Forthcoming at *The Review of Economics and Statistics*

# The Schooling Costs of Teenage Out-of-Wedlock Childbearing: Analysis with a Within-School Propensity Score Matching Estimator

by

David I. Levine
Haas School of Business
University of California
Berkeley, CA 94720-1900
(510) 642-1697
levine@haas.berkeley.edu

and

Gary Painter
School of Policy, Planning, and Development
University of Southern California
Los Angeles, CA 90089-0626
(213) 740-8754
gpainter@usc.edu

*Abstract*: Teen out-of-wedlock mothers have lower education and earnings than do peers who have children later. This study uses the National Educational Longitudinal Survey of 1988 (NELS) to examine the extent to which the apparent effects of out-of-wedlock teen childbearing are due to preexisting disadvantages of the young women and their families. We use a novel method that matches teen mothers to similar young women in their junior high school (that is, prior to pregnancy). We find that out-of-wedlock fertility reduces education substantially, although far less than the cross-sectional comparisons of means suggest. We further find that this effect is larger among those with lower probabilities of having a child out of wedlock.

*Acknowledgment*: Paul Gertler suggested and Bryan Lincoln helped code the estimator we used. The authors acknowledge helpful comments from Mark Gilkey, Suzanne O'Keefe, Jeffrey Owings, Janet Yellen, seminar participants at U.C. Berkeley, and several referees. The second author will provide

SAS code on request.

*Our most serious social problem [is] the epidemic of teen pregnancies and births where there is no marriage.*

-- President Clinton, 1995 State of the Union Address

*Little of [their disadvantage] would be changed just by getting teen mothers to delay their childbearing into adulthood.*

-- Hotz, McElroy, and Sanders, 1999

Perhaps surprisingly, the authors of both of the quotations above agree that unwed teen mothers have severe disadvantages (for example, lower average education and earnings) compared with their peers who have children later. The disagreement that underlies the above quotations concerns how much (if any) of the correlations are causal. Estimates as to the size of this effect vary widely (e.g., Ribar, 1999). Importantly, a number of studies find that much of the apparent ill effects of teen parenthood are not causal (Geronimus and Korenman, 1992 and 1993; Hotz, Mullin, and Sanders, 1997, Hoffman *et al.*, 1993a, b; Hotz, McElroy, and Sanders, 1999; Ribar, 1994; also see the excellent review in Hoffman 1998). They find that most teen mothers were disadvantaged before motherhood. On average, even had these young mothers delayed childbearing, they would not have avoided all (or perhaps any) of the poor outcomes for themselves or their children.

This study uses the National Education Longitudinal Survey (NELS) of 1988 to examine the extent to which the links between teen out-of-wedlock childbearing and the young mothers' poor outcomes could have been predicted using the women's pre-motherhood characteristics. We examine these issues using both parametric methods and a novel within-school semi-parametric method based on matching. That is, we match each teen mother with a young woman who attended the same junior high school and who was similar on many observable characteristics in the eighth grade. We then compare the outcomes of teen mothers with their matches.

This method permits larger sample sizes than most previous studies. Taking advantage of the larger sample size, we also extend previous research by permitting the effect of teen out-of-wedlock motherhood on later educational attainment to vary depending on the pre-existing advantages and

disadvantages of the young women.

## Theory and Methods

The vast literature on teen pregnancies leads us to believe that teen out-of-wedlock mothers were disadvantaged in terms of parental income and education and other resources prior to their first childbirth. This familiar result reappears in the NELS dataset we examine here (Table 1). Moreover, in part due to these observable disadvantages, we expect young women who become teen out-of-wedlock mothers to have poor outcomes before their first childbirth--for example, low test scores and high rates of smoking and drug use.

Complementing these simple cross-tabs, the large literature on the "underclass" lists a number of factors that can cause both teen out-of-wedlock childbearing and low educational attainment. This literature emphasizes that America's least advantaged neighborhoods often combine low adult employment rates, high crime and gang activity, few fully employed and married adult role models, and poor schools. These factors, in turn, lead to a set of outcomes for youth including high rates of dropping out of high school, using drugs, committing crimes, and having a child out of wedlock. (Jencks and Peterson [1991] review this literature.)

Even in neighborhoods without such disadvantages, young women who are doing poorly academically are likely to find school more burdensome and to perceive the rewards of additional education as lower than do their classmates. Thus, precisely the young women at the highest risk of dropping out are also often the ones who perceive the lowest costs of out-of-wedlock teen pregnancy. Cross-sectional comparisons finding higher dropout rates for teen mothers overstate the causal links between out-of-wedlock childbearing and low education for teen mothers-to-be already disadvantaged prior to giving birth. As noted by Hoffman, et al. (1993a and b), most analyses of how teen motherhood affects young women's achievement lack adequate measures of family background and parental involvement in education. A number of studies have either used a socioeconomic status index provided by the dataset (e.g., Lee et al, 1994), created an ad hoc index of parents' characteristics (e.g., Herrnstein and Murray, 1994), or used a limited set of family background measures.

Fortunately, the topic of the effects of teen pregnancy has attracted some of the most careful

studies in the social sciences. Unfortunately, results often differ across datasets, across outcomes, or with modest changes in specification. Moreover, as the studies authors note, even with very good control groups, some selection based on unobservables remains.

One set of studies compares the children of teen mothers with the children of the teen mothers' sisters, who had children at a later age. Such a comparison implicitly controls for all aspects of the sisters' shared family background. In two of the three datasets examined, the children of the teen mother were not substantially disadvantaged compared to their cousins, whose mother had children at a later age (Geronimus and Korenman, 1993). Moreover, in one dataset, the young mothers were not disadvantaged compared with their sisters who delayed childbearing (Geronimus and Korenman, 1992). These results were not conclusive, however, as standard errors were often large and results varied by dataset. Hoffman et al. (1993a and b), agree with the Geronimus and Korenman findings that much of the cross-sectional correlation of teen childbearing and poor outcomes is not causal, but they emphasize the reasons to believe results from the dataset that indicates the largest harmful effects of early motherhood (*contra* see Geronimus and Korenman, 1993).

A second set of studies uses almost-natural experiments to identify nearly exogenous variability in teen childbearing. One set of studies examines the effects of miscarriage, an almost-natural experiment that typically delays the age of first birth by several years (Hotz, Mullin, and Sanders, 1997; Hotz, McElroy, and Sanders, 1999). In general, teenagers who became pregnant but whose first birth was delayed by miscarriage did not have systematically better outcomes than their peers who carried their baby to term. Teen mothers had lower high school graduation rates but were advantaged on several other outcomes. For example, teen mothers had less education and correspondingly more years of work experience. Also, the relatively good earnings of teen mothers during their twenties may not persist as the nonmothers with more education gain labor market experience.

In separate studies, David Ribar (1994) and Klepinger, Lundberg, and Plotnick (1995) examine age at menarche, noting that earlier age at menarche leads to more years at risk of becoming pregnant (1994). (They also used several other instruments.) Ribar, but not Klepinger et al., has found that controlling for the endogeneity of teen childbearing eliminates any negative impact of teen births on high

school completion.

The conclusions of both sets of studies indicate that the apparent disadvantages of teen motherhood are due in large part to the disadvantages of the mothers involved, not to their young age. Unfortunately, all studies have the problem of identifying a good control group. For example, the Geronimus and Korenman (1992, 1993) studies examine pairs of sisters to control for unobserved family background. At the same time, the sister who had a child as a teenager often differed systematically in other ways from the sister who did not (Geronimus and Korenman, 1993). In addition, this sample was necessarily drawn from larger families (families with at least two children), and therefore may not be representative of all families.

The studies by Hotz and his colleagues compare young women who became pregnant but had a miscarriage with women who experienced the birth of a child. These studies depend on the assumption of miscarriages (particularly reported miscarriages) being random events, yet "there are important reasons for believing that this is not the case" (Wolfe, Pence, and Haveman, 1999; also see Hotz, Mullin, and Sanders, 1997, who address some of these issues by bounding the importance of noncausal channels). Similarly, age at menarche may be correlated with unobservable factors; in other settings it has correlated with race and poverty.

Both of these methods identify clever control groups, but neither method is applicable in our dataset. Thus, we use a propensity score matching method, described below, to identify a suitable comparison group that is comparable on a rich set of family and youth characteristics.. We extend existing propensity matching methods to incorporate the counterpart of school fixed effects. That is, we compare the outcomes of a teen out-of-wedlock mother with someone of the same race from her junior high school. This matching controls for many observable and unobservable features of the family and neighborhood.

An advantage of this approach over those mentioned previously is that we are able to utilize a significantly larger sample of teen mothers than do most previous studies. (Hotz et al., 1999, for example, examines a sample with more teen mothers but with only 69 controls – that is, women who miscarry.) Moreover, our method selects a more-similar control group than do standard regression

analyses. At the same time, as in any nonexperimental study, additional unobserved factors may affect both a young woman's decision to have a child out of wedlock and her decision to continue her education. The method described below details our attempt to control for many of these unobserved factors.

## *Methods*

An "ideal" experiment to measure the effect of out-of-wedlock childbearing would pick matched pairs of young unmarried women with identical schools, race, academic ability, family income, smoking behavior, etc., and randomly have half of them carry a baby to term. To describe the ideal experiment is to assure both its impossibility and ethical undesirability.

The challenge, then, is to identify a good control group. Below, we introduce a within-school propensity-score matching model and contrast its results with a standard parametric regression method. The standard parametric method estimates the coefficient of teen out-of-wedlock motherhood when predicting youth outcomes and then examines how the estimated coefficient declines as additional controls are added. Thus, we, like authors of the previous literature, estimated several logit models:

$$Pr(y=1) = F(b_1 \cdot \text{teen childbearing}), \quad (1)$$

$$Pr(y=1) = F(b_2 \cdot \text{teen childbearing} + C_2 \cdot X), \quad (2)$$

where

y = Educational outcomes: dropping out of high school or attending college,

X = characteristics that preceded the birth of the child, such as parental education and demographics, as well as eighth-grade characteristics of family and child such as family income and child test scores in 1988,

and

F(.) is the cumulative logistic distribution: $F(z) = e^z / (1 + e^z)$.

We transformed the logit coefficients from models (1) and (2) into predicted changes in probabilities of each outcome for teen out-of-wedlock mothers compared with similar others. To the

extent that the correlation between teen out-of-wedlock childbearing and poor outcomes is causal, the estimated effect of teen childbearing should not change much when controlling for preexisting characteristics of the family. Conversely, if the estimated effects are strongly affected by the inclusion of preexisting conditions, then most of the measured effects of teen childbearing are due to pre-childbearing disadvantages. This method is used by many prospective studies (e.g., Painter and Levine, 2000, and the studies cited in Wolfe et al., 1999).

This standard method of estimating a logit or probit regression imposes strong restrictions on the functional form. Importantly, most women in the sample are quite different from most mothers-to-be. Nevertheless, in a typical regression, the non-teen-mother sample is quite important in estimating the counter-factual behavior of the out-of-wedlock mothers-to-be if they had not given birth out of wedlock. The assumption of a linear or logistic function permits data from all observations to be combined into one estimate, but the validity of that estimate is suspect when the "combining" function operates over people with very different characteristics. Thus, results are often sensitive to choice of functional form. Moreover, the estimating procedures create estimates that are complex averages of the typical effect of the treatment on the treated (that is, unwed teen mothers) and the effect of the treatment on those unlikely to ever become unwed teen mothers.

We used a variant of the method proposed by Rosenbaum and Rubin (1983) that requires weaker assumptions about functional forms. To see their result, consider the extreme case where we can observe all factors that affect both teen motherhood and education. (Actually we need a somewhat weaker assumption that conditional on the observable factors X, assignment to the treatment group--in this case, becoming a teen mother out of wedlock-- is not correlated with unobservables that predict later education.) In that case, all one must do to estimate the effects of teen out-of-wedlock childbearing is to match each treatment youth with a control who has the same observable characteristics. The mean difference in the treatment and matched controls' outcomes equals the true effect of teen out-of-wedlock childbearing on unmarried teen mothers. Note we are estimating the effect of the "treatment on the treated" -- a distinction that will arise again below.

Even if all important characteristics are observable, this method has the problem that the dataset

contains many characteristics. A problem arises here, as in many contexts: "Selecting a subset of comparison units similar to the treatment units is difficult because units must be compared across a high-dimensional set of pre-treatment characteristics" (Dehejia and Wahba, 2002). Thus, few of the mothers-to-be have a control with precisely the same junior high school, maternal education, family income, and other characteristics. Rosenbaum and Rubin (1983) suggest using the propensity score to make matching feasible. The propensity score is a young woman's estimated probability of receiving the treatment (in this case, becoming an unmarried teen mother) given her observable characteristics. Rosenbaum and Rubin prove that matching on the propensity score provides as powerful a control as matching on all observable characteristics. This technique reduces the problem from matching on the number of family and youth characteristics to matching on one dimension, the propensity score.

Dehejia and Wahba (2002) provide an important example where the matching method closely estimates the true treatment effects of a training program. Importantly, they find that the results from the matching method are closer to the experimental results than are the estimates from a standard regression.

**Matching with a clustered sample:** The NELS tracked for six years a sample with an average of eight female students per junior high school. A matching model that did not restrict itself to one junior high school would almost always match mothers-to-be to women from different junior high schools. At the same time, even conditioning on observable features of a family, knowing a youth's school is very useful for predicting youth outcomes (Jencks and Peterson 1991; Levine and Painter, 2001). This effect is due both to the causal features of schools and neighborhoods and due to the sorting that families engage in when choosing when to live. Moreover, observable characteristics of the school, the other students, and their families do not capture most of the actual effects of junior high schools (Hanushek, Rivkin, and Taylor, 1996).[1] Thus, important information is lost if school fixed

1. Similarly, neighborhoods vary enormously in terms of predicting the outcomes of youth, but observable qualities of a neighborhood explain only a small portion of that variance (Solon, Page, and Duncan, 2000). To illustrate the importance of school characteristics vs. fixed effects in predicting a typical outcome in the NELS, we regressed test scores normalized to have a standard deviation of unity against a long list of measures of the family, youth, and school. (Variables are listed in Table 1. The sample included only schools with at least 10 students.)

After extensive controls for the characteristics of the family, the school, and the families whose children attend the school, the standard error of the equation was .911. When we replaced the characteristics of the school

effects are ignored.

The presence of clustering by schools implies that matching within the junior high can often improve the match compared to matching nationwide. The junior high school match captures all the observable and unobservable features of the school and neighborhood. Within-school matching also controls for all unobservable characteristics of the family that led them to live in that neighborhood and send children to that school.

In many studies, this sorting by families is a problem that leads to bias in estimating the causality underlying observed neighborhood effects. In this study, our method of matching within junior high schools is useful regardless of whether the school fixed effect is causal or due to sorting. Similarly, in the context of labor markets, Heckman, Ichimura, and Todd (1997) note the increased efficiency of matching within the same region.

To take advantage of the importance of junior high schools, we performed a two-stage matching that restricted all matches within the junior high school. Specifically, we estimated the propensity score with a conditional (fixed-effects) logit regression that included a separate intercept $a_i$ for each junior high school (Chamberlin, 1980). Letting $T_{ij} = 1$ if observation $j$ at junior high $i$ is an unmarried teen mother (that is, treatment group), we have:

$$\Pr(T_{ij} = 1 \mid X_{ij}, a_i) = F(a_i + d \cdot X_{ij}). \quad (3)$$

The coefficient estimates d, but not the school-specific fixed effects $a_i$, can be recovered from this estimation. Fortunately, the differences in predicted probabilities for two women in the same junior high school can be recovered because the school-specific fixed effects $a_i$ cancel out.

Thus, for each young woman $i$ at high school $j$, we estimated her predicted probability of having a child out of wedlock ($T_{ij} = 1$) conditional on there being no other unmarried teen mother in her junior high school sample:

---

with a complete set of fixed effects for each junior high school to the regression, the standard deviation of the junior high fixed effects was .51 – that is, the average test scores of two randomly chosen schools differ about half as much as the test scores of two randomly chosen youth. This figure is inflated slightly by sampling error as we have an average of 17 youths per junior high. Adjusting for sampling error does not change the results meaningfully. The standard error of the equation declined to .795. That is, even after controlling for observable characteristics, fixed effects lowered the standard error by far more than did a long list of observable characteristics of the family and school.

$$\Pr(T_{ij}=1,\ T_{ik}=0,\ k \neq j) = \frac{\exp(X_{ij} \cdot d)}{\sum_{k=i}^{N_i} \exp(X_{ik} \cdot d)},$$

where $N_i$ is the number of classmates at junior high school $i$. We then matched each young mother-to-be with the young woman of the same race at her junior high school with the nearest propensity score.

As an additional screen, we required that each treatment woman have a match at her junior high school with a propensity score within 10 percentage points. Otherwise we did not analyze the outcome for that unwed mother-to-be. Heckman, Ichimura, and Todd (1997) stress the importance of the probability distributions of, in our case, being a teen mom coming from the same support. Below we discuss results with alternative bandwidths.[2]

Intuitively, consider an eighth grader who will soon have a child out of wedlock and already has low-income parents, low test scores, and many behavioral problems in an otherwise advantaged junior high school where all the young women in the NELS sample were academically successful. In this case, we had no good control group for this mother-to-be. A parametric method uses assumptions about functional forms to utilize information on the quite-different girls in the high school, while our method is less dependent on such assumptions.

We permitted a single control to match more than one treatment. This method minimizes the distance between treatments and their controls, but at the possible loss of some efficiency. Dehejia and Wahba (2002) find that in their sample this nearest-match algorithm performed better than algorithms that permit several "fairly near" controls to match a single treatment. Less than 17 percent of the controls were used more than once.

Assuming that a good match was found within the junior high school, the estimated percentage

2. A gap of 10 percentage points in the predicted probability of becoming a teen mother would be quite high if these were predicted probabilities from a logit or probit. Instead, these predicted probabilities are derived from the fixed-effect logit method that calculates the *relative* odds of each particular woman (conditional on exactly one woman in the sample of each junior high school) becoming a teen mother. To see the difference, consider a junior high with ten young women with a 2 percent chance of becoming a teen mother and one young woman with a 5 percent chance of becoming a teen mother. The high-risk woman has a 20 percent chance of being the teen mother if exactly one woman in this high school gives birth, while the other ten women have an 8 percent chance. Thus, the 3 percentage point gap in raw odds grows to a 12 percentage point gap in conditional probabilities, and our method would not use the high-risk woman as a control for the others.

point effect of teen out-of-wedlock childbearing on education ($B_{match}$) outcomes Y is the mean graduation rates of mothers-to-be (treatments) who have controls minus the mean graduation rates of those controls (with some controls entering more than once):

$$\left[\sum_{i=1}^{N_{match}} Y_{i,treatment} - Y_{i,control}\right] \Big/ N_{match}$$

where $N_{match}$ is the number of matched pairs.

Matching within a junior high school largely captures physical neighborhood effects. At the same time, a small proportion (less than 1 percent) of the mothers-to-be attended a private junior high school in 1988. Thus, the control women for these teens probably do not come from the same physical neighborhood as the mothers-to-be. At the same time, both the students in private school and their families probably resemble others in the private school more than others in their neighborhoods. In any case, the number of teen mothers in private schools was very small, and results were unchanged if they were dropped from the sample.

**Hard-to-match mothers-to-be:** Any matching method is less likely to find a close match for people who were most different from the typical member of the control group. In our setting, because we required a match within a junior high school, the relatively disadvantaged teen mothers-to-be were the least likely to have a close match. Intuitively, teen mothers were disadvantaged compared to most women. Thus, among the teen mothers, it was the most disadvantaged who were likely to be quite different from their classmates. In contrast, the relatively advantaged teen mothers were more similar to the average nonmother. Thus, the within-school matching method examines a less-disadvantaged set of teen mothers than the average teen mother.

To investigate the effect of out-of-wedlock motherhood on the group with no good match in their junior high schools, we applied our matching algorithm without regard to school. In the first stage equation that predicted teen out-of-wedlock childbearing, we replaced the logit with school fixed effects (equation 3) with a logit regression that controlled for a number of characteristics of the school and its students (the $Z_i$):

$$\Pr(T_{ij} = 1 \mid X_{ij}, Z_i) = F(d' \cdot X_{ij} + f Z_i). \quad (4)$$

We used the resulting coefficients to predict each young woman's probability of out-of-wedlock childbearing. We then identified the young woman of the same race with the closest predicted probability of out-of-wedlock motherhood, regardless of school.

**Combining within- and across-school matching:** An alternative way to view the within-school match is to consider it as one indicator of a good match but not a requirement. In a linear model with fixed effects, for example, it is easy to trade off slightly less similar school fixed effects to achieve slightly closer matches on family characteristics. The problem is more difficult with a discrete outcome such as having a child out of wedlock, as the fixed effects from the regression cannot be recovered.

Our solution is to use a composite estimator that matches all students who have a match whose predicted probability of becoming a teen mother is less than some bandwidth $P_0$ within their junior high school, but to use the closest match regardless of junior high school for others. To estimate the probabilities regardless of junior high school, we estimated a logit with a number of controls for characteristics of the school. We report results for the within-school bandwidth $P_0$ equal to 10 percentage points and perform sensitivity analyses with respect to choice of bandwidth.

**Testing if effect sizes differ for the most disadvantaged:** The standard parametric technique as well as the semi-parametric techniques described above both assume that the effect of teen out-of-wedlock motherhood on the proportion of young women graduating high school is similar for more- and for less-advantaged youth. In fact, this result may not hold. Theory is ambiguous as to whether the effect size will be largest for the most advantaged or for the most disadvantaged. For example, if more of the disadvantaged young women are closer to the margin of dropping out and if the more advantaged are able to utilize greater resources to provide resiliency after a shock, then the effect size would be largest for the disadvantaged. Working the other direction, it is possible that the relatively advantaged unmarried teen mothers women face greater stigma and other disadvantages if they have a child out of wedlock. If this is true, the effect size for the most advantaged would be larger than for the average female teen mother.

To test the possibility of different effect sizes, we divided the sample into quartiles based on the predicted probability of out-of-wedlock teen childbearing from equation (4). We then tested for

differences in effect sizes among the quartiles. We performed this test with the composite estimator to maximize sample size.

**Comparison with within-family matching:** To see if within-school propensity-score matching is identifying similar matches, we compared our results with the "gold standard" of matching estimators: within-family matches. Propensity-score matching methods match on observables and (in our case) schools. As such, there is always concern that mothers-to-be differ from their controls in unobservable features. Social scientists have used within-family matching to control for all aspects of a young woman's family background and community that she shares with her sisters (e.g., Hoffman et al., 1993a, b; Geronimus and Korenman, 1992, 1993). An important question is whether within-family matching provides very different results from our within-school matching.

To answer that question, we applied our method to geocoded data from the National Longitudinal Study of Youth (NLSY) used by previous researchers and compared those results with a procedure that matched sisters. If the two procedures provide similar results, our confidence that we are identifying good matches is greatly enhanced.

### *Data*

The National Education Longitudinal Study of 1988 (NELS) was sponsored by the National Center for Education Statistics and carried out by the National Opinion Research Center. NELS was designed to provide trend data about critical transitions experienced by young people as they develop, attend school, and embark on their careers. The base year (1988) survey was a multifaceted study with questionnaires for students, teachers, parents, and the school.

Sampling was first conducted at the school level and then at the student level within schools. The data were drawn from a sample of 1,000 schools (800 public schools and 200 private schools, including parochial institutions). Within this school sample, 25,000 eighth grade students were selected. The three follow-ups revisited the majority of the same sample of students in 1990, 1992, and 1994; that is, when the respondents were typically in the tenth grade, in the twelfth grade, and roughly two years after high school graduation. A randomized sample of approximately 14,000 students was

interviewed in the 1994 survey. These form the base sample for the estimation.

The NELS sample was stratified, clustered, and over sampled rare groups. The NELS provided sampling weights to control for the effects of sampling design. While the primary analysis was done using unweighted estimates, the results are unchanged when using weighted estimates.

**Teen motherhood**: The results we present are for those teens who experienced an out-of-wedlock birth. We also reran our models to include young mothers who married prior to giving birth, and the results were not changed substantively. All regressions dropped young women who gave birth prior to the first wave of the survey in 1988.

**Socioeconomic Status and Family Background**: Compared with most past studies, this study employs a much more detailed measure of family background and family involvement in education, which is intended to better isolate the effect of out-of-wedlock teen childbearing on outcomes. Variables were selected because past research (typically confirmed here) has found that they predicted teen pregnancy, low educational attainment, or both. All variables also had theoretical links to these outcomes, although we do not review all of the theory here.

The measures of socioeconomic status were created from both the parent and student questionnaire. The set of variables included occupational status (using Duncan's index), parental education, and family income. These variables were converted into z-scores with mean zero and standard deviation equal to one. When there were missing values for parental education because of a missing parent, these were given a z-score of 0 and categorical variables were included to note these important missing values.[3] To adjust family income for its size, family income was divided by the poverty line adjusted for family size. This is an improvement over most studies, which simply include some measure of family income in their estimated models. The log of this income/needs ratio (hereafter,

3. For father's education, this procedure is far from perfect. Most of these missing values are in female-headed households. Furthermore, it may be the case that these values are missing in precisely those families that are the most disadvantaged because of the least connection to the father. This will cause the coefficient on single parent to be biased upward. In addition, it is not clear in the NELS whether the value for a stepfamily is taken from the stepfather or the biological father. For these reasons, the analysis was replicated without the variable father's education, and the differences in the results were small and not statistically significant.

called income:needs ratio) was included for the student's eighth grade year. In addition, a composite measure of socioeconomic status (see Levine and Painter, 1999) was interacted with the racial/ethnic categories to allow for disparate effects.

Descriptors of a youth's family structure in eighth grade were included as well as indicators if there were any transitions in family structure during high school. The six family structures were intact families, single parent families with the biological mother present and with the father present, step-families with either the biological mother or father present, and those families with no biological parent present. In addition, four possible family transitions were included: divorce, remarriage, both divorce and remarriage, and death of a parent.

To supplement this fairly standard list, we included a wide range of measures that prior research suggests are indicators of advantages or disadvantages for young women. From the student questionnaire, a number of variables are potentially important predictors of education. A first set of variables controlled for standard demographic characteristics: region, rural vs. urban vs. suburban, and a female categorical variable. A second set of variables were indirectly related to parental involvement in education but were not exogenous to the outcome variable. These included whether a foreign language was spoken in the home, whether the mother or father was foreign born, the number of siblings, and whether the home had a library card, magazines, and many books.

From the parental questionnaire, indicators were obtained for whether the family was one of five religions and any of four levels of religious observance. These variables may proxy for how closely a family is knit as well as proxy for the social capital available to the children (Coleman, 1990). A categorical variable indicating whether the teen's mother had also been a teen when giving birth was also included. (Unfortunately, the dataset does not indicate whether the parents were married when the teen was born.)

Four variables partially capture parents' involvement in the young woman's life and education. The first variable was equal to one if the parent belonged to a parent-teacher association or related organization, or volunteered at school. The second variable was equal to one if the parent helped the child with homework. Next, a categorical variable for whether the child had participated in clubs such

as Boy or Girl Scouts during elementary school was included to proxy for the quantity of time spent with the child outside of the home. Finally, a measure of the parent's expectation was included. It takes the value one if the parent expected the student to obtain education levels beyond high school.

Three variables measured whether the student often came to school without a pen/pencil and paper, without homework, and without books. Another variable indicated whether the student was ever held back in school, and a final indicator variable equals one if the student changed schools. This final variable is a strong univariate predictor of dropping out and may measure low social capital.

**Eighth Grade Status**: We used several measures of student status in eighth grade: whether the teen had behavioral problems (coded as present if the student had been disciplined at school more than three times or if the parents considered the child to have severe behavioral problems), emotional problems (coded as present if the parent said that the student had an emotional problem that could inhibit learning), smoked cigarettes, and used drugs (marijuana, and harder drugs), as well as the student's test scores. The student's test scores were taken from a set of cognitive math and reading tests taken in eighth grade (see Levine and Painter, 1999, for a full description of the cognitive tests). Further, variables were included that were identified by the National Center for Education Statistics (1992) as good predictors of a student dropping out. These include whether the youth changed schools previous to junior high, was ever held back a grade, cut class, or often came to school without books, homework, or pen and paper.

*Educational outcomes:* We examined two educational outcomes of the youth in 1994, when they were roughly age 20. The first was whether the young woman dropped out of high school; that is, had no high school diploma by age 20. Second, we examined the proportion that had started college by 1994. We discuss the subset that received a GED below.

Summary statistics for the analysis variables are presented in Table 1. The means are for the entire sample we analyze. Approximately 14 percent of the sample dropped out of high school, while 70 percent of the sample (and a higher proportion of the high school graduates) had attended some college by age 20. Thirteen percent of the young women had a child out of wedlock while a teenager.

## Results

Unmarried teen mothers suffered far worse outcomes than did other young women. Teen out-of-wedlock mothers had a dropout rate of 44 percent, 5 times the rate of other young women (9 percent). Among high school graduates, young mothers' rate of entering college by age 20 was less than half that of their peers (31 vs. 76 percent).

Although prior researchers have not achieved consensus on the precise extent to which the correlation is causal, all agree that much or most of the correlation is not causal.[4] Consistent with these prior findings, the NELS data shows unwed mothers-to-be were disadvantaged in eighth grade, before they gave birth (Table 1). Compared to young women who would not give birth out of wedlock before age 20, in the eighth grade, teen mothers-to-be were twice as likely to be living with a single mother (27 vs. 14 %), both of their parents' education was .4 standard deviation lower than their peers' parents, and their parents reported somewhat lower parental involvement. The families' income:needs ratios were only a third of the average and they had changed schools twice as frequently as other young women (41 versus 21%).

Moreover, prior to giving birth out of wedlock, the teen mothers-to-be exhibited less socially desirable behaviors and lower academic achievement than did their peers. By eighth grade, they had a half of a standard deviation lower test scores than did young women who would not become teenage unwed mothers. They were also twice as likely to smoke (11 vs. 5 %). Their parents and teachers were more than twice as likely to report behavior problems (18 vs. 7 %), and their rate of severe emotional problems, although low, was more than triple that of their peers (5.1 vs. 1.8 %). The teen mothers-to-be were also much more likely to come to school unprepared or cut class. In addition, they were also more likely to have been held back in school (29 vs. 11 %).

**Logit results:** The logit results show that the effect of out-of-wedlock teen motherhood on the

4. Most past researchers have examined all teen births, while we examine only teen births out of wedlock. Some past researchers have looked at long-term effects on teen mothers, while our dataset only contains data on short-term effects. In addition, most past researchers have compared teen mothers to mothers who had first births in their twenties. Our comparison group includes all other women. For all of these reasons, we probably have a larger gap in education outcomes than those in other datasets. Nevertheless, these differences in data should not affect our main result. For example, when we pooled both married and unmarried teen mothers, our basic results were unchanged.

high school dropout rate fell from 35 percentage points in the raw data to 12.8 percentage points when controlling for demographic and eighth grade characteristics of the young women and their families (Table 2). These are the estimated logit effects when the logit coefficients were evaluated at the sample mean, as most social scientists do. As such, they attempt to answer the question of how unmarried teen childbearing affects a young woman who is unlikely to become a young mother (that is, this estimate attempts to recover the effect of the treatment on the untreated.). This 65 % decline is roughly consistent with findings from quasi-experimental methods (Hotz et al., 1997) and from methods using sisters as matches (Geronimus and Korenman, 1992, 1993).

Importantly, the estimated effects of teen childbearing are larger when the logit coefficients were evaluated at the mean of the sample of mothers-to-be. These estimates address a question that is closer to what the data can actually answer, as we can not estimate the effects of out-of-wedlock childbirth on nonmothers. Because the logit function is nonlinear, the simulated effect size of having a child out of wedlock differs by the probability of being a teen mother (see Figure 1). The estimated increase in the probability of dropping out due to having a child out of wedlock is 19.7 percentage points when evaluated using the mean of the sample of mothers-to-be, instead of the 12.8 percentage points when evaluated using the average characteristics of all women. Correspondingly, even our very good controls reduce less of the gap when we evaluate the logit coefficients at the average characteristics of the mothers-to-be.

Similarly, the effect of teen pregnancy on college attendance was 44.3 percentage points in the raw data (Table 1) and declined by more than half to 18.7 percentage points when the logit coefficients were evaluated at the sample mean. As with dropouts, the effect size rose to 28.7 percentage points when evaluated at the characteristics of the average teen mother-to-be.

**Semi-parametric within-school matching method:** A primary contribution of this paper is to compare the estimated effect size using the alternative within-school matching method. Our first-stage conditional logit estimates of the probability of teen motherhood are in the Appendix. As others have found and as showed up in the means, young women were more likely to become teen mothers if they came from single-parent homes, if they were black, and if they had low incomes.

To identify appropriate matches, we first set the cutoff for "similar" probability to be within 10 percentage points in the likelihood of being the sole teen mother (conditional on the junior high producing exactly one teen mother), and experimented to be sure other values did not appreciably change the results. We also required that matches be of the same race and attend the same junior high school. Fifty-five percent (470 of 840) of the young mothers-to-be had matches that met these criteria.[5]

As expected, mothers-to-be and their matches were much closer on observable pre-fertility behaviors than were mothers-to-be with other young women who did not match a teen mother. (comparing Tables 1 and 3). Of the comparisons we made between mothers-to-be and their matches, only two of the differences were statistically significant at the 5 percent level. This figure is roughly what one would expect by chance, given the many comparisons. In contrast, teen mothers-to-be were statistically significantly disadvantaged relative to the average young woman on 37 of the measures (Table 1).

**Results:** Using the within-school matching method, the estimated gap in dropout rates between teen mothers and their matches was 22.1 percentage points, a bit over half the 35 percentage point raw gap from the entire sample (Table 4). This 22.1 percentage points effect size of teen pregnancy was larger than the 12.8 percentage point estimate from the logit evaluated at the sample means, and the difference is statistically significant. At the same time, the 22.1 percentage point effect size was close to the effect size from the logit when the logit coefficients were evaluated at the characteristics of the mean mother-to-be. This convergence is to be expected as the latter logit results, like the matching model, estimated the effect of the treatment on the treated, while the former logit estimated the effect of the treatment on the average.

The matching method's 95 percent confidence interval stretches about 5 percentage points in each direction, over twice the confidence interval from the logit. The higher standard errors of the matching estimates are due to the sample size of 940 young women (470 pairs) for the matching as

5. The cutoff of .10 in predicted probability of teen motherhood is substantively neither enormous nor small. As discussed previously, these units are based on the conditional logit estimation. If we compared the matches to each other based on logit predictions, the differences between a teen mother and the selected match is less than .01.

opposed to almost 6,500 women in the logit. At the same time, most of the additional women analyzed in the logit sample were quite different from the mothers-to-be. Thus, the standard errors from the logit may be misleadingly narrow.

The raw gap in college attendance was 44 percentage points, while the gap estimated by the matching method was a much lower 26 percentage points (Table 5). The estimated effect of teen pregnancy on college attendance from the matching model (26 percentage points) was slightly smaller than the logit effect evaluated at the mean characteristics of teen mothers (29 percentage points). Although the results are similar in the matching model and when the logit effect is evaluated at the mean characteristics of teen mothers, the matching method is more credible as it relies on weaker assumptions concerning functional form.

**Using information on those without a close within-school match:** Our method requires that fairly similar matches be found within the junior high school of each teen mother-to-be. We would like to use additional information from the sample without a close match. To motivate the match-anywhere model, we first show that an estimate that matches regardless of school replicates our results on the sample that has close within-school matches. We then use the closest match anywhere on the hard-to-match students to create a composite estimator with a larger sample size and a representative sample.

Consider first the sample that has a close within-school match. The estimated decline in high school completion was 22.1 percentage points using within-school matching, which was almost identical to the 21.7 percentage point effect size estimated using the closest match anywhere (Table 6, rows A and B). The college results were almost as close, with a 25.7 percentage point gap using within-school matching and a 23.0 percentage point gap matching anywhere. Thus, it appears that the gain from controlling for unobservable school and neighborhood characteristics with the within-school match were largely offset by the somewhat more distant matches in terms of observable factors (as summarized by the propensity score).

The results concerning high school completion and college attendance were much different in the sample for which there was no close match within the school (Table 6, rows C vs. D). As noted earlier, these were disproportionately the most disadvantaged young women. Not surprisingly, if one matches

within school, the estimated effect of teen motherhood on education was very large. This result was due to the fact that these women were very different from their classmates.

If one instead identifies the closest match at any school, the difference in the effect size reverses. The teen mothers' difference in high school completion (compared with their matches) fell to 0.176, and the difference in college attendance rates fell to -0.143. The estimated effect size on college attendance rates for this sample was close to 9 percentage points lower than for teen mothers with close within-school matches. It is possible that the more-disadvantaged women were in situations where a child out-of-wedlock carried less stigma and was less disruptive, but we do not have direct evidence on this point. It is also possible that the women who were not teen mothers-to-be but were more likely to have adverse education outcomes were less affected by the out-of-wedlock childbirth due to the fact that they were more likely to have poor outcomes before the birth.

Finally, we present estimates concerning high school completion using a composite estimator on the entire sample (Table 6, rows E and F). The composite estimator matches within the junior high school if there is a match whose estimated probability of becoming a teen mother is within $P_0$; otherwise, the match is the closest at any school. Using the composite estimator and a bandwidth of $P_0 = .10$, the predicted probability of high school completion was 20.1 percentage points lower for teen mothers than for their matches. This estimate was similar to the 22.1 percentage point gap using only the within-school matches.

The two estimates of the effects of teen motherhood on the probability of starting college results were somewhat more distinct. The predicted probability of starting college was 20.7 percentage points lower for teen mothers than for their matches using the composite estimator, which was smaller than the 25.7 percentage point gap using only the within-school matches. This rise from 20.7 to 25.7 percentage points was both economically meaningful and statistically significant.

**Results by quartile of disadvantage:** When we analyzed the larger sample, our estimated effect sizes declined. We thought this difference might be due to mixing of two methods: within schools and between school matching. In fact, the difference was driven by effects of teen motherhood being smaller for the least advantaged.

To estimate separate effects for the more and less disadvantaged, we took advantage of our large sample and divided it into quartiles based on their estimated probability of becoming a teen mother. We then estimated the composite estimator separately on each quartile (Figure 2).

The estimated effects of teen parentage on education were larger for the more advantaged. For high school completion, the effect sizes were near 0.16 for the most disadvantaged quartile and about 0.20 for the most advantaged. For college attendance, the effect size differences were even larger (0.11 vs. 0.22). In both cases, these differences among quartiles were statistically significant at the 5 percent level.

For college, these results were expected because less than half of the matches from the most disadvantaged quartile attended college. Thus, it is not too unusual for the effect size to be smaller even though the proportionate decline in college attendance was similar.

For high school completion, the larger effect sizes for the more advantaged were more surprising. On average, the teen mothers who were from the upper quartile of the distribution were presumably further from the margin of dropping out had they not had a child. It is possible that the more-disadvantaged women were in situations where a child out of wedlock was less disruptive or the matches were more likely to drop out, but we have no direct evidence on this point.

**Comparison with within-family matching**

We applied our method to the NLSY data used by previous researchers (Geronimus and Korenman, 1992) and compared propensity-score matching results with results that come from matched sisters. If the two procedures provide similar results, our confidence that we are identifying good matches is greatly enhanced. Our procedures differ somewhat from our analysis in the NELS because we measure starting college in the NLSY up to age 28; this cutoff permits more catch-up than our younger NELS sample (which must have started college by age 20). More importantly, we can only match at the county level, as we do not have information on neighborhood or school in the NLSY.

We first identified the sample of 83 unmarried teen mothers with sisters present in the NLSY. (Previous authors often looked at all teen mothers; results were similar.) The dropout rate for these young mothers was 27 percent, while 22 percent started college by age 20 (Table 7). These figures are

far worse than the NLSY averages of 16 percent dropouts and 46 percent starting college.

As others have found, matching on sisters eliminates most, but not all, of the gap in educational attainment. Sisters of teen mothers dropped out 17 percent of the time and started college 35 percent of the time. The within-region propensity score matching method provided controls who were almost identical on average to the method of matching based upon sisters. The matches had a dropout rate of 15 percent (versus 17 percent for sisters) and a college starting rate of 39 percent (versus 35 percent for sisters). None of the gap between the propensity score matches and the sisters was close to statistically significant.

In brief, results were almost identical using the within-region propensity-score matching method as the sister matching. This fact increases our confidence that the within-school propensity score matching method identifies appropriate control groups.

***Robustness tests***

We performed a number of robustness tests of both the logit and matching results.

**GED:** It is possible that some of the higher dropout rate we observe in teen women who had children is a short-run effect due to disruption and that the effect of teen childbearing later declines. High GED rates for teen mothers, for example, is a main result found in Hotz et al. (1999). If the effects of teen childbearing decline as women age, then teen mothers who dropped out of high school would be more likely to return for a GED degree than other female dropouts.

We found no evidence that teen mothers were more likely to return to school. In fact, among those without a high school diploma by 1994 (that is, roughly at age 20), 26 % of the teen mothers and 36 % of other female dropouts had a GED (Table 1). The relative advantage of the non-teen-mother dropouts remained when looking at the matched sample (Table 2). Studies with more years of data can examine longer-term catch-up, as in Geronimus and Korenman (1992) and Hotz et al. (1999).

**Varying coefficients by race:** Several studies find that the effects of teen motherhood on graduation vary by race (GAO, 1998) . Both the matching and logit procedures should correctly reproduce the average result across races, but the result may not hold for any single race. In fact, the point estimates for the effects of teen motherhood on graduation were similar, with estimated effects 2

percentage points higher for blacks and 2 percentage points lower for Hispanics than for whites. These small differences were well under one standard error.

**Wider bandwidths:** We reran the results using the somewhat larger sample of young women who had a match within .20, not .10, in the predicted probability of becoming a teen out-of-wedlock mother. The advantage of this cutoff was that the sample grew from 470 with the .10 cutoff to 581 with .20 cutoff. The disadvantage was that the mothers-to-be and their matches now differed more on observable characteristics. The gap in the two groups' mean predicted probability of out-of-wedlock motherhood was 3.4 percentage points, which was statistically significant at the 5 percent level.

With the cutoff of .20 and slightly poorer matches but a larger sample size, the estimated effect of motherhood out of wedlock on dropping out of high school was 23.6 percent, which is substantively and statistically similar to the results with cutoff equal to the more conservative .10. This effect size after matching remained a bit over half the total cross-sectional effect of teen motherhood in the representative sample. Thus, the controls explain less than in the naive logit and about the same as with the smaller cutoff. As we expect, the less-perfect matching implies a slightly larger estimated gap.

We also re-estimated the composite estimator using cutoffs of .05 and .20 (instead of .10) for the bandwidth to use the within-school estimator instead of the closest match anywhere. Results were similar regardless of bandwidth.

**Alternative treatment of timing**: As Upchurch and McCarthy (1990) have emphasized, the timing of dropping out and teen birth can be important in determining the impact of teen out-of-wedlock childbearing on high school completion. Unfortunately, the NELS includes only annual data on dropouts. Thus, we erroneously classified some young women as having dropped out after becoming pregnant or giving birth, when, in fact, they gave birth (or at least knew they were pregnant) before leaving school. We reran the analysis dropping all cases where the departure from high school may have preceded the pregnancy. Results were quite similar.

**Controlling for the Propensity Score**: In spite of our matching, teen mothers were slightly less advantaged on most measures than their matched controls. We included the propensity score as a covariate in a regression using teen motherhood to predict educational attainment. Our specification

was a conditional logit with a fixed effect for each pair. The coefficient on the propensity score was small and not significant, while the effect size on teen motherhood was similar to that reported in Tables 4 and 5.

**Reuse of Some Control Observations:** In some cases, a single control was matched to multiple teen mothers. This procedure biased our standard errors because we did not take into account the non-independence of observations. When we dropped a teen mother if she was the second to match to a control, standard errors were almost identical.

## Discussion

Rates of teen pregnancy are very high in the United States. Almost two in five young women will become pregnant before they are 20. About half of these pregnancies will end in abortion or miscarriage, and about half in a live birth (Sylvester, 1994). Moreover, approximately one in five white children is born out of wedlock, roughly the same rate of childbearing out of wedlock that black women had when Moynihan decried the death of the black family in 1967. Moreover, about three out of five black children are born out of wedlock.[6]

The results we report support prior findings that a substantial portion of the relation between teen childbearing and high school completion is due to preexisting disadvantages of the young women, not due to the childbirth itself. At the same time, about half of the very large disadvantages remain using all methods regardless of controls. Moreover, the causal part of the effect appears largest for the most advantaged mothers-to-be relative to the least advantaged.

This analysis has provided several contributions to the existing literature on how out-of-wedlock childbearing affects education. First, we use the NELS, which has extremely good measures of the characteristics of young women and their families. The junior high match provides complete controls for school, neighborhood, and many unobserved family characteristics -- an important advance on most previous studies. Second, we use a propensity-score method that is less sensitive to functional forms

6. Importantly, the rising share of black births that are out of wedlock is due to a small increase in rates of out-of-wedlock births over the last 30 years and a dramatic decline in births within marriage—falling by two-thirds since the 1950s. Akerlof, Yellen, and Katz review the evidence and provide innovative theory for the rise in out-of-wedlock fertility (1996).

than is standard regression analysis. Third, we have more pairs available to us (470 or more) than do the other matching studies that use sisters as controls or studies of the contaminated natural experiment of miscarriage. Finally, we use our larger sample to permit effect sizes to vary by characteristics of the young women.

In spite of our good dataset and appropriate methods, our analysis does not control for all possible characteristics of the young women. Similar critiques hold, for example, in studies that use sisters as matches. Unwed teen mothers differ from their sisters and their matched classmates in observable ways; thus, it is likely they also differ in unobservable ways.

Finally, an important advantage of our method is that it uses a more representative sample of teen mothers than do other methods. The effect of sample selection can be quite large; in the NLSY, the dropout rate for unmarried teen mothers with sisters is 27 percent, while the dropout rate for the other unmarried teen mothers is 39 percent. In the sister's studies, on average the young women are relatively disadvantaged because of above-average family sizes. In the miscarriage studies, the teen mothers are relatively disadvantaged if miscarriages are more prevalent among the disadvantaged. In both cases, our results suggest that effect sizes within these disadvantaged populations are smaller than for the average teen mother, a result consistent with the literature's findings.

**Methodological implications:** Because clustering reduces costs of data collection, most household datasets have observations that are clustered geographically. At the same time, this clustering implies that many analyses can control for unobserved characteristics correlated with neighborhood (and usually school) by matching within the enumeration area. Solon, Page, and Duncan (2000), for example, use this method to study neighborhood effects. Other studies can use the matching method described here to control for neighborhood effects.

**Policy implications:** From a policy perspective, we (like others) find young unmarried mothers end up with lower education but have many disadvantages prior to giving birth that already predict low education. Thus, half or more of young mothers' disadvantages would not have been eliminated by the young women waiting until their twenties to have children.

At the same time, almost all estimates in this study and in many of its predecessors indicate that

substantial disadvantages remain that are plausibly due to becoming a teen mother. Thus, policymakers should not ignore the potential effectiveness of policies that delay first births in affecting young women's education and other outcomes. The question is what to do with these findings.

Out-of-wedlock teen childbearing is the result of a complex set of factors. Many of these factors reflect disadvantages that society should reduce regardless of their effects on education. For example, roughly half of teen out-of-wedlock births are to women who were sexually molested at some time (Sylvester, 1994). In addition, many young women (and men) do not believe that they are likely to be able to succeed academically in high school, or that a high school diploma will lead to further education or career success. Many young women (and men) lack basic information on pregnancy and sexuality, are not supported by peer groups that encourage wise choices such as delaying the start of sexual activity, and (when sexually active) do not have access to contraception.

On the one hand, the precise cost-benefit analysis for policies to address these problems depends in part on whether reducing teen childbearing increases education. On the other hand, reducing sexual molestation, improving our nation's worst high schools, making it clear to young people that graduating high school increases expected living standards, and giving young people the skills, knowledge, and resources to handle their sexuality wisely are policies that make sense regardless of how much of the correlation between teen pregnancy and educational attainment is causal.

## Table 1: Summary Statistics by Fertility Status

Entire sample

| | All Females | Non teen mothers | Mothers-to-be |
|---|---|---|---|
| N | 6,476 | 5,636 | 840 |
| **Family Structure** | | | |
| Intact in eighth grade | 0.673 | 0.705 | 0.456 * |
| Single - Female headed in eighth grade | 0.153 | 0.136 | 0.268 * |
| Single - Male headed in eighth grade | 0.014 | 0.014 | 0.012 |
| Stepfather family in eighth grade | 0.096 | 0.089 | 0.143 * |
| Stepmother family in eighth grade | 0.018 | 0.018 | 0.018 |
| Resided with no biological parents in eighth grade | 0.047 | 0.039 | 0.104 * |
| Divorced during high school | 0.065 | 0.062 | 0.086 * |
| Remarried during high school | 0.024 | 0.023 | 0.031 |
| Both a divorce and remarriage occurred during high school | 0.008 | 0.007 | 0.011 |
| Parent died during high school | 0.011 | 0.011 | 0.013 |
| **Family in 1988** (Young woman in eighth grade) | | | |
| Ethnicity - African American (omitted is Caucasian) | 0.109 | 0.087 | 0.261 * |
| Ethnicity - Latino American | 0.131 | 0.120 | 0.204 * |
| Ethnicity - Asian American | 0.069 | 0.077 | 0.020 * |
| SES*African American interaction term | -0.051 | -0.025 | -0.228 * |
| SES*Latina American interaction term | -0.110 | -0.093 | -0.219 * |
| SES*Asian American interaction term | 0.028 | 0.033 | -0.004 |
| Parental involvement in education | 0.519 | 0.539 | 0.389 * |
| Parents and children are involved in clubs | 0.834 | 0.847 | 0.748 * |
| Parents help with homework | 0.406 | 0.404 | 0.418 |
| Mother's education (z) | -0.065 | -0.004 | -0.475 * |
| Father's education (z) | -0.043 | 0.013 | -0.418 * |
| Mother was a teen when this daughter was born | 0.126 | 0.114 | 0.213 * |
| Eighth grade income/needs | 0.858 | 0.942 | 0.299 * |

**Table 1: Summary Statistics by Fertility Status**

Entire sample

| | All Females | Non teen mothers | Mothers-to-be |
|---|---|---|---|
| Father foreign born | 0.164 | 0.167 | 0.149 |
| Mother foreign born | 0.170 | 0.172 | 0.156 |
| Live in the south (omitted category is northeast) | 0.358 | 0.355 | 0.381 |
| Live in the west | 0.198 | 0.195 | 0.220 |
| Live in the central | 0.274 | 0.273 | 0.282 |
| Live in urban area (omitted category is suburb) | 0.247 | 0.243 | 0.274 |
| Live in rural area | 0.326 | 0.322 | 0.352 |
| Oldest child | 0.311 | 0.317 | 0.264 |
| Father's occupation {z} | -0.042 | 0.003 | -0.344 * |
| Father unemployed | 0.069 | 0.062 | 0.113 * |
| Mother's occupation {z} | -0.047 | 0.002 | -0.380 * |
| Mother unemployed | 0.295 | 0.288 | 0.344 * |
| Religious affiliation - Baptist (omitted religion is other Protestant) | 0.205 | 0.192 | 0.294 * |
| Religious affiliation - Catholic | 0.325 | 0.330 | 0.292 |
| Religious affiliation - other religion | 0.121 | 0.119 | 0.131 |
| Religious affiliation - missing religion | 0.036 | 0.035 | 0.043 |
| Religious affiliation - no religion | 0.028 | 0.028 | 0.033 |
| Religiosity - very religious (omitted religiosity is "Not at all religious") | 0.416 | 0.440 | 0.252 * |
| Religiosity - religious | 0.155 | 0.155 | 0.155 |
| Religiosity - somewhat religious | 0.172 | 0.172 | 0.171 |
| Number of siblings | 2.318 | 2.224 | 2.950 * |
| More than 50 books in home | 0.882 | 0.898 | 0.780 * |
| Has at least one magazine subscription | 0.741 | 0.765 | 0.579 * |
| Family has a public library card | 0.820 | 0.834 | 0.726 * |
| Parents expect the youth to continue in school past high school | 0.898 | 0.915 | 0.786 * |

## Table 1: Summary Statistics by Fertility Status

Entire sample

| | All Females | Non teen mothers | Mothers-to-be |
|---|---|---|---|
| **Young woman in 1988** (That is, in eighth grade) | | | |
| Behavioral problems reported by teacher or parents | 0.086 | 0.072 | 0.177 * |
| Emotional problems | 0.022 | 0.018 | 0.051 * |
| Cigarette smoking | 0.055 | 0.048 | 0.106 * |
| Eighth grade test scores (z) | -0.024 | 0.045 | -0.484 * |
| Cuts classes at school | 0.093 | 0.081 | 0.169 * |
| Often comes to school without a pen/pencil and paper | 0.156 | 0.150 | 0.196 * |
| Often comes to school without homework | 0.160 | 0.145 | 0.263 * |
| Often comes to school without books | 0.067 | 0.057 | 0.130 * |
| Ever held back in school | 0.133 | 0.110 | 0.287 * |
| Changed schools during elementary school or junior high | 0.239 | 0.213 | 0.410 * |
| **<Variables above this point are controls in Table 2.>** | | | |
| **Predicted probability of having a child out of wedlock** | | | |
| Predicted probability of a having a child out of wedlock based on characteristics of the young woman and her family; coefficients from the Appendix. | 0.127 | 0.096 | 0.239 * |
| **Outcomes 1992-94** (Aged roughly 18 to 20) | | | |
| Dropout | 0.136 | 0.090 | 0.440 * |
| College attender | 0.699 | 0.756 | 0.313 * |
| College attender (among those with a high school diploma) | 0.783 | 0.813 | 0.449 * |
| Received a GED (among those without a high school diploma) | 0.316 | 0.358 | 0.257 * |

* represents that the value for mothers-to-be is significantly different from non-teen mothers at the 5 percent level.

All variables above the row "Predicted Probability of having a child out of wedlock" are controls in tables 3 and 4.

Students attended 919 junior high schools. {Z} refers to z-coded variables.

**Table 2**
**Logit Results on How Controls Affect The Coefficient on Teen Childbearing for Dropout and Started College**

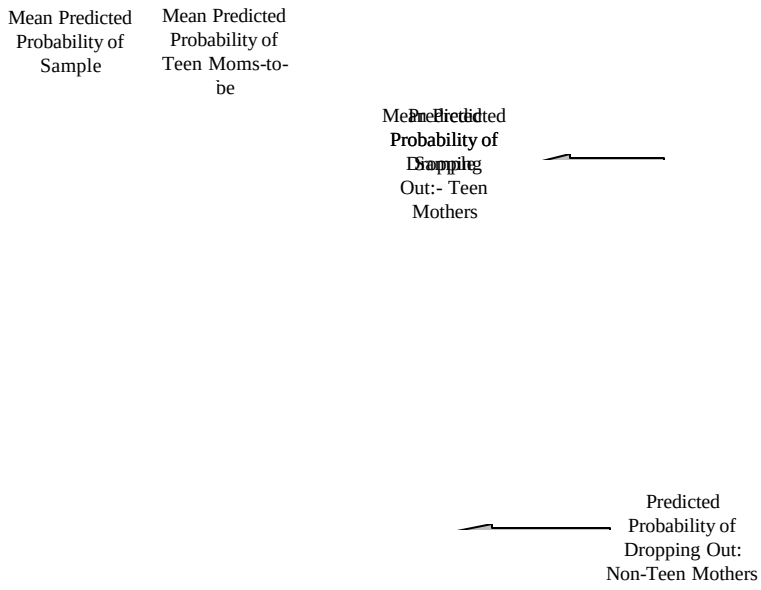
Reference group is young women who did not have a child out of wedlock.

| | No controls | Controlling for demographic and eighth grade characteristics of family and child (Evaluated at the mean of the sample) | Controlling for demographic and eighth grade characteristics of family and child (Evaluated at the mean of teen mothers-to-be) |
|---|---|---|---|
| ***Dropout*** (N = 6486) | | | |
| Had a Child out of Wedlock | 0.350 ** (0.014) | 0.128 ** (0.011) | 0.197 ** (0.011) |
| ***Started college*** (N = 6486) | | | |
| Had a Child out of Wedlock | -0.443 ** (0.019) | -0.187 ** (0.016) | -0.287 ** (0.016) |

*Notes:* Eighth grade characteristics of family and child include all controls listed as such in Table 1.
** represents different from zero at the 5 percent level.

**Figure 1**

**The Probability of Dropping Out of High School as a Function of the Predicted Probability of Having a Child Out of Wedlock**



## Table 3: Summary Statistics by Fertility Status

Matched sample

| | Mothers-to-be | Matched non-teen-mothers |
|---|---|---|
| N | 470 | 470 |
| **Family Structure** | | |
| Intact in eighth grade | 0.545 | 0.604 |
| Single - Female Headed in eighth grade | 0.219 | 0.226 |
| Single - Male Headed in eighth grade | 0.011 | 0.021 |
| Stepfather family in eighth grade | 0.128 | 0.094 |
| Stepmother family in eighth grade | 0.019 | 0.019 |
| Resided with no Biological Parents in eighth grade | 0.079 | 0.036 * |
| Divorced during High School | 0.087 | 0.081 |
| Remarried during High School | 0.034 | 0.023 |
| Both a Divorce and Remarriage Occurred During High School | 0.011 | 0.013 |
| Parent died during High School | 0.015 | 0.009 |
| **Family in 1988** (Young woman in eighth grade) | | |
| Ethnicity - African American (Omitted is Caucasian) | 0.164 | 0.164 |
| Ethnicity - Latino American | 0.177 | 0.177 |
| Ethnicity - Asian American | 0.021 | 0.021 |
| SES Index*African American interaction | -0.113 | -0.129 |
| SES Index*Latino American interaction | -0.181 | -0.193 |
| SES Index*Asian American interaction | 0.003 | -0.005 |
| Parental Involvement in Education | 0.430 | 0.440 |
| Parents and children are involved in clubs | 0.796 | 0.794 |
| Parents help with homework | 0.409 | 0.411 |
| Mother's education (z) | -0.378 | -0.351 |
| Father's education (z) | -0.347 | -0.374 |
| Mother was a teen when this daughter was born | 0.187 | 0.166 |
| Eighth grade income/needs (that is, poverty line) | 0.517 | 0.568 |
| Father foreign born | 0.123 | 0.153 |
| Mother foreign born | 0.136 | 0.166 |
| Live in the south (Omitted category is northeast) | 0.368 | 0.368 |
| Live in the west | 0.226 | 0.226 |

## Table 3: Summary Statistics by Fertility Status

Matched sample

| | Mothers-to-be | Matched non-teen-mothers |
|---|---|---|
| Live in the central | 0.287 | 0.287 |
| Live in urban area (Omitted category is suburb) | 0.245 | 0.245 |
| Live in rural area | 0.404 | 0.404 |
| Oldest child | 0.291 | 0.309 |
| Father's occupation {z} | -0.326 | -0.309 |
| Father unemployed | 0.104 | 0.106 |
| Mother's occupation {z} | -0.277 | -0.232 |
| Mother unemployed | 0.328 | 0.306 |
| Religious affiliation - Baptist (Omitted religion is other Protestant) | 0.251 | 0.234 |
| Religious affiliation - Catholic | 0.300 | 0.302 |
| Religious affiliation - Other religion | 0.140 | 0.121 |
| Religious affiliation - Missing religion | 0.040 | 0.053 |
| Religious affiliation - No religion | 0.032 | 0.023 |
| Religiosity - very religious (Omitted religiosity is “Not at all religious”) | 0.311 | 0.338 |
| Religiosity - religious | 0.170 | 0.168 |
| Religiosity - somewhat religious | 0.194 | 0.189 |
| Number of siblings | 2.651 | 2.538 |
| More than 50 books in home | 0.836 | 0.836 |
| Has at least one magazine subscription | 0.619 | 0.666 |
| Family has a public library card | 0.787 | 0.753 |
| Parents expect the youth to continue in school past high school | 0.838 | 0.872 |
| **Young woman in 1988** (That is, in eighth grade) | | |
| Behavioral problems reported by teacher or parents | 0.100 | 0.089 |
| Emotional problems | 0.030 | 0.019 |
| Cigarette smoking | 0.072 | 0.055 |
| Eighth grade test scores (z) | -0.397 | -0.319 |
| Cuts classes at school | 0.123 | 0.098 |
| Often comes to school without a pen/pencil and paper | 0.172 | 0.155 |
| Often comes to school without homework | 0.191 | 0.177 |
| Often comes to school without books | 0.077 | 0.070 |

## Table 3: Summary Statistics by Fertility Status

Matched sample

| | Mothers-to-be | Matched non-teen-mothers |
|---|---|---|
| Ever held back in school | 0.198 | 0.151 * |
| Changed schools during elementary school or junior high | 0.336 | 0.344 |
| **Predicted probability of a having a child out of wedlock** | | |
| Predicted probability of a having a child out of wedlock based on characteristics of the young woman and her family. Calculated based on coefficients from Appendix. | 0.129 | 0.113 |
| **Educational Outcomes 1992-94** (Aged roughly 18 to 20) | | |
| Dropout | 0.370 | 0.149 * |
| College attender | 0.355 | 0.613 * |
| College attender (among those with a diploma – not necessarily matched, N = 296 and 400) | 0.473 | 0.693 * |
| Received a GED (among those without a diploma -- not necessarily matched, N = 174 and 70) | 0.310 | 0.400 |

***** represents that the t-test on the mean value for mothers to be is significantly different from matched non-teen mothers at the 5 percent level. Students in the matched sample attend 492 distinct junior high schools.

## Table 4: Within-School Propensity Score Matching and Dropout Rates

Two-by-two matrix of matched pairs' outcomes at end of high school
Entries are proportions of matched pairs with similar or dissimilar outcomes
N = 470 pairs

| | **Young women who would not soon become unwed mothers (matched controls group)** | |
|---|---|---|
| **Mothers-to-be (treatment group)** | **Dropped out** | **Graduated high school** |
| **Dropped out** | 0.079 | 0.291 |
| **Graduated high school** | 0.070 | 0.560 |

**Proportion who dropped out of high school:**

| | | | | |
|---|---|---|---|---|
| Teen mothers | 0.370 | | | |
| Matched controls | 0.149 | [95% conf. interval] | | |
| | | | --- | --- |
| Difference | 0.221** | | 0.169 | 0.274 |
| Ratio | 2.485** | | 1.972 | 3.133 |
| | | | | |
| Odds ratio | 4.152 | 2.821 | 6.270 | |
| McNemar's $?^2(1)$ | 63.62** | | | |

Notes: Odds ratio = % of pairs where control graduated and mother-to-be dropped out / % of pairs where mother-to-be graduated and control dropped out (that is, 0.291 / 0.070).
McNemar's $?^2$ tests if the odds ratio equals 1.
Confidence intervals and test statistics are described further in the text.

** implies rejects the hypothesis that the ratio or odds ratio of proportions equals one or that the difference in proportions equals zero at the 1% level.

## Table 5: Within-School Propensity Score Matching and College Attendance

**Two-by-two matrix of possible college attendance**
Entries are proportions of matched pairs with similar or dissimilar college attendance by 1994 (roughly age 20).
N = 470 pairs

| | **Young women who would not soon become unwed mothers (Matched control group)** | |
|---|---|---|
| **Mothers-to-be (treatment group)** | **Did not attend College** | **Attended College** |
| **Did not attend College** | 0.272 | 0.373 |
| **Attended College** | 0.115 | 0.240 |

**Proportion attending college at age 20**

| | | |
|---|---|---|
| Teen mothers | 0.355 | |
| Matched controls | 0.613 | [95% conf. interval] |
| | | --- |
| Difference | -0.257 ** | -0.318 -0.197 |
| Ratio | 0.580 ** | 0.507 0.663 |
| Odds ratio | 0.309 | 0.223 0.421 (exact) |
| McNemar's $?^2(1)$ | 63.93 ** | |

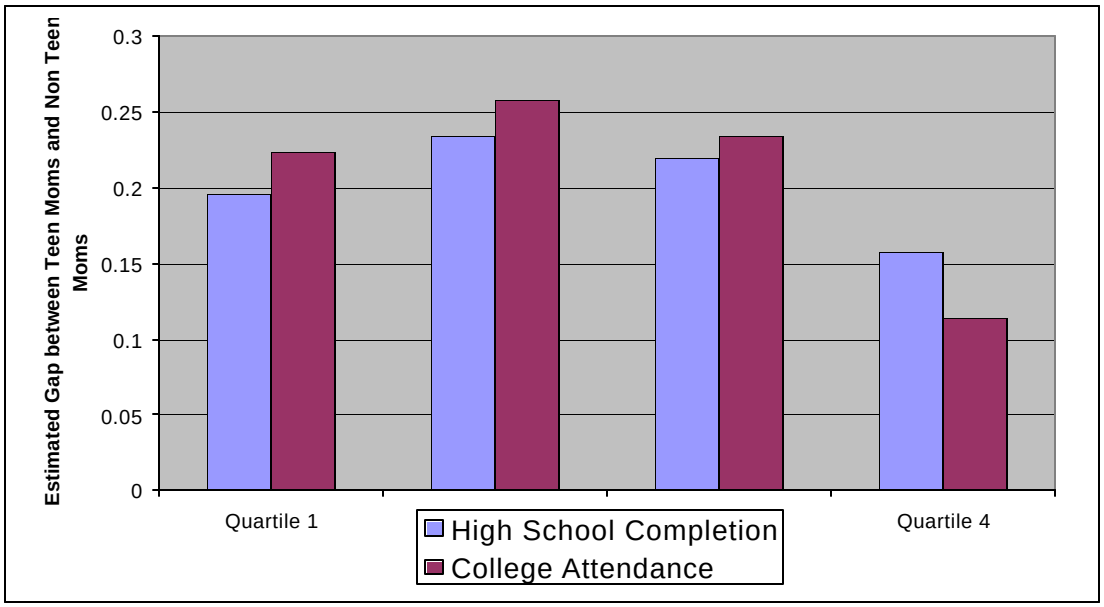
See notes to Table 4.

## Table 6: Comparing Within-School and Closest Anywhere Matching

| | Within-School Matching | Closest Anywhere Matching |
|---|---|---|
| ***Close Matches Within-School (N = 470 pairs)*** | | |
| A) Gap in Dropout Rate | 0.221 | 0.217 |
| B) Gap in College Attendance | -0.257 | -0.230 |
| ***No Close Matches Within-School (N = 370 pairs)*** | | |
| C) Gap in Dropout Rate | 0.330 | 0.176 |
| D) Gap in College Attendance | -0.354 | -0.143 |
| | **Composite Estimator** | |
| ***Full Sample (N = 840 pairs)*** | | |
| E) Gap in Dropout Rate | 0.201 | |
| F) Gap in College Attendance | -0.207 | |

Note: For the Composite Estimator, if a close match (estimated probability of teen pregnancy of control and match < .1) is available within the school sample, then that match is chosen; if not the closest match anywhere is chosen.

**Figure 2: Effect Sizes of Teen Motherhood Out of Wedlock, Estimated Separately for Each Quartile of the Predicted Probability of Teen Motherhood Out of Wedlock**




Note: Quartile 1 is the most advantaged, and Quartile 4 is the most disadvantaged, as measured by the predicted probability of having a child out of wedlock.

**Table 7: Comparing Within-Family Matching to the Propensity Score Matching in the NLSY**

| | N | Proportion dropout | Proportion starting college by age 28 |
|---|---|---|---|
| Unmarried teen mothers with sisters in the NLSY | 83 | .27 | .22 |
| Sisters | 83 | .17 | .35 |
| Within-region propensity score match | 83 | .15 | .39 |

Gaps between sisters and teen mothers are statistically significantly different at the 8 percent level for dropouts and 4 percent level for college. Gaps between propensity score matches and teen mothers are significantly different at the 4 percent level for dropouts and 2 percent level for college. Gaps between sisters and propensity score matches were not statistically significant.

Note: For the Propensity Score Estimators, if a close match (estimated probability of teen pregnancy of control and match < .1) is available within the county and with the same race, then that match is chosen; if not the closest match anywhere is chosen. For the propensity score matching related to the sister's sample, 71 of 83 teens with out of wedlock births had close matches within the county. The other 12 were matched by race but outside of the county.

## References


Akerlof, George A., Janet L. Yellen, and Michael L. Katz, "An analysis of out-of-wedlock childbearing in the United States," *Quarterly Journal of Economics*, May 1996, 111(2): 277-317.

Chamberlin, G. (1980): "Analysis of Covariance with Qualitative Data," *Review of Economic Studies*, 47, pp. 225-238.

Coleman, James. 1990. *Foundation of Social Theory*. Cambridge: Harvard University Press.

Dehejia, Rajeev, and Sadek Wahba (2002), "Propensity Score Matching Methods for Non-experimental Causal Studies," *Review of Economics and Statistics*, 84(1), 151-61.

U.S. General Accounting Office (GAO), *Teen Mothers: Selected Socio-Demographic Characteristics and Risk Factors*, Washington DC, June 30, 1998, HEHS-98-141.

Geronimus, Arline T.; Korenman, Sanders. "The Socioeconomic Consequences of Teen Childbearing Reconsidered," *Quarterly Journal of Economics*, 107, 4 (Nov 1992):1187-1214.

Geronimus, Arline T; Korenman, Sanders. "The socioeconomic costs of teenage childbearing: Evidence and interpretation." *Demography*, 30, 2 (May 1993): 281-290.

Hanushek, Eric A.; Rivkin, Steven G.; Taylor, Lori L. (1996): "Aggregation and the Estimated Effects of School Resources," *Review of Economics and Statistics*, 78 (4), 611-27.

Herrnstein, Richard, and Charles Murray (1995). *The Bell Curve.* New York: Free Press.

Heckman, James J., Hidehiko Ichimura, and Petra E. Todd (1997): "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme," *Review of Economic Studies*, 64, 605-654.

Hoffman, Saul, "Teenage Childbearing Is Not So Bad After All ...Or Is It? A Review of the New Literature." *Family Planning Perspectives*, v. 30, n. 5, Sep.-Oct. 1998: 236-239.

Hoffman, Saul D., Foster, E Michael, Furstenberg, Frank F Jr., "Reevaluating the costs of teenage childbearing." *Demography*, 30, 1 (Feb 1993): 1-13.

Hoffman, Saul D., Foster, E Michael, Furstenberg, Frank F Jr., "Reevaluating the costs of teenage childbearing: Response to Geronimus and Korenman." *Demography* 30, 2 (May 1993): 291-296.

Hotz, V. Joseph, C. Mullin, and Seth G. Sanders (1997): "Bounding Causal Effects Using Data from a Contaminated Natural Experiment: Analyzing the Effects of Teenage Childbearing," *Review of Economic Studies*; 64(4), 575-603.

Hotz, V. Joseph, Susan Williams McElroy, and Seth G. Sanders (1999): "Teenage Childbearing and its Lifecycle Consequences: Exploiting a Natural Experiment," NBER w.p. 7397, Cambridge, MA.

Jencks, C,. and P.E. Peterson (1991), eds., *The Urban Underclass*, Brookings, Washington, DC.

National Center for Education Statistics (1992). *Characteristics of At-Risk Students in the National Education Longitudinal Study of 1988*, Washington, DC, August.

Klepinger, Daniel, Shelly Lundberg and Robert Plotnick, "Adolescent Fertility and the Educational Attainment of Young Women," *Family Planning Perspectives*, 27, 1, Jan.-Feb. 1995: 23-28.

Lee, Valerie, David Burkam, Herbert Zimiles, and Barbara Ladewski (1994). "Family Structure and Its Effect on Behavioral and Emotional Problems in Young Adolescents," *Journal of Research on Adolescence* 4(3): 405-437.

Levine, David I. and Gary Painter (1999): "The NELS Curve: Replicating The Bell Curve with the National Education Longitudinal Survey," *Industrial Relations* 38(3): 364-401.

Levine, D. I. and G. Painter (2001). "How Much of School Effects is Just Sorting? Identifying Causality in the National Education Longitudinal Survey." University of Southern California Working Paper. [http://www-rcf.usc.edu/~gpainter/schools.pdf]

Painter Gary, and David I. Levine (2000), "Family Structure and Youths= Outcomes: Which Correlations are Causal? " *Journal of Human Resources*, 35(3): 524-549.

Ribar, David C., "The socioeconomic consequences of young women's childbearing: Reconciling disparate evidence," *Journal of Population Economics*, (1999) 12: 547-565.

Ribar, David C., "Teenage Fertility and High School Completion, " *Review of Economics and Statistics*, Vol. 76, No. 3 (Aug. 1994), pp. 413-424.

Rosenbaum, P. and D. Rubin (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects, " *Biometrika*, 70: 41-55.

Stata, *Stata Reference Manual Release 6*, College Station TX, 1999.

Solon, Gary, Marianne E. Page, and Greg J. Duncan: "Correlations between Neighboring Children in their Subsequent Education Attainment, " *Review of Economics and Statistics*, Aug 2000, 82(3): 383-392.

Sylvester, Kathy, *Teenage Pregnancy: a Preventable Calamity*, Progressive Policy Institute Policy Report, Washington DC, November 1994.[http://www.dlcppi.org/texts/social/teenpreg.txt]

Upchurch, Dawn M., and James McCarthy (1990): "The Timing of a First Birth and High School Completion, " *American Sociological Review*, 55(2), 224-234.

U.S. Department of Health and Human Services, "Fertility, Family Planning, and Women's Health: New Data From the 1995 National Survey of Family Growth, " Vital and Health Statistics Series, 23(19) [DHHS Publication No. (PHS) 97-1995], (June 1997). http://www.cdc.gov/nchswww/data/ sr23_19.pdf.

U.S. General Accounting Office [GAO], *Teen Mothers: Selected Socio-Demographic Characteristics and Risk Factors*, Letter Report, 06/30/98, GAO/HEHS-98-141, 1998.

Wolfe, Barbara, Karen Pence, and Robert Haveman (1999). "Youth Responses to Expected Income and 'Relationship' Consequences in Nonmarital Childbearing Choices: Are Youths Rational? " University of Wisconsin, Madison.

## Appendix: Predicted Probability of a Having a Child Out of Wedlock

Conditional Logit Estimates

Fixed effects for each junior high school

(N=3882)

| | Odds Ratio | Standard Error |
|---|---|---|
| **Family Structure** | | |
| Single - Female Headed in eighth grade | 1.657 * | 0.238 |
| Single - Male Headed in eighth grade | 1.237 | 0.509 |
| Stepfather family in eighth grade | 1.401 * | 0.210 |
| Stepmother family in eighth grade | 0.856 | 0.294 |
| Resided with no Biological Parents in eighth grade | 1.682 * | 0.309 |
| Divorced during High School | 1.132 | 0.191 |
| Remarried during High School | 1.041 | 0.295 |
| Both a Divorce and Remarriage Occurred During High School | 0.943 | 0.443 |
| Parent died during High School | 0.727 | 0.303 |
| **Family in 1988** (Young woman in eighth grade) | | |
| Ethnicity - African American (Omitted is Caucasian) | 2.161 * | 0.458 |
| Ethnicity - Latino American | 1.860 * | 0.388 |
| Ethnicity - Asian American | 0.437 * | 0.148 |
| SES*African American interaction term | 0.743 * | 0.116 |
| SES*Latino American interaction term | 1.056 | 0.149 |
| SES*Asian American interaction term | 1.105 | 0.287 |
| Parental Involvement in Education | 0.847 | 0.086 |
| Parents and children are involved in clubs | 1.008 | 0.129 |
| Mother's education (z) | 1.033 | 0.077 |
| Father's education (z) | 0.912 | 0.076 |
| Mother was a teen when this daughter was born | 1.314 * | 0.170 |
| Eighth grade income/needs | 0.995 | 0.067 |
| Father foreign born | 1.208 | 0.258 |
| Mother foreign born | 0.674 * | 0.140 |
| Oldest child | 1.015 | 0.115 |
| Father's occupation {z} | 0.876 * | 0.060 |
| Father unemployed | 1.142 | 0.188 |

## Appendix: Predicted Probability of a Having a Child Out of Wedlock

Conditional Logit Estimates

Fixed effects for each junior high school

(N=3882)

| | Odds Ratio | Standard Error |
|---|---|---|
| Mother's occupation {z} | 0.948 | 0.053 |
| Mother unemployed | 1.137 | 0.121 |
| Religious affiliation - Baptist (Omitted religion is other Protestant) | 1.220 | 0.183 |
| Religious affiliation – Catholic | 1.165 | 0.178 |
| Religious affiliation - Other religion | 1.192 | 0.207 |
| Religious affiliation - Missing religion | 1.121 | 0.288 |
| Religious affiliation - No religion | 1.506 | 0.421 |
| Religiosity - very religious (Omitted religiosity is "Not at all religious") | 0.488 * | 0.062 |
| Religiosity - religious | 0.753 * | 0.110 |
| Religiosity - somewhat religious | 0.770 * | 0.107 |
| Number of siblings | 1.137 * | 0.033 |
| More than 50 books in home | 0.955 | 0.125 |
| Has at least one magazine subscription | 0.776 * | 0.080 |
| Family has a public library card | 1.137 | 0.136 |
| Parents expect the youth to continue in school past high school | 0.929 | 0.126 |
| **Young woman in 1988** (That is, in eighth grade) | | |
| Behavioral problems reported by teacher or parents | 1.393 * | 0.206 |
| Emotional problems | 1.095 | 0.274 |
| Cigarette smoking | 1.537 * | 0.279 |
| Eighth grade test scores (z) | 0.473 * | 0.045 |
| Cuts classes at school | 1.257 | 0.187 |
| Often comes to school without a pen/pencil and paper | 0.757 * | 0.105 |
| Often comes to school without homework | 1.281 * | 0.160 |
| Often comes to school without books | 1.565 * | 0.285 |
| Ever held back in school | 1.409 * | 0.173 |
| Changed schools during elementary school or junior high | 1.121* | 0.057 |

Pseudo-$R^2$ = .212 ***** represents statistically significant at the 5 percent level.

Odds ratio > 1 indicates that variable increases the predicted probability of teen out-of-wedlock birth.