# The Educational Consequences of Teen Childbearing

**Jennifer B. Kane**,
Carolina Population Center, University of North Carolina at Chapel Hill, 123 W. Franklin Street, Chapel Hill, NC 27516

**S. Philip Morgan**,
Department of Sociology and Carolina Population Center, University of North Carolina at Chapel Hill

**Kathleen Mullan Harris**, and
Department of Sociology and Carolina Population Center, University of North Carolina at Chapel Hill

**David K. Guilkey**
Department of Economics and Carolina Population Center, University of North Carolina at Chapel Hill

Jennifer B. Kane: jbkane@unc.edu

## Abstract

A huge literature shows that teen mothers face a variety of detriments across the life course, including truncated educational attainment. To what extent is this association causal? The estimated effects of teen motherhood on schooling vary widely, ranging from no discernible difference to 2.6 fewer years among teen mothers. The magnitude of educational consequences is therefore uncertain, despite voluminous policy and prevention efforts that rest on the assumption of a negative and presumably causal effect. This study adjudicates between two potential sources of inconsistency in the literature—methodological differences or cohort differences—by using a single, high-quality data source: namely, The National Longitudinal Study of Adolescent Health. We replicate analyses across four different statistical strategies: ordinary least squares regression; propensity score matching; and parametric and semiparametric maximum likelihood estimation. Results demonstrate educational consequences of teen childbearing, with estimated effects between 0.7 and 1.9 fewer years of schooling among teen mothers. We select our preferred estimate (0.7), derived from semiparametric maximum likelihood estimation, on the basis of weighing the strengths and limitations of each approach. Based on the range of estimated effects observed in our study, we speculate that variable statistical methods are the likely source of inconsistency in the past. We conclude by discussing implications for future research and policy, and recommend that future studies employ a similar multimethod approach to evaluate findings.

### Keywords

Teen childbearing; Educational attainment; Semiparametric maximum likelihood; Propensity score matching; Add Health

## Introduction

After 40 years of research on educational consequences of teen childbearing, the magnitude of the causal relationship is still unclear. Some scholars have found no discernible difference

in the educational attainment of teen mothers compared with others (Fletcher and Wolfe 2009; Hotz et al. 2005, 2008; Jones et al. 1999; Olsen and Farkas 1989; Ribar 1994; Rindfuss et al. 1980). When studies have found differences, the estimates range from 0.15 to 2.6 fewer years of schooling among teen mothers (Ashcraft et al. 2013; Ashcraft and Lang 2006; Card and Wise 1978; Furstenberg 1976; Furstenberg Jr. 2003; Furstenberg et al. 1989; Grogger and Bronars 1993; Hofferth and Moore 1979; Hoffman 2008; Klepinger et al. 1995, 1999; Lee 2010; Levine and Painter 2003; Marini 1984; McElroy 1996; Mollborn 2007; Moore and Hofferth 1980; Moore and Waite 1977; Mott and Marsiglio 1985; Rich and Kim 1999; Sanders et al. 2007; Upchurch and McCarthy 1990; Waite and Moore 1978; Wellings 2007). This range encompasses a differential close to zero and values (2+ years) reflecting substantively important distinctions—such as completing high school or a work-based training program—that are logical targets for policy efforts.

Policy makers generally operate under the assumption that teen childbearing has negative consequences for women and, as a result, organize efforts to prevent teen childbearing. As noted earlier, this assumption may not be true in the important case of educational attainment. Because policy makers are continually searching for ways to strategically allocate scarce resources, evaluating the impact of teen childbearing on education is of great practical importance.

The range of findings in this literature could be produced by variable statistical methods or populations studied. With respect to method, common approaches used in the past include instrumental variables, propensity score matching, sibling fixed effects, and structural equation modeling. Such strategies are necessary because it is generally agreed that endogeneity in teen childbearing exists; that is, common unobservable factors are present in the error terms of the regression equations that predict both teen childbearing and educational attainment. Endogeneity is problematic in that selection into "treatment" (i.e., teen childbearing) and "control" (i.e., not teen childbearing) groups is not random, producing incorrect estimates of treatment effects (i.e., the effect of teen childbearing on educational attainment) using non-experimental strategies, such as ordinary least squares (OLS) regression. Different estimates across methods may arise because each addresses endogeneity differently and imposes different assumptions. A second possibility is that inconsistency arises from estimating the causal relationship across different data sets, some of which examine women born in the 1940, 1950s, and 1960s; some examine women born in the 1980s.

We adjudicate between these two sources of inconsistency by replicating analyses across four different statistical strategies using the same population-based data set. If our findings are consistent and robust across methodologies, this would suggest that cohort differences may underlie past observed inconsistencies. On the other hand, if results replicate the range of estimates observed in past research, then we will argue that past inconsistency is likely due to choice of statistical method.

In this second scenario, a range of estimates begs for a "preferred estimate" produced by weighing the strengths and limitations of each approach. If the preferred model estimate is "no effect," this would indicate an underlying selection mechanism (i.e., socioeconomic disadvantage) is the root cause of both teen childbearing and diminished educational attainment. A second possibility is a salient effect that is diminished in magnitude relative to non-experimental methods (i.e., OLS). This would lend support to both causation and selection arguments: an underlying selection mechanism contributing to both factors is present, but an important part of the relationship remains that can be attributed to a causal effect of teen childbearing. In a third scenario, a salient but inflated effect (relative to OLS) would emerge, potentially suggesting some version of favorable self-selection. That is,

young women may purposefully choose to have a teen birth but also benefit from high levels of social or structural support that allow her to achieve her desired level of education. We discuss this later.

In the following sections, we first review past research. We organize this review around choice of statistical method to highlight how assumptions, strengths, and limitations of each method affect estimates of the educational penalty. We also introduce each method with a brief description of the estimation procedure. We then introduce two additional strategies not used in this literature—namely, parametric and nonparametric maximum likelihood (Rindfuss et al. 2007, 2010), which provide (1) an explicit estimate of the distribution of unobserved heterogeneity affecting both teen childbearing and educational attainment, and (2) a direct test of the endogeneity of teen childbearing with educational attainment. Next we conduct analyses that document a range of estimates (between 0.7 and 1.9 fewer years of schooling among teen mothers) and select our preferred estimate (0.7), derived from semiparametric maximum likelihood estimation, on the basis of weighing the strengths and limitations of each approach. Based on the range observed in our study, we conclude that variable statistical methods are the likely source of inconsistency in past research. We also discuss implications for future research and policy.

## Background

### Statistical Approaches in Past Research

Most research employs one of three distinct strategies—propensity score matching (PSM), within-family fixed effects (FE), or instrumental variables (IV)—to assess the educational consequences of teen childbearing. To make the differences in the methods more explicit, we write the statistical specification of the model as follows:

$$TCB_i^* = \mathbf{X}_{i1}\beta + \varepsilon_{i1} \quad (1)$$

$$E_i = \alpha\, TCB_i + \mathbf{X}_{i2}\delta + \varepsilon_{i2}, \quad (2)$$

where the dependent variable in Eq. (1) is a latent variable related to teen childbearing. The corresponding observed variable is $TCB_i$, which is a binary variable indicating whether respondent $i(i = 1, 2, …, N)$ had a teen birth that takes on a value of 1 when $TCB_i^*$ is positive and 0 otherwise. $\mathbf{X}_{i1}$ represents observed exogenous variables that affect teen childbearing, $\beta$ is an unobserved set of coefficients to be estimated, and $\varepsilon_{i1}$ represents unobserved variables that affect teen childbearing.

The dependent variable in Eq. (2) is educational attainment for respondent $i$ and is assumed to be a function of whether the respondent had a teen birth, a set of exogenous variables, and an unobserved error term. The set of exogenous variables in Eqs. (1) and (2) ($\mathbf{X}_{i1}$ and $\mathbf{X}_{i2}$) have a great deal of overlap (i.e., race/ethnicity, parental socioeconomic status). However, there are variables included in $X_{i1}$ that are omitted from Eq. (2) (i.e., presence of abortion clinics, neighborhood teen childbearing rates). In fact, some of the estimation methods discussed later require that at least one variable is included in $\mathbf{X}_{i1}$ that is not included in $\mathbf{X}_{i2}$ in order for statistical identification of Eq. (2). We discuss this later.

Correct estimation of $\alpha$ in Eq. (2), the parameter of primary interest, depends on assumptions made about the error terms. Some methods used in the past assume that there is no overlap in the set of unobservables affecting Eqs. (1) and (2) (i.e., $E(\varepsilon_{i1}\varepsilon_{i2}) = 0$). However, we previously argued that the error terms are likely to be correlated (i.e., $E(\varepsilon_{i1}\,\varepsilon_{i2})$ $\varepsilon$ 0), which means that these methods could yield misleading results.

## Statistical Approaches Assuming No Correlation Between Error Terms: OLS and PSM

Early studies used descriptive and OLS regression techniques to show that teen mothers complete high school at lower rates than their classmates from similar socioeconomic backgrounds (Furstenberg 1976) and complete 1–4 fewer years of schooling (Moore and Waite 1977; Waite and Moore 1978). OLS regression is analogous to specifying this association using only Eq. (2). Lower rates of high school completion and college enrollment among teen mothers were also observed using basic matching strategies (Card and Wise 1978). Despite the foundational nature of these studies, findings may be misleading given that they do not directly address the potential endogeneity of teen childbearing.

A second approach, PSM, indirectly addresses potential endogeneity by comparing educational outcomes across two subgroups that exhibit similar teen childbearing propensities. This estimation requires four steps: (1) regressing the risk of teen childbearing on observed covariates that may approximate the selection mechanism, (2) using predicted values of teen birth risk to identify matches between "treatment" and "control" cases, (3) comparing matched cases to ensure that groups are "balanced" on observed covariates, and (4) calculating outcome differences between matched cases to produce average treatment effects. Generally, studies using this method find small adverse effects of teen childbearing on educational attainment (Levine and Painter 2003; Sanders et al. 2007). One particularly relevant study using PSM and OLS regression with Add Health data (Waves I, II, and III) found diminished consequences of teen childbearing (relative to OLS) (Lee 2010).

While powerful, PSM is limited by (1) not explicitly modeling the endogeneity of teen childbearing and (2) including only observed factors that may be part of the selection mechanism. This approach rests on the "ignorable treatment assignment" assumption: namely, that assignment into control and treatment groups is random, meaning that all pertinent factors are observed or, at the very least, that relevant unobserved characteristics are correlated with observables (Rosenbaum and Rubin 1983). Because this form of matching may not eliminate selection bias (Heckman et al. 1996), potential violation must be explored within sensitivity analyses (Rosenbaum 2002; Rosenbaum and Rubin 1985). The "common support condition" calls attention to a third limitation—that average treatment effects are calculated based on the matched sample only. This condition omits cases in which the estimated propensity score is so small or so large that a sufficient match cannot be identified, and is particularly restrictive if the effect of the treatment depends on the propensity score (Heckman et al. 1996). No strategy is without its limitations, however, which further emphasizes the importance of comparing findings across multiple methodologies.

## Statistical Approaches That "Difference-Out" Common Unobservables: Fixed Effects

A third strategy uses within-family fixed-effects (FE) models. This approach compares educational outcomes across sister pairs, one of whom became a teen mother (the treatment) and one of whom did not (the control), thus correcting for family-specific characteristics that are identical for both sisters. This includes factors—observed or unobserved—in the family environment that may contribute to the selection mechanism such as socioeconomic status (SES), neighborhood environment, parenting styles, parent's personality characteristics, or genetic traits.

Studies using FE generally document diminished effects of teen childbearing compared with non-experimental approaches. For example, one study examined sister pairs in three surveys (National Longitudinal Survey of Young Women, Panel Study of Income Dynamics, and National Longitudinal Survey of Youth 1979) and found that detriments related to high

school graduation and postsecondary schooling rates were attenuated, although they did not disappear entirely (Geronimus and Korenman 1992). Another study using data from sisters in the PSID reached essentially the same conclusions: teen mothers exhibited lower rates of high school graduation and college attendance, but differences were smaller within FE relative to basic regression (Hoffman et al. 1993).

This strategy is limited in four ways. First, its estimation can be inefficient because sister pairs are retained only if they differ on both the outcome (education) and explanatory variable (teen birth). This produces small sample sizes—particularly when examining binary outcomes, such as high school graduation—which limit statistical power to detect group differences (see Hoffman et al. 1993). Second, it cannot, by definition, account for unobserved family experiences that differ across siblings, which is important because within-family heterogeneity, such as differential academic performance across sisters, can bias estimates of the consequences of teen childbearing (Holmlund 2005). Third, FE results are not generalizable beyond the analytic sample (Allison 2009). Fourth, as an approximation of a classic experimental design, this design suffers from a contamination effect in which the "treatment" potentially results in family resources being diverted from the sibling without a teen birth to the sibling with a teen birth (Burton 1990). In sum, the extent to which these findings can inform our understanding of population-level trends is limited.

### Statistical Approaches Assuming the Error Terms Are Correlated: IV, Parametric, and Semiparametric Maximum Likelihood

A fourth strategy uses instrumental variables (IV). This approach directly addresses the endogeneity of teen childbearing by estimating two separate equations. The first equation uses linear regression to regress the endogenous variable (teen childbearing) on sociodemographic controls (or exogenous covariates) and variables that "identify" teen childbearing and serve as "instruments." In this context, identifying or instrumental variables are related to the endogenous variable but are not otherwise directly related to the outcome. Then, predicted values of teen childbearing are used in a second equation to estimate educational attainment. This approach successfully breaks the correlation of teen childbearing with the error term of the second equation, which produces an unbiased estimate of the treatment effect.[1]

Many IV studies use miscarriage as an identifying variable and limit the sample to young women who become pregnant as teenagers. This approach compares outcomes across two groups of women, both of whom would have become teen mothers but differ only on the random event of a miscarriage (Hoffman 2008; Hotz et al. 2005, 2008).[2] Some studies using this approach have found little evidence of any educational penalty of early childbearing (e.g., Fletcher and Wolfe 2009), although others have shown small penalties: 0.15 fewer years of schooling for teen mothers (Ashcraft et al. 2013; Ashcraft and Lang 2006). These latter studies argued that this is an underestimate because teens who miscarry are more likely to come from disadvantaged backgrounds than teens who abort or give birth. Rindfuss and colleagues (1980) also used teen miscarriage to identify age at first birth, although their sample included all women at risk of birth (not just pregnant teens). They found a strong recursive effect—that educational success reduces teen birth risk—and concluded that these decisions are made jointly, an important point to which we will return later.

---

[1]Two-step probit estimation is sometimes used instead to accommodate the binary nature of teen childbearing.
[2]Whether miscarriage is a random event is the subject of some debate (see Ashcraft et al. 2013; Fletcher and Wolfe 2009).

One limitation of these IV studies is that, by definition, treatment effects are estimated conditional on becoming pregnant as a teenager (Rindfuss et al. (1980) excepted). Because we are interested in estimating educational consequences of teen childbearing at the national level, we therefore use a full population-based sample. Although this necessitates a different comparison, it should not be problematic provided that we use strategies that account for both observed and unobserved factors that are part of the underlying selection process.

Another shortcoming of these IV studies is that they use only one identifying variable. Other studies use multiple instruments in a single model, such as age at menarche, county-level physician availability, county-level abortion rates, and area-level economic well-being. We draw on this work to select multiple instruments for use in our analyses. Each factor is a plausible instrument for teen birth because each is correlated with teen childbearing and arguably has no direct effect on educational attainment. For example, local abortion rates may reflect the availability of abortion clinics and whether the medical option to terminate a pregnancy is available locally. Similar indicators include state-level laws regulating public funding for abortion clinics and whether parental consent is required for abortions among minors. Abortion laws are linked with teen childbearing rates (Fletcher and Wolfe 2009), but no research has linked these (directly) to young women's educational attainment. In the same way, the availability of obstetricians in young women's neighborhoods contextualizes the health care environment in which teens may seek prenatal care or prevention services, but it does not directly impact young women's educational attainment. Ribar (1994) argued that the number of neighborhood physicians affects both the cost of prenatal care and prevention services but shows an overall negative association with teen births, suggesting that the cost of prevention for teenagers is more relevant. In addition, higher average payments to low-income families made through programs such as Medicaid have been found to be positively associated with (Ribar 1994), or unrelated to (An et al. 1993), teen births. Medicaid payments may also proxy for more generous social safety net programs and have a negative association with teen births.

Studies using these variables as instruments produce varying estimates of educational consequences. Klepinger et al. (1999) cited 2.6 fewer years of schooling among teen mothers; Marini (1984) estimated that women complete 0.16 additional years of education for each year of delayed childbearing. Not all studies cite adverse effects, however. Ribar (1994) found no difference in schooling between teen and nonteen mothers, and Olsen and Farkas (1989) found no significant effect on high school dropout rates.

Note that all studies using IV necessarily assume that the treatment effect is homogeneous across the sample. In practice, this assumption may be violated. For example, some women may elect to have a teen birth because they have low educational aspirations relative to other women in the population, as some research suggests (Brien et al. 1999; Musick 1995; Rindfuss et al. 1980; Upchurch et al. 2002b; Upchurch and McCarthy 1990). In addition, some studies have cited IV estimates of educational consequences that are larger in magnitude than OLS estimates (Ashcraft et al. 2013; Ashcraft and Lang 2006; Klepinger et al. 1999). This situation can arise when the effect of teen childbearing on educational attainment varies across subgroups in the sample (Angrist and Evans 1996), which is also referred to as "essential heterogeneity" (Basu et al. 2007; Heckman et al. 2006). Essential heterogeneity arises from two sources: when people respond differently to a treatment, and when people exhibit idiosyncratic gains from treatment (Heckman et al. 2006) for reasons that are correlated with knowledge about how the outcome would affect them. A test for solving the problem of essential heterogeneity was recently proposed by Basu et al. (2007), but little research has rigorously evaluated it. This is an important area for future research that is beyond the scope of this article.

Although IV methods are the most frequently used method to correct for endogeneity, a full information maximum likelihood (ML) approach has the advantage of being efficient relative to IV as long as model assumptions are satisfied. A fully parametric ML estimation strategy makes a specific distributional assumption about the joint error term distribution. The most common assumption is that $\varepsilon_{i1}$ and $\varepsilon_{i2}$ follow a bivariate normal distribution with mean zero and standard deviations $\sigma_1$ and $\sigma_2$, respectively, for all $N$ observations. (Because Eq. (1) is in probit form, the standard identifying assumption that $\sigma_1 = 1$ is imposed.) It is further assumed that the correlation between the two errors ($\rho$) is nonzero. With these assumptions, parametric ML is straightforward and can be implemented using the *treatreg* procedure in Stata. Along with the other parameters of the model, $\rho$ is estimated, providing a direct test of the endogeneity of teen birth. Because the model is nonlinear, it is identified without exclusion restrictions (i.e., $\mathbf{X}_{i1}$ and $\mathbf{X}_{i2}$ can be identical). However, parameter estimates tend to be more stable if valid exclusion restrictions are imposed.

The other full-information ML estimator that could be used is semiparametric. In this case, it is assumed that the error term can be written with two components:

$$\varepsilon_{i1} = \mu_{i1} + \upsilon_{i1} \quad (3)$$

$$\varepsilon_{i2} = \mu_{i2} + \upsilon_{i2}, \quad (4)$$

where the assumptions are that $E(\upsilon_{i1}\upsilon_{i2}) = 0$ and that these error terms follow the normal distribution. The assumptions for the other error components are (1) that they are correlated and (2) that the joint distribution can be approximated by discrete factors along the lines originally suggested by Heckman and Singer (1984) and evaluated in a Monte Carlo setting by Mroz (1999) and Guilkey and Lance (2012). Mroz found that in a two-equation model with one discrete and one continuous dependent variable, the discrete factor method worked almost as well as parametric ML when the true error distribution was bivariate normal and substantially better than parametric ML (that assumes normality) when the true error distribution is nonnormal. Guilkey and Lance documented similar results for a model in which both dependent variables are binary. The actual implementation of the discrete-factor method that Guilkey and Lance used is the logistic cumulative distribution for $\upsilon_{i1}$ rather than normality. However, this has been found to make little difference empirically. We use the logistic distribution in estimations that follow.

## Contributions of the Current Study

Our study advances understanding of the educational consequences of teen childbearing in four ways. First, we compare results across multiple statistical strategies, each of which takes a different approach to endogeneity. Past work compares only one quasi-experimental approach with OLS regression. Second, we make these comparisons using a single, rich data source: Add Health. This decision isolates the effect of different estimation strategies and allows us to speculate on whether estimates in past studies are related more to choice of statistical strategy or data set. No other study has taken this approach nor has been able to speculate on the source of inconsistency. Third, we draw conclusions from a nationally representative population-based sample that provides an estimate of the years of education lost as a result of having a teen birth for all those at risk of teen childbearing, which proves a useful platform to spur future policy discussions. Many prior studies condition the sample on women who became pregnant as teenagers, which indicates years of education lost *only* for those who become pregnant as teenagers. Fourth, very few have used data from contemporary cohorts. Our approach estimates the contemporary, population-level educational penalty, along with upper and lower bounds.

## Data, Measures, and Methods

### Data

We use data from Add Health (the National Longitudinal Study of Adolescent Health), a school-based, nationally representative sample of 20,745 seventh through twelfth graders in 1994–1995. Respondents were reinterviewed in 1996 (Wave II), 2001–2002 (Wave III), and 2008–2009 (Wave IV) (Harris 2010; Harris et al. 2009). Our sample includes only female respondents who participated in Waves I and IV ($n$ = 8,352). We examine women because the burden of teen pregnancy falls more heavily on mothers (than fathers) and because young men's reports of parenthood are unreliable (Upchurch et al. 2002a). We further constrain our sample to those with a valid sampling weight ($n$ = 7,870) to produce nationally representative estimates.[3] In most cases, missing data on analytic variables are minimal (less than 4 %).[4] We use a single imputation procedure in Stata to replace missing data on all independent variables. Results from listwise deletion (available upon request) produce substantively and statistically similar results.

### Measures

Respondent's schooling is measured at Wave IV. Offered responses were a series of ordinal categories (1 = 8th grade or less, 13 = completed post baccalaureate professional education), which we coded as a continuous variable measuring years of completed education (mean = 14.4; range = 8–26). At Wave IV, female respondents were aged 24–34 (mean = 28.7). We replicated analyses using only the older half of the sample (aged 28–34) and found substantively and statistically similar results.

Teen childbearing is a binary variable indicating whether respondents had a live birth before age 19 (1 = yes). Much prior work examines births prior to age 18. We replicate our models accordingly and find identical results.

Consistent with prior research (see Hoffman and Maynard 2008), we include sociodemographic covariates of teen birth and educational attainment, measured at Wave I, including family structure, parent's education, household income-to-needs ratio, race/ethnicity, nativity status, and age. We also include the respondent's age-standardized score on the Add Health Picture Vocabulary test (AHPVT), an indicator of cognitive ability.

We use additional variables in the parametric and semiparametric ML models to identify teen childbearing: statewide abortion laws regarding parental consent for, and public funding of, abortion; the abortion rate among women aged 15–44; the average Medicaid payment per recipient; the number of Ob/Gyn physicians per 100,000 women aged 15–44; and the percentage of family planning clients younger than age 20.[5] All variables except for the latter have been used as instruments in past research. Descriptive statistics and variable descriptions for all measures are shown in Table 1.

Results from a test of overidentifying restrictions indicate that these instruments are valid ($\chi^2$ = 4.57, $p$ = .47).[6] We do not employ teen miscarriage as an instrument for two reasons. First, it is not necessarily exogenous to the system when all young women at risk of birth are included in the sample (in contrast to a sample of pregnant teens). Second, recent evidence

---

[3]Respondents without sampling weights were recruited in the field to complete cells for the genetic oversample of sibling pairs, or attended a school outside the 80 school communities selected in the Add Health design.

[4]Variables with more than 4 % missingness include parent's education and income, and the respondents' grades in math, history, and science.

[5]We replicated analyses using an additional instrument—age at menarche—but found no differences.

[6]This test was performed using a linear IV regression model. It rests on the assumption that one instrument is valid and then goes on to test the validity of remaining instruments.

suggests that even in a sample limited to pregnant teens, miscarriage can be related to community-level effects (Fletcher and Wolfe 2009).

## Methods

First, we estimate OLS regressions using exploit four statistical strategies: one non-experimental regression method (OLS regression), and three quasi-experimental strategies (propensity score matching, parametric maximum likelihood (P-MLE), and semiparametric maximum likelihood (SP-MLE)). Sibling fixed-effects models were initially explored, but the small sample size called into question the robustness of the results. Analyses were performed in Stata with the exception of the SP-MLE models, which were estimated in Fortran-based software. We present unweighted models, although we replicated our analysis adjusting for survey weighting and stratification and found similar results (available upon request).the statistical specification outlined in Eq. (2). OLS does not account for endogeneity and therefore produces a biased estimate of the treatment effect of teen childbearing on education, provided that there are common unobservables across Eqs. (1) and (2). This is true even asymptotically (Heckman 1997).

Second, we use propensity score matching based on Eqs. (1) and (2). Bootstrapped standard errors are used that approximate average treatment effects on the treated group. Several matching strategies were explored; we present results based on a 1 % caliper radius because these are the most conservative. Only cases that lie within the region of common support are retained. Replications of analyses with and without replacing cases did not change results. We conduct a Rosenbaum bounds sensitivity analysis to examine the extent to which the outcome is homoskedastic within each group (treatment and control) or that no unobserved heterogeneity remains (Rosenbaum 2002). This assumption is satisfied when the model contains an exhaustive set of covariates that differ between groups, which proves to be a demanding and unrealistic endeavor. Rosenbaum bounds examine the extent to which failure to satisfy this assumption may bias results by estimating how large the effect of unobserved heterogeneity would have to be in order to nullify propensity score estimates. These bounds are also useful because the algorithms used in the propensity score estimation do not produce consistent estimators of the treatment effects if the treatment is endogenous (DiPrete and Gangl 2004).

Third, a treatment effects model (a parametric ML estimator (P-MLE)) is estimated, again based on Eqs. (1) and (2). We use the Stata *treatreg* command, which is based on full information maximum likelihood and assumes joint normality across error terms of both equations. This method also provides an estimate of the correlation between the error terms from both equations ($\rho$), which provides a direct test for the presence of common unobservables across Eqs. (1) and (2). Thus, $\rho \neq 0$ implies that OLS and PSM yield biased results.

Fourth, semiparametric ML estimator (SP-MLE), or discrete factor models, are estimated based on Eqs. (1)–(4). This procedure is similar to a treatment effects approach because it estimates a reduced-form equation (predicting teen birth) and a structural equation (predicting educational attainment based on whether the respondent had a teen birth along with control variables). However, this approach has three distinct advantages. First, it does not impose the multivariate normality assumption on the correlated unobservables; instead, it imposes a nonparametric error assumption. Because this method is less restrictive in this sense, the SP-MLE results can inform our understanding of whether the multivariate normality assumption is violated in the parametric model. For example, if similar results from the two strategies emerge, this suggests that the normality assumption is supported. However, if different results emerge, this might indicate a potential violation of this assumption. Second, it estimates the distribution of unobserved heterogeneity. This will be

used as a benchmark to determine whether many unobserved variables are left out of other models. This also provides more precise estimates of the treatment effect because it is explicitly estimated net of all unobserved factors. Third, it identifies a series of subgroups of women identified by their underlying probability to have a teen birth.

We present two models for each method: one including only teen childbearing (null model) and one that adds sociodemographic controls (full model). We use estimates from each full model to calculate the average treatment effect. We then compare these results (along with lower and upper bounds of the estimated treatment effect based on 95% confidence intervals (CI)) across all four strategies. Detailed results for each method are presented in the appendices and referenced throughout the text.

## Results

Weighted descriptive analyses, presented in Table 1, indicate that the analytic subsample is similar to other population-level samples of contemporary young women. Most important for our study, respondents completed an average of 14 years of education by Wave IV, when women were aged 24–34 (the average age was 29). Approximately 12 % of young women in our sample had a birth before age 19.

Table 2 presents a summary of results across the four methods. OLS results (detailed in the Online Resource 1, Table S1) show that teen mothers complete nearly 2 fewer years of schooling than nonteen mothers when no other covariates are included (Table S1, Model 1, $b = -1.75$). Adjusting for sociodemographic covariates reduces this difference to less than one year (Table S1, Model 2, $b = -0.94$). As expected, women living in two-biological-parent families, who have a higher household income-to-needs ratio, with higher AHPVT scores, and whose parents have higher levels of education completed more years of education. The point estimates for control variables are of the same sign and of similar magnitude across estimation methods. Therefore, we do not further discuss these effects.

The propensity score matching began with a (logistic) regression of teen birth risk on observed sociodemographics (see Online Resource 1, Table S2). Many significant differences in the log odds of having a teen birth versus delaying (or forgoing) childbirth emerged. For example, non-Hispanic black women exhibit 32 % greater odds of having a teen birth relative to non-Hispanic white women (odds ratio (OR) = $e^{0.28} = 1.32$). Young women residing with two biological parents exhibit lower odds than those residing in any other family type, as do women whose parents report higher levels of education.[7] Matching results are nearly identical to the OLS results—a difference of just under one year of education ($b = -0.93$). Rosenbaum sensitivity analyses (available upon request) suggest that an unmeasured factor affecting both the risk of teen childbearing and poor educational attainment would have to exhibit a magnitude of at least 2.9 to render the estimated average treatment effect from the matching results null. In other words, the unmeasured factor would need to have an effect far greater than that of non-Hispanic black ethnicity, family structure, or nativity.

Parametric maximum likelihood results (see Online Resource 1, Tables S3 and S4) suggest larger educational consequences. The baseline difference in educational attainment for teen and nonteen mothers is nearly three years (Table S3, Model 1, $b = -2.8$); this difference is reduced to just under two years with sociodemographic covariates added (Table S3, Model 2, $b = -1.87$). Exploratory analyses examined the inclusion of various combinations of

---

[7]Nearly all cases were retained (6,849 control and 890 treatment cases); the propensity score estimation was sufficiently balanced (results available upon request).

identifying variables to test the sensitivity of our results to various specifications. Results were extremely similar across specifications, suggesting that our results are rather robust with respect to the identifying variables included here. Note also that $\rho$ —the correlation between the error terms of the two equations—is significantly different from zero (95 % CI = 0.13, 0.33), thus indicating that teen childbearing is endogenous. Therefore, these results are preferable to a single-equation approach (OLS).

Semiparametric maximum likelihood results (see Tables 3–5) suggest an educational penalty of less than three-quarters of a year (Table 3, $b = -0.7$), an estimate similar to OLS and PSM results. Tables 3 and 5 report the parameters of the estimated discrete distribution as laid out in Eqs. (3) and (4). We follow Mroz (1999) and add points of support until the improvement in the likelihood function is smaller than the increase in the number of parameters. This leads us to a model with five points of support (see Table 4). The heterogeneity parameters are, for the most part, precisely estimated. For the teen birth equation, one of the mass points is a very large negative number (–27.98). Its corresponding probability weight is 0.58, suggesting that approximately 58 % of the sample would not have a teen birth regardless of values of control variables given that this mass point would overwhelm the effects of observed covariates and result in a predicted probability of a teen birth for these women close to zero.

