Essentially Heterogeneous: The Consequences of Teen Childbearing on Ecuadorian Mothers and Children

Leonel Borja-Plaza*

October, 2023

Click here for latest version

Abstract

I use deviations from the expected age at menarche to estimate the marginal treatment effects of teen childbearing on schooling and labor outcomes for Ecuadorian mothers and schooling and health outcomes for their firstborn children. Findings suggest that women whose unobservable characteristics make them *less likely* to become teen mothers have fewer years of schooling, are less likely to finish high school and to participate in the labor force. Women whose unobservables characteristics make them *more likely* