the overall ATE is difficult to sign, unless there is no heterogeneity in the conditional ATEs, i.e., $\Delta(X) = \Delta$ a.s. In this case underreporting introduces negative bias in the estimator of the overall ATE.

B.2 Reporting abortions as miscarriages

We show that misreporting abortions as miscarriages affects the identification of the effects of interest, even if the misreporting is random conditional on characteristics. Let R^*_{AM} denote the indicator for abortion reported as miscarriage. We assume that misreporting is independent of potential outcomes and scheduled weeks conditional on the characteristics X.

Assumption 8 (Conditionally random misreporting of abortion as miscarriage)

$$R^*_{AM} \perp\!\!\!\perp [Y^*(0), Y^*(1), T^*_M, T^*_A] \mid X \quad a.s.$$

We denote again with tilde all the probabilities and survival functions in the population with misreporting. In this case the analysis is slightly more complicated because misreporting produces misclassification of types.

The following result shows that the weighted average of the OLS and IV estimands in Theorem 2 underestimates the conditional ATE under Assumption 5.

Theorem 5 (Conditional ATE with misreporting of abortions as miscarriages) *Under Assumptions 1–5, and 8, if* $F^*_A = f \in \{0,1\}$ *and* $\mathbb{P}(R^*_{AM} = 1 \mid X) < 1$ *a.s.,*

$$\frac{[1-\tilde{\gamma}^*_f(X)]\tilde{\Delta}_{ols}(X) + \tilde{\gamma}^*_f(X)\tilde{S}^*_{fM}(X,\bar{T}-1)\tilde{\Delta}_{iv}(X)}{1-\tilde{\gamma}^*_f(X) + \tilde{\gamma}^*_f(X)\tilde{S}^*_{fM}(X,\bar{T}-1)} \leq \Delta_f(X), \quad a.s. \tag{59}$$

Proof To simplify the notation, we drop the dependence on X and f from all the functions and assume that there are no ties, $\mathbb{P}(T^*_M = T^*_A \mid X) = 0$ a.s.

Note that $\tilde{\mu}(1) = \mu(1)$ because misreporting does not produce misclassification of treatment, $\tilde{D} = D$, so that it does not affect women of birth type, $A^* = M^* = 0$. However, misreporting might affect $\tilde{\mu}(0)$ because it alters some of the relationships between the observed and structural variables in Assumption 1, and introduces misclassification of types.

By the same argument as in the proof of Theorem 4

$$\begin{aligned}\tilde{\mu}(0) \;&=\; \frac{\tilde{\gamma}^*[1-\tilde{S}^*_A(\bar{T}-1)]}{\tilde{S}^*_A(\bar{T}-1)}\{\mathbb{E}[\tilde{Y}^*(0) \mid \tilde{A}^* = 1, \tilde{T}^*_M < \tilde{T}^*_A] - \mathbb{E}[\tilde{Y}^*(0) \mid \tilde{A}^* = 1, \tilde{T}^*_M > T^*_A]\} \\ +\mathbb{E}[\tilde{Y}^*(0) \;&\mid\; \tilde{A}^* = 0, \tilde{M}^* = 1] \geq \mathbb{E}[Y^*(0) \mid A^* = 0, M^* = 1] = \mu(0),\end{aligned} \tag{60}$$

where the second inequality follows from the equalities of events: $\{\tilde{A}^* = 1, \tilde{T}_M^* > \tilde{T}_A^*\} = \{A^* = 1, T_M^* > T_A^*, R_{AM}^* = 0\}$ and $\{\tilde{A}^* = 1, \tilde{T}_M^* < \tilde{T}_A^*\} = \{A^* = 1, T_M^* < T_A^*, R_{AM}^* = 0\}$ so that

$$\mathbb{E}[\tilde{Y}^*(0) \mid \tilde{A}^* = 1, \tilde{T}_M^* < \tilde{T}_A^*] - \mathbb{E}[\tilde{Y}^*(0) \mid \tilde{A}^* = 1, \tilde{T}_M^* > T_A^*] = 0, \tag{61}$$

by Assumptions 2 and 8; and $\{\tilde{A}^* = 0, \tilde{M}^* = 1\} = \{A^* = 0, M^* = 1\} \cup \{A^* = 1, R_{AM}^* = 1\}$ so that

$$\mathbb{E}[\tilde{Y}^*(0) \mid \tilde{A}^* = 0, \tilde{M}^* = 1] \geq \mathbb{E}[Y^*(0) \mid A^* = 0, M^* = 1], \tag{62}$$

by Assumptions 5 and 8. The last equality follows from definition of $\mu(0)$.

The conclusion of the theorem then follows from

$$\frac{[1 - \tilde{\gamma}^*]\tilde{\Delta}_{ols} + \tilde{\gamma}^* \tilde{S}_M^*(\bar{T} - 1)\tilde{\Delta}_{iv}}{1 - \tilde{\gamma}^* + \tilde{\gamma}^* \tilde{S}_M^*(\bar{T} - 1)} = \tilde{\mu}(1) - \tilde{\mu}(0) \leq \mu(1) - \mu(0) = \Delta. \tag{63}$$

□

The effect of misreporting on the overall ATE is difficult to sign, unless there is no heterogeneity in the conditional ATEs, i.e., $\Delta(X) = \Delta$ a.s. In this case misreporting introduces negative bias in the estimator of the overall ATE.

B.3 Underreporting of miscarriages

We show that underreporting miscarriages does not affect the identification of the conditional ATEs, provided that underreporting is random conditional on characteristics. Let R_M^* denote the indicator for reported miscarriage. We assume that reporting is independent of potential outcomes and scheduled weeks conditional on the characteristics X.

Assumption 9 (Conditionally random miscarriage underreporting)

$$R_M^* \perp\!\!\!\perp [Y^*(0), Y^*(1), T_M^*, T_A^*] \mid X \quad a.s.$$

We denote with tilde all the variables, probabilities and survival functions in the population with underreporting; for example, $\tilde{p}_M(x, t) := \mathbb{P}(\tilde{T}_M = t \mid \tilde{X} = x) = \mathbb{P}(T_M = t \mid X = x, R_M^* = 1)$. We assume that women that do not report miscarriages are not observed in the population of pregnant teenagers. An implication of this assumption is that there is no misclassification of types and treatment in the population with underreporting.

The following result shows that the conditional ATE is identified by the weighted average of the OLS and IV estimands of Theorem 2 with underreporting of miscarriages.

Theorem 6 (Conditional ATE with underreporting of miscarriages) *Under Assumptions 1–4, and 9, if* $F_A^* = f \in \{0,1\}$ *and* $\mathbb{P}(R_M^* = 1 \mid X) > 0$ *a.s.,*

$$\frac{[1-\tilde{\gamma}_f^*(X)]\tilde{\Delta}_{ols}(X) + \tilde{\gamma}_f^*(X)\tilde{S}_{fM}^*(X,\bar{T}-1)\tilde{\Delta}_{iv}(X)}{1-\tilde{\gamma}_f^*(X) + \tilde{\gamma}_f^*(X)\tilde{S}_{fM}^*(X,\bar{T}-1)} = \Delta_f(X), \ \ a.s. \tag{64}$$

Proof To simplify the notation, we drop the dependence on X and f from all the functions and assume that there are no ties, $\mathbb{P}(T_M^* = T_A^* \mid X) = 0$ a.s.

Note that $\tilde{\mu}(1) = \mu(1)$ because underreporting does not affect women of birth type, $A^* = M^* = 0$. However, underreporting might affect $\tilde{\mu}(0)$ because it alters some of the relationships between the observed and structural variables in Assumption 1.

By the same argument as in the proof of Theorem 4

$$\begin{aligned}\tilde{\mu}(0) \ &= \ \frac{\tilde{\gamma}^*[1-\tilde{S}_A^*(\bar{T}-1)]}{\tilde{S}_A^*(\bar{T}-1)}\{\mathbb{E}[\tilde{Y}^*(0) \mid \tilde{A}^*=1, \tilde{T}_M^* < \tilde{T}_A^*] - \mathbb{E}[\tilde{Y}^*(0) \mid \tilde{A}^*=1, \tilde{T}_M^* > T_A^*]\} \\ +\mathbb{E}[\tilde{Y}^*(0) \ &\mid \ \tilde{A}^*=0, \tilde{M}^*=1] = \mathbb{E}[Y^*(0) \mid A^*=0, M^*=1, R_M^*=1] = \mu(0),\end{aligned} \tag{65}$$

where the second equality follows from the equalities of events: $\{\tilde{A}^* = 1, \tilde{T}_M^* > \tilde{T}_A^*\} = \{A^* = 1, T_M^* > T_A^*\}$ and $\{\tilde{A}^* = 1, \tilde{T}_M^* < \tilde{T}_A^*\} = \{A^* = 1, T_M^* < T_A^*, R_M^* = 1\}$ so that

$$\mathbb{E}[\tilde{Y}^*(0) \mid \tilde{A}^*=1, \tilde{T}_M^* < \tilde{T}_A^*] - \mathbb{E}[\tilde{Y}^*(0) \mid \tilde{A}^*=1, \tilde{T}_M^* > T_A^*] = 0, \tag{66}$$

by Assumption 9; and $\{\tilde{A}^* = 0, \tilde{M}^* = 1\} = \{A^* = 0, M^* = 1, R_M^* = 1\}$ so that

$$\mathbb{E}[\tilde{Y}^*(0) \mid \tilde{A}^*=0, \tilde{M}^*=1] = \mathbb{E}[Y^*(0) \mid A^*=0, M^*=1, R_M^*=1]. \tag{67}$$

The last equality follows from Assumption 9.

The conclusion of the theorem then follows from

$$\frac{[1-\tilde{\gamma}^*]\tilde{\Delta}_{ols} + \tilde{\gamma}^*\tilde{S}_M^*(\bar{T}-1)\tilde{\Delta}_{iv}}{1-\tilde{\gamma}^* + \tilde{\gamma}^*\tilde{S}_M^*(\bar{T}-1)} = \tilde{\mu}(1) - \tilde{\mu}(0) = \mu(1) - \mu(0) = \Delta. \tag{68}$$

□

The effect of underreporting miscarriages on the overall ATE is difficult to sign, unless there is no heterogeneity in the conditional ATEs, i.e., $\Delta(X) = \Delta$ a.s. In this case underreporting does not affect the probability limit of the estimator of the overall ATE.

C Multiple Teen Pregnancies

Let D_2 be an indicator for women who gave birth from a second teen pregnancy. We assume that this treatment indicator is independent of potential outcomes and miscarriage type in the first pregnancy for women who did not give birth in the first pregnancy. The treatment that defines potential outcomes is giving birth in either the first or second teen pregnancy, i.e. $\max(D, D_2)$.

Assumption 10 (Conditionally random birth from second pregnancy)

$$D_2 \perp\!\!\!\perp [Y^*(1), Y^*(0), M^*] \mid X, A^*, D = 0.$$

This assumption imposes that, for women with characteristics X who did not give birth in the first pregnancy, the event of giving birth in a second pregnancy is independent of potential outcomes and miscarriage type. It permits, however, the probability of giving birth in a second pregnancy to be different for abortion and non-abortion types. Let $\pi_f^*(X, A^*) = \mathbb{E}[D_2 \mid X, A^*, D = 0]$ be the probability of giving birth from a second teen pregnancy for women who did not give birth in the first pregnancy, with characteristics X, and of abortion type A^*, when $F_A^* = f$.[21] The following Lemma adapts Lemma 2 to multiple teen pregnancies. We focus on the case where Assumption 4 holds. A similar analysis applies without imposing this assumption, but the notation is more cumbersome.

Lemma 3 (Moment conditions) *Under Assumptions 1–4, and 10, if* $F_A^* = f \in \{0, 1\}$,

$$\mathbb{E}[D_2 \mid M = 1, X][1 - S_M(X, \bar{T} - 1)][1 - S_{fM}^*(X, \bar{T} - 1)]^{-1}$$
$$= S_{fA}^*(X, \bar{T} - 1)\pi_f^*(X, 0) + \gamma_f^*(X)[1 - S_{fA}^*(X, \bar{T} - 1)]\pi_f^*(X, 1) \quad (69)$$
$$\mathbb{E}[D_2 \mid A = 1, X] = \pi_f^*(X, 1), \quad (70)$$
$$\mathbb{E}[Y \mid M = 1, X][1 - S_M(X, \bar{T} - 1)][1 - S_{fM}^*(X, \bar{T} - 1)]^{-1}$$
$$= S_{fA}^*(X, \bar{T} - 1)\{\pi_f^*(X, 0)\mathbb{E}[Y^*(1) \mid A^* = 0, X] + [1 - \pi_f^*(X, 0)]\mathbb{E}[Y^*(0) \mid A^* = 0, X]\}$$
$$+\gamma_f^*(X)[1 - S_{fA}^*(X, \bar{T} - 1)]\{\pi_f^*(X, 1)\mathbb{E}[Y^*(1) \mid A^* = 1, X]$$
$$+[1 - \pi_f^*(X, 1)]\mathbb{E}[Y^*(0) \mid A^* = 1, X]\}, \quad (71)$$
$$\mathbb{E}[Y \mid A = 1, X] = \pi_f^*(X, 1)\mathbb{E}[Y^*(1) \mid A^* = 1, X] + [1 - \pi_f^*(X, 1)]\mathbb{E}[Y^*(0) \mid A^* = 1, X], \ a.s. \quad (72)$$

Proof The proof is omitted because it follows from similar arguments to the proofs of Lemma 2 and Theorem 2. □

[21] We do not impose the restriction that the probability of giving birth in the second pregnancy for abortion types is zero, $\pi_f^*(X, 1) = 0$, because in our data we observe that about an eight of teens who terminated their first pregnancy, gave birth in a second teen pregnancy.

We show that we can rescale the weighted average of the OLS and IV estimands in Theorem 2 to identify the conditional ATE. The rescaling factor is greater than one and adjusts the weighted average for the possibility of giving birth in the second pregnancy. It depends only on the probability of giving birth in the second pregnancy for miscarriage types and is identified from the data and the competing risk model.

Theorem 7 (Conditional ATE with multiple pregnancies) *Under Assumptions 1–4, and 10, if* $F_A^* = f \in \{0,1\}$ *and* $\pi_f^*(X,0) < 1$ *a.s.,*

$$\Delta_f(X) = \frac{1}{1-\pi_f^*(X,0)} \frac{[1-\gamma_f^*(X)]\Delta_{ols}(X) + \gamma_f^*(X) S_{fM}^*(X,\bar{T}-1)\Delta_{iv}(X)}{1-\gamma_f^*(X) + \gamma_f^*(X) S_{fM}^*(X,\bar{T}-1)}, \ \text{a.s.} \tag{73}$$

where

$$\pi_f^*(X,0) = \frac{[1-S_M(X,\bar{T}-1)]\mathbb{E}[D_2 \mid M=1,X]}{S_{fA}^*(X,\bar{T}-1)[1-S_{fM}^*(X,\bar{T}-1)]} - \frac{\gamma_f^*(X)[1-S_{fA}^*(X,\bar{T}-1)]\mathbb{E}[D_2 \mid A=1,X]}{S_{fA}^*(X,\bar{T}-1)}, \ \text{a.s.} \tag{74}$$

Proof We use the same notation as in the proof of Theorem 2.

Note that the possibility of multiple teen pregnancies does not affect the identification of $\mu(1)$, so that $\mu(1) = Y_D$. For $\mu(0) = \mathbb{E}[Y^*(0) \mid A^* = 0]$, by Lemma 3

$$\mu(0) = \frac{[1-S_M(\bar{T}-1)]Y_M - [1-S_M^*(\bar{T}-1)]S_A^*(\bar{T}-1)\pi^*(0)Y_D - \gamma^*[1-S_M^*(\bar{T}-1)][1-S_A^*(\bar{T}-1)]Y_A}{[1-S_M^*(\bar{T}-1)]S_A^*(\bar{T}-1)[1-\pi^*(0)]}. \tag{75}$$

Substituting the expressions of $\mu(1)$ and $\mu(0)$,

$$\Delta = \mu(1) - \mu(0) = \frac{1}{1-\pi^*(0)} \left\{ Y_D - Y_M + \frac{\gamma^*[1-S_A^*(\bar{T}-1)]}{S_A^*(\bar{T}-1)]} [Y_A - Y_M] \right\}. \tag{76}$$

The result then follows because by the same argument as in Theorem 2

$$\frac{[1-\gamma^*]\Delta_{ols} + \gamma^* S_M^*(\bar{T}-1)\Delta_{iv}}{1-\gamma^* + \gamma^* S_M^*(\bar{T}-1)} = Y_D - Y_M + \frac{\gamma^*[1-S_A^*(\bar{T}-1)]}{S_A^*(\bar{T}-1)}[Y_A - Y_M]. \tag{77}$$

The expression for $\pi^*(0)$ follows from solving the first two equations of Lemma 3. □

References

Ashcraft, A., Fernandez-Val, I., and K. Lang (2012), "The Consequences of Teenage Childbearing: Consistent Estimates When Abortion Makes Miscarriage Nonrandom," *Economic Journal*.

Kalbfleisch, J. D., and R. L. Prentice (2002), *The Statistical Analysis of Failure Time Data, 2nd ed.* Hoboken, New Jersey: John Wiley.