Finally, given the estimated discrete distribution, we calculate the correlation between the error terms of the two equations ($\rho$) as .15, which is substantially lower than the estimate from the P-MLE model. Unfortunately, attaching a standard error to this estimate is difficult.

Parameter estimates and corresponding 95 % CI are graphed in Fig. 1. Note that confidence intervals for three estimates—OLS, PSM, and SP-MLE—overlap with one another; however, one estimate—P-MLE—does not. Taking the outer bands of the confidence intervals across all four methods produces a probable range of the educational consequences associated with teen childbearing: between 0.08 and 2.3 fewer years of schooling among teen mothers. The probable range based on point estimates across methods is between 0.7 and 1.9 fewer years of schooling.

## Discussion

Given the broad-ranging causal estimates of teen childbearing on educational attainment, we estimate these effects in a contemporary cohort of young adults using four statistical strategies: ordinary least squares (OLS), propensity score matching, parametric maximum likelihood estimation, and semiparametric maximum likelihood estimation. The effect sizes in the literature range from no difference in years of education, to 2.6 fewer years among teen mothers (Ashcraft et al. 2013; Ashcraft and Lang 2006; Fletcher and Wolfe 2009; Hofferth and Moore 1979; Holmlund 2005; Klepinger et al. 1999; Marini 1984; Moore and Hofferth 1980; Moore and Waite 1977; Ribar 1994; Rindfuss et al. 1980; Waite and Moore 1978). Our estimates also vary greatly, from 0.7 to 1.9 fewer years of schooling.

Nearly identical educational consequences were predicted using OLS (0.94 fewer years famong teen mothers) and propensity score matching (0.93). This similarity arises because (1) neither method directly accounts for the endogeneity of teen childbearing, and (2) propensity score matching differs from OLS only in using a semiparametric approach to estimate the treatment effect. Supplementary analyses (available upon request) revealed that adding numerous individual-, family-, school- and neighborhood-level characteristics reduced the treatment effect to 0.60 fewer years, or by 35 %, a number that perhaps is not at as large of a reduction as one might expect given the rich set of covariates added. However,

these results are similar to past research documenting diminished educational consequences comparing propensity score matching to OLS (Lee 2010).[8]

Parametric maximum likelihood results indicate that teen mothers complete nearly two fewer years of schooling (1.87). Observing an average treatment effect that is larger relative to non-experimental methods (e.g., OLS) is consistent with some past research using an instrumental variables specification with a population-level sample (Klepinger et al. 1995, 1999) but dissimilar to others that show either no effect or a much smaller effect (Marini 1984; Ribar 1994). It is also dissimilar to research using samples of pregnant teenagers that cite smaller educational consequences, ranging from no difference to 0.15 fewer years (Ashcraft et al. 2013; Ashcraft and Lang 2006; Fletcher and Wolfe 2009; Hoffman 2008; Hoffman et al. 1993; Hotz et al. 2005, 2008). Supplementary analyses limited our sample to pregnant teens and found that teen mothers completed 0.53 fewer years of education.[9] Although our estimate is greater in magnitude than some, it is lower relative to OLS, which is largely consistent with past research.

Semiparametric maximum likelihood results are more similar to those garnered from OLS and propensity score matching: just under three-quarters of a year less schooling among teen mothers (−0.70). We can use these results as leverage to evaluate the multivariate normality assumption presumed under parametric maximum likelihood. The dissimilarity between the two model results suggests that the normality assumption is questionable in this case. This conclusion is consistent with past research demonstrating that the multivariate normality assumption can provide misleading results within a Monte Carlo simulation (Mroz 1999) when the true error distribution is not normal. Bivariate probit models in particular, compared with semiparametric maximum likelihood models, can produce misleading findings if this assumption is violated (Guilkey and Lance 2012).

On the basis of these results, in combination with our replications of past research, we speculate that the wide-ranging estimates in past research are related more to the choice of statistical strategy than to the use of different data sets (focused on different cohorts). We make this assertion because we are able to reproduce nearly the identical range of estimates using four methods and a single data set. Thus, it does not appear that there is much of a (net) cohort difference in the "educational penalty" experienced by our sample (born in the 1980s) and studies of cohorts born in the 1940s, 1950s, and 1960s. It is possible that meaningful cohort differences are masked by factors working in opposing directions, but this remains speculation given that we do not examine data for other cohorts here.

Ultimately, we select the semiparametric maximum likelihood results as the preferred estimate of the treatment effect in the case of the educational consequences of teen childbearing. This method directly addresses the endogeneity of teen childbearing, does not impose stringent assumptions (e.g., joint normality assumption, ignorable treatment assignment assumption), and is not limited by the common support condition (as is propensity score matching); further, the coefficient estimates of other variables in this model are substantively plausible. Unlike OLS and propensity score matching, the semiparametric maximum likelihood model takes into account the fact that decisions regarding teen childbearing and educational attainment are made jointly, which is an important point emphasized in past research (Brien et al. 1999; Musick 1995; Rindfuss et al. 1980; Upchurch et al. 2002b; Upchurch and McCarthy 1990).

---

[8]Lee (2010) estimated the treatment effect of teen childbearing on college attendance at Wave III ($b = -0.16$, $p < .001$). A replication using nearly identical covariates and following women through Wave IV showed a smaller treatment effect ($b = -0.05$, $p < .001$), which is consistent with the life cycle argument further elaborated in the Discussion section.
[9]This reduced our sample to 972 women, 890 of whom had a teen birth and 120 of whom had a teen miscarriage.

The semiparametric maximum likelihood point estimate is smaller in magnitude than the OLS corollary, suggesting evidence of an underlying selection mechanism, contributing to both the risk of teen childbearing and diminished educational attainment, *in addition to* a causal effect of teen childbearing on young women's schooling. This point estimate also lies outside the 95 % CI for the OLS and propensity score models, which is consistent with the notion that endogeneity exists and that OLS produces a biased estimate of the treatment effect. It also underscores the fact that semiparametric maximum likelihood estimation adds something new to this series of replications. Based on the 95 % CIs shown in Fig. 1, it is also the most conservative method to select (if one were to choose only one of these four methods) given that its 95 % CI entirely encompasses the respective 95 % CIs for the OLS and propensity score models. The parametric maximum likelihood model is clearly quite different because its 95 % CI does not overlap with any other.

Here, results from semiparametric maximum likelihood are similar to OLS and propensity score matching; however, this finding is not generalizable to other substantive applications. Prior applications have shown these results can differ dramatically from OLS estimates (Angeles et al. 1998; Mroz 1999; Rindfuss et al. 2007, 2010). Therefore, comparing results across methodologies such as those examined here is a recommended strategy for future research. Currently, semiparametric maximum likelihood is available only in Fortran-based software, but work is underway to create a similar program in Stata that will make this method more accessible.

The larger magnitude of the parametric maximum likelihood estimate, albeit consistent with some studies using a similar strategy (Klepinger et al. 1995, 1999), raises the question of whether this modeling strategy indicates the presence of essential heterogeneity. In other words, the treatment effect may be heterogeneous across the sample because young women respond differently to the "treatment" of having a teen birth and because decisions to have a teen birth and continue one's education are jointly determined (Angrist and Evans 1996; Basu et al. 2007; Heckman et al. 2006; Rindfuss et al. 1980). This makes sense both substantively and statistically. In this scenario, some unobserved characteristics are positively related to both teen childbearing and educational attainment. Subsumed within this "black box" are factors related to a decision-making process about if or when to complete one's education. If this explanation holds, we cannot rule out the possibility that essential heterogeneity is also present in the semiparametric maximum likelihood model. However, the burden of evidence presented throughout this study (across numerous specifications in the full population sample as well as in a sample limited to pregnant teens) suggests that the educational consequences associated with teen childbearing are less than one year. Therefore, we are less concerned with the possibility that the semiparametric maximum likelihood results are misleading. Although we cannot test this directly, future work should rigorously evaluate the Basu et al. test for the problem of essential heterogeneity to continue advancing this literature.

We can, however, speculate on potential unobserved factors subsumed in this "black box." A likely contender is favorable self-selection. For example, if the adolescent's mother had a teen pregnancy that did not negatively affect the mother's educational outcomes, the daughter may be more likely to have a teen birth but also work to achieve her desired level of education. She may benefit from high levels of familial and structural support that assist her in doing so. Similar evidence is found on adverse self-selection into prenatal care (Rosenzweig and Schultz 1991; Wehby et al. 2009, 2012). Qualitative research on early childbearing also provides corroborating evidence. That is, teen births are not necessarily desirable; but after a young woman becomes pregnant, her family increases its support of her and the baby (Edin and Kefalas 2005). This can influence a young woman to have an

early birth because of the knowledge that she will have support in pursuing her preferred level of education.

Substantively speaking, our semiparametric maximum likelihood results suggest that at the population level, teen mothers complete approximately three-quarters of a year less schooling (0.7) than young women who delay childbearing until at least age 20. Although our results across each statistical strategy unilaterally show a negative effect of teen childbearing, this is perhaps not as large of a differential as one might expect, particularly because many young women in our sample may continue to pursue higher education in the future. Several studies suggest that teen mothers often "recoup" educational losses in their late 20s and early 30s by going back to school or by gaining specialized employment training (Furstenberg et al. 1989; Hoffman 2008; Hotz et al. 2005, 2008). The argument here is that having a child at any point is difficult and has consequences for women's labor market participation and economic independence that can depend, in part, on the time of observation. Thus, observing our sample 10 years into the future may reveal smaller educational consequences than are suggested by the current analysis. Future research could further test this possibility using subsequent waves of Add Health data or replicate these analyses in another contemporary cohort.

Finally, a related question for future research is how we should conceptualize an "early" birth. Is a "teen" birth, in and of itself, problematic? Or, is simply transitioning to motherhood before the normative age at first birth consequential for young women? Contextualized within current fertility norms of delayed age at childbirth in the United States, this suggests that recent cohorts transitioning to motherhood at age 20 or 21 might be similarly disadvantaged as teen mothers in the past. This could also be particularly reticent for recent cohorts given rising levels of schooling across the population (particularly for young women), increased necessity of dual incomes to maintain family stability, and increasing demand for educational credentials in the labor market. Indeed, research has documented negative social and economic consequences of childbearing in the early 20s, which is often composed largely of nonmarital childbearing (Guzzo and Hayford 2011; Hoffman and Maynard 2008; McLanahan 2009; Raymo et al. 2011). Therefore, future research on this topic may benefit from broadening the scope to include teen births and those to women aged 20 or 21.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

# References

Allison, PD. Fixed effects regression models. Newbury Park, CA: Sage Publications; 2009.

An CB, Haveman R, Wolfe B. Teen out-of-wedlock births and welfare receipt: The role of childhood events and economic circumstances. Review of Economics and Statistics. 1993; 75:195–208.

Angeles G, Guilkey DK, Mroz TA. Purposive program placement and the estimation of family planning program effects in Tanzania. Journal of the American Statistical Association. 1998; 93:884–899.

Angrist, JD.; Evans, WN. NBER Working Paper No 5406. Cambridge, MA: National Bureau of Economic Research; 1996. Schooling and labor market consequences of the 1970 state abortion reforms.

Ashcraft A, Fernández-Val I, Lang K. The consequences of teenage childbearing: Consistent estimates when abortion makes miscarriage nonrandom. The Economic Journal. 2013 Advance online publication. 10.1111/ecoj.12005

Ashcraft, A.; Lang, K. NBER Working Paper No 12485. Cambridge, MA: National Bureau of Economic Research; 2006. The consequences of teenage childbearing.

Basu A, Heckman JJ, Navarro-Lozano S, Urzua S. Use of instrumental variables in the presence of heterogeneity and self-selection: An application to treatments of breast cancer patients. Health Economics. 2007; 16:1133–1157. [PubMed: 17910109]

Brien MJ, Lillard LA, Waite LJ. Interrelated family-building behaviors: Cohabitation, marriage, and nonmarital conception. Demography. 1999; 36:535–551. [PubMed: 10604080]

Burton LM. Teenage childbearing as an alternative life-course strategy in multigeneration black families. Human Nature. 1990; 1:123–143. [PubMed: 24222049]

Card JJ, Wise LL. Teenage mothers and teenage fathers: The impact of early childbearing on the parents' personal and professional lives. Family Planning Perspectives. 1978; 10:199–205. [PubMed: 567590]

DiPrete TA, Gangl M. Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. Sociological Methodology. 2004; 34:271–310.

Edin, K.; Kefalas, M. Promises I can keep: Why poor women put motherhood before marriage. Berkeley, CA: University of California Press; 2005.

Fletcher JM, Wolfe BL. Education and labor market consequences of teenage childbearing. Journal of Human Resources. 2009; 44:303–325.

Furstenberg FF. The social consequences of teenage parenthood. Family Planning Perspectives. 1976; 8:148–164. [PubMed: 964349]

Furstenberg, FF.; Brooks-Gunn, J.; Morgan, SP. Adolescent mothers in later life. New York: Cambridge University Press; 1989.

Furstenberg FF Jr. Teenage childbearing as a public issue and private concern. Annual Review of Sociology. 2003; 29:23–39.

Geronimus AT, Korenman S. The socioeconomic consequences of teen childbearing reconsidered. Quarterly Journal of Economics. 1992; 107:1187–1214.

Grogger J, Bronars S. The socioeconomic consequences of teenage childbearing: Findings from a natural experiment. Family Planning Perspectives. 1993; 25:156–174. [PubMed: 8405341]

Guilkey, DK.; Lance, PM. Unpublished manuscript. Department of Economics, University of North Carolina; Chapel Hill: 2012. Estimation of non-random program impact when both the program variable and outcome variable are binary indicators: Monte Carol results for alternative estimators and empirical examples.

Guzzo KB, Hayford S. Fertility following an unintended first birth. Demography. 2011; 48:1493–1516. [PubMed: 21842326]

Harris KM. An integrative approach to health. Demography. 2010; 47:1–22. [PubMed: 20355681]

Harris, KM.; Halpern, CT.; Whitsel, E.; Hussey, J.; Tabor, J.; Entzel, P.; Udry, JR. The National Longitudinal Study of Adolescent Health: Research design. Chapel Hill: Carolina Population Center, University of North Carolina at Chapel Hill; 2009. Retrieved from http://www.cpc.unc.edu/projects/addhealth/design

Heckman J. Instrumental variables: A study of implicit behavioral assumptions used in making program evaluations. Journal of Human Resources. 1997; 32:441–462.

Heckman JJ, Singer B. A method for minimizing the impact of distributional assumptions in econometric models for duration data. Econometrica. 1984; 52:271–320.

Heckman JJ, Ichimura H, Smith J, Todd P. Sources of selection bias in evaluating social programs: An interpretation of conventional measures and evidence on the effectiveness of matching as a program evaluation method. Proceedings of the National Academy of Sciences. 1996; 93:13416–13420.

Heckman JJ, Urzua S, Vytlacil EJ. Understanding instrumental variables in models with essential heterogeneity. The Review of Economics and Statistics. 2006; 88:389–432.

Hofferth SL, Moore KA. Early childbearing and later economic well-being. American Sociological Review. 1979; 44:784–815. [PubMed: 533035]

Hoffman, SD. Updated estimates of the consequences of teen childbearing for mothers. In: Hoffman, SD.; Maynard, RA., editors. Kids having kids: Economic costs & social consequences of teen pregnancy. Washington, DC: Urban Institute Press; 2008. p. 74-92.

Hoffman SD, Foster EM, Furstenberg FF Jr. Reevaluating the costs of teenage childbearing. Demography. 1993; 30:1–13. [PubMed: 8379973]

Hoffman, SD.; Maynard, RA. The study, the context, and the findings in brief. In: Hoffman, SD.; Maynard, RA., editors. Kids having kids: Economic costs and social consequences teen pregnancy. Washington, DC: Urban Institute Press; 2008. p. 1-24.

Holmlund H. Estimating long-term consequences of teenage childbearing. Journal of Human Resources. 2005; XL:716–743.

Hotz J, McElroy SW, Sanders SG. Teenage childbearing and its life cycle consequences: Exploiting a natural experiment. Journal of Human Resources. 2005; 40:683–715.

Hotz, VJ.; McElroy, SW.; Sanders, SG. Consequences of teen childbearing for mothers through 1993. In: Hoffman, SD.; Maynard, RA., editors. Kids having kids: Economic costs and social consequences of teen pregnancy. Washington, DC: Urban Institute Press; 2008. p. 51-73.

Jones AS, Astone NM, Keyl PM, Kim YJ, Alexander CS. Teen childbearing and educational attainment: A comparison of methods. Journal of Family and Economic Issues. 1999; 20:387–418.

Klepinger DH, Lundberg S, Plotnick RD. Adolescent fertility and the educational attainment of young women. Family Planning Perspectives. 1995; 27:23–28. [PubMed: 7720849]

Klepinger D, Lundberg S, Plotnick R. How does adolescent fertility affect the human capital and wages of young women? Journal of Human Resources. 1999; 34:421–448.

Lee D. The early socioeconomic effects of teenage childbearing: A propensity score matching approach. Demographic Research. 2010; 23(article 25):697–736.10.4054/DemRes.2010.23.25

Levine DI, Painter G. The schooling costs of teenage out-of-wedlock childbearing: analysis with a within-school propensity-score-matching estimator. Review of Economics and Statistics. 2003; 85:884–900.

Marini MM. Age and sequencing norms in the transition to adulthood. Social Forces. 1984; 63:229–244.

McElroy SW. Early childbearing, high school completion, and college enrollment: Evidence from 1980 high school sophomores. Economics of Education Review. 1996; 15:303–324.

McLanahan S. Fragile families and the reproduction of poverty. The Annals of the American Academy of Political and Social Science. 2009; 621:111–131. [PubMed: 20204078]

Mollborn S. Making the best of a bad situation: Material resources and teenage parenthood. Journal of Marriage and Family. 2007; 69:92–104. [PubMed: 21921967]

Moore KA, Hofferth SL. Factors affecting early family formation: A path model. Population & Environment. 1980; 3:73–98.

Moore KA, Waite LJ. Early childbearing and educational attainment. Family Planning Perspectives. 1977; 9:220–225. [PubMed: 902716]

Mott FL, Marsiglio W. Early childbearing and completion of high school. Family Planning Perspectives. 1985; 17:234–237. [PubMed: 3842664]

Mroz TA. Discrete factor approximations in simultaneous equation models: Estimating the impact of a dummy endogenous variable on a continuous outcome. Journal of Econometrics. 1999; 92:233–274. [PubMed: 12322578]

Musick, JS. Young, poor, and pregnant: The psychology of teenage motherhood. New Haven, CT: Yale University Press; 1995.

Olsen RJ, Farkas G. Endogenous covariates in duration models and the effect of adolescent childbirth on schooling. Journal of Human Resources. 1989; 24:39–53.

Raymo, JM.; Carlson, MJ.; Perelli-Harris, B.; Lim, S-J.; Iwasawa, M. Paper presented at the spring meeting of the ISA RC28. University of Essex; UK: 2011 Apr. Educational differences in early childbearing: A cross-national comparative study. Retrieved from http://www.iser.essex.ac.uk

Ribar DC. Teenage fertility and high school completion. Review of Economics and Statistics. 1994; 76:413–424.

Rich LM, Kim SB. Patterns of later life education among teenage mothers. Gender & Society. 1999; 13:798–817.

Rindfuss RR, Bumpass L, St John C. Education and fertility: Implications for the roles women occupy. American Sociological Review. 1980; 45:431–447. [PubMed: 7406358]

Rindfuss RR, Guilkey DK, Morgan SP, Kravdal ÃY. Child-care availability and fertility in Norway. Population and Development Review. 2010; 36:725–748. [PubMed: 21174867]

Rindfuss RR, Guilkey D, Morgan SP, Kravdal ÃY, Guzzo KB. Child care availability and first-birth timing in Norway. Demography. 2007; 44:345–372. [PubMed: 17583309]

Rosenbaum, PR. Observational studies. New York: Springer-Verlag; 2002.

Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983; 70:41–55.

Rosenbaum PR, Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. American Statistician. 1985; 39:33–38.

Rosenzweig MR, Schultz TP. Who receives medical care? Income, implicit prices, and the distribution of medical services among pregnant women in the United States. Journal of Human Resources. 1991; 26:473–508.

Sanders, S.; Smith, J.; Zhang, Y. Unpublished manuscript. Department of Economics, University of Maryland; College Park: 2007. Teenage childbearing and maternal schooling outcomes: Evidence from matching.

Upchurch DM, Lillard LA, Aneshensel CS, Li NF. Inconsistencies in reporting the occurrence and timing of first intercourse among adolescents. Journal of Sex Research. 2002a; 39:197–206. [PubMed: 12476267]

Upchurch DM, Lillard LA, Panis CWA. Nonmarital childbearing: Influences of education, marriage, and fertility. Demography. 2002b; 39:311–329. [PubMed: 12048954]

Upchurch DM, McCarthy J. The timing of a first birth and high school completion. American Sociological Review. 1990; 55:224–234.

Waite LJ, Moore KA. The impact of an early first birth on young women's educational attainment. Social Forces. 1978; 56:845–865.

Wehby GL, Murray JC, Castilla EE, Lopez-Camelo JS, Ohsfeldt RL. Prenatal care demand and its effects on birth outcomes by birth defect status in Argentina. Economics & Human Biology. 2009; 7:84–95. [PubMed: 19059012]

Wehby GL, Prater K, McCarthy AM, Castilla EE, Murray JC. The impact of maternal smoking during pregnancy on early child neurodevelopment. Journal of Human Capital. 2012; 5:207–254. [PubMed: 22272363]

Wellings, K. Causes and consequences of teenage pregnancy. In: Baker, P.; Guthrie, K.; Hutchinson, C.; Kane, R.; Wellings, K., editors. Teenage pregnancy and reproductive health. London, UK: The Royal College of Obstetricians and Gynaecologists; 2007. p. 69-80.
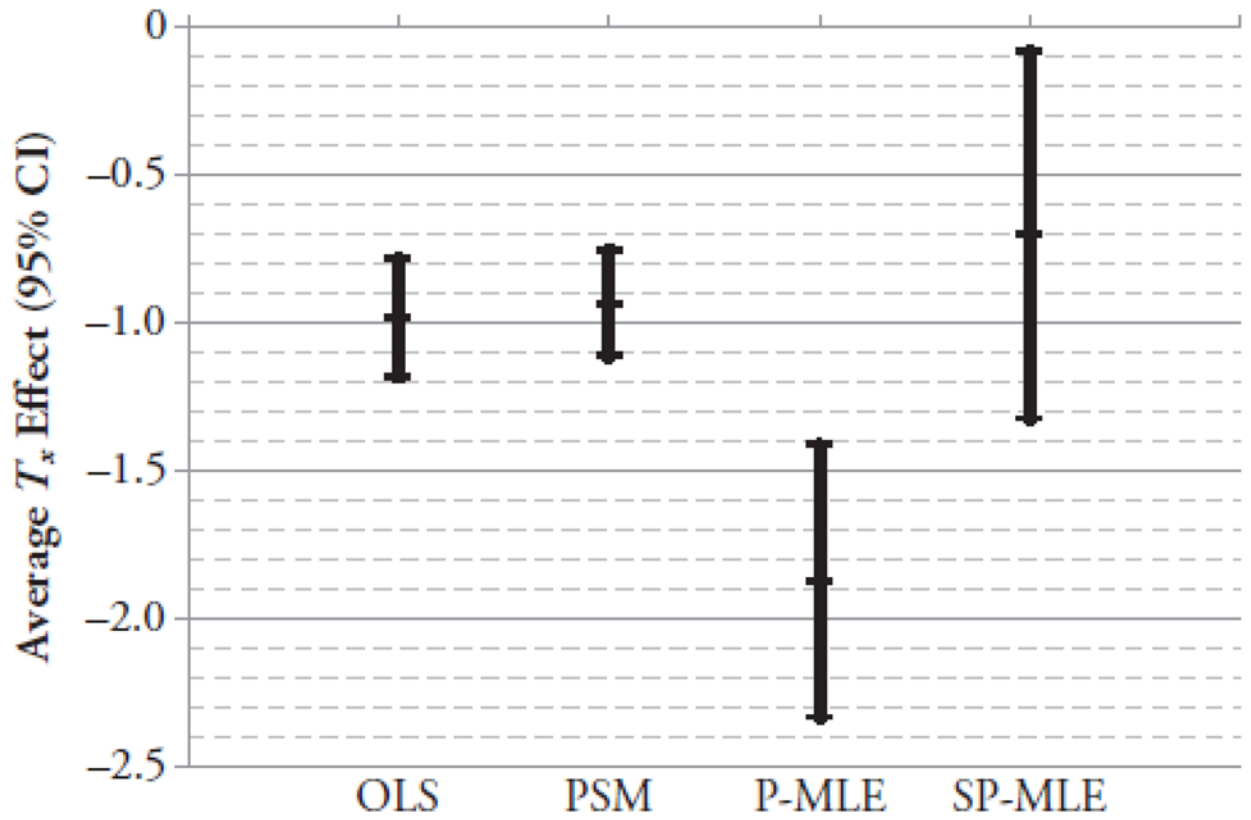
**Fig. 1.**
Predicted average treatment effects and 95 % confidence intervals for each method (OLS, PSM, P-MLE, and SP-MLE)

**Table 1**

Weighted descriptive statistics

|  | Mean | SE | Minimum | Maximum |
|---|---|---|---|---|
| Years of Completed Education (Wave IV) | 14.39 | 0.10 | 8 | 26 |
| Had Teen Birth (by age 19) | 0.12 | 0.01 | 0 | 1 |
| Sociodemographic Variables |  |  |  |  |
| Two-parent family (1 = yes) | 0.53 | 0.01 | 0 | 1 |
| Parent education | 5.28 | 0.09 | 1 | 9 |
| Income-to-needs ratio | 2.86 | 0.10 | 0 | 83.75 |
| Race/ethnicity |  |  |  |  |
| Non-Hispanic black | 0.17 | 0.02 | 0 | 1 |
| Hispanic | 0.12 | 0.02 | 0 | 1 |
| Non-Hispanic other | 0.04 | 0.01 | 0 | 1 |
| Non-Hispanic white (ref.) | 0.66 | 0.03 | 0 | 1 |
| Nativity (1 = foreign-born) | 0.04 | 0.01 | 0 | 1 |
| Add Health PVT score | 100.19 | 0.67 | 15 | 146 |
| Age (Wave I) | 15.40 | 0.12 | 11 | 21 |
| Per Capita Income | 12.96 | 0.42 | 1.84 | 68.82 |

*Notes:* Sample size is 7,870 women. Statistics are weighted and adjusted for clustering.

**Table 2**

Summary of analyses predicting years of completed education

| Method | Difference ($T_x$ – *Control*): Average Treatment Effect | 95 % Confidence Interval: Lower Bound | 95 % Confidence Interval: Upper Bound | N (*Control* vs. $T_x$) |
|---|---|---|---|---|
| OLS | −0.98[***] (0.10) | −1.18 | −0.78 | $N_C$ = 6,980 $N_{Tx}$ = 890 |
| Propensity Score Matching | −0.93[***] (0.09) | −1.11 | −0.75 | $N_C$ = 6,849 $N_{Tx}$ = 890 |
| Parametric Maximum Likelihood | −1.87[***] (0.23) | −2.33 | −1.41 | $N_C$ = 6,980 $N_{Tx}$ = 890 |
| Semiparametric Maximum Likelihood | −0.70[*] (0.31) | −1.32 | −0.08 | $N_C$ = 6,980 $N_{Tx}$ = 890 |

*Note*: Propensity score matching results indicate matches based on 1 % caliper radius method.

[*]
$p < .05$;

[***]
$p < .001$

**Table 3**

Results from semiparametric maximum likelihood models predicting years of education

| | 1 |
|---|---|
| Teen Birth (by age 19) | −0.70[*] (0.31) |
| Two-Parent Family (1 = yes) | 0.57[***] (0.05) |
| Parent's Education | 0.29[***] (0.01) |
| Income-to-Needs Ratio | 0.05[**] (0.02) |
| Race/Ethnicity | |
|    Non-Hispanic black | 0.61[***] (0.06) |
|    Hispanic | 0.46[***] (0.06) |
|    Non-Hispanic other | 0.17[*] (0.09) |
| Nativity (1 = foreign-born) | 0.63[***] (0.10) |
| Add Health PVT Score | 0.04[***] (0.002) |
| Age (Wave I) | 0.03[*] (0.01) |
| Per Capita Income | 0.02[***] (0.004) |
| Intercept | 5.63[***] |
| $\rho$ | .15 |
| $\sigma$ | 1.47 |
| Heterogeneity Parameters for Each Point of Support (1–5) | |
|   1 (normalized at 0) | 0 |
|   2 | 11.94 (0.27) |
|   3 | 2.58 (0.27) |
|   4 | 6.93 (0.40) |
|   5 | 2.60 (0.20) |

*Notes:* Analyses are unweighted. $N = 7,870$ women. Rho indicates the correlation between the error terms of the two equations. Sigma indicates the standard error of the outcome regression.

[*]
$p < .05$;

[**]
$p < .01$;

[***]
$p < .001$ (two-tailed tests)

**Table 4**

Probability weight for each point of support in semiparametric maximum likelihood model

| Point of Support | Probability Weight |
|:---:|:---:|
| 1 | .07 |
| 2 | .02 |
| 3 | .58 |
| 4 | .03 |
| 5 | .30 |

**Table 5**

Results from semiparametric maximum likelihood models predicting teen childbearing

| | 1 |
|---|---|
| State Requires Parental Consent for Abortions to Minors (1 = yes) | 0.22 (0.16) |
| State Provides Public Funding for Abortion in Most Cases (1 = yes) | −0.09 (0.14) |
| % Family Planning Clients Younger Than 20 | 3.06[*] 1.26 |
| Abortion Rate Among Women Aged 15–44 | −0.07 (0.05) |
| Average Medicaid Payment per Recipient | −0.15[**] (0.05) |
| Number of Ob/Gyn Physicians per 100,000 Women Aged 15–44 | 0.01 (0.01) |
| Two-Parent Family (1 = yes) | −0.99[***] (0.13) |
| Parent's Education | −0.22[***] (0.04) |
| Income-to-Needs Ratio | −0.11[**] (0.04) |
| Race/Ethnicity | |
|    Non-Hispanic black | 0.47[**] (0.15) |
|    Hispanic | 0.35[†] (0.18) |
|    Non-Hispanic other | 0.15 (0.27) |
| Nativity (1 = foreign-born) | −1.60[***] (0.33) |
| Add Health PVT Score | −0.03[***] (0.01) |
| Age (Wave I) | −0.08[**] (0.03) |
| Intercept | 3.54[***] |
| Heterogeneity Parameters for Each Point of Support (1–5) | |
|    1 (normalized at 0) | 0 |
|    2 | −0.58 (0.59) |
|    3 | −27.98 (1.00) |
|    4 | −1.74 (1.14) |
|    5 | 1.02 (0.57) |

*Notes:* Analyses are unweighted. $N$ = 7,870 women. Identifying variables are measured at census-tract level unless otherwise indicated.

[†] $p < .10$;

[*] $p < .05$;

[**] $p < .01$;

[***] $p < .001$ (two-tailed tests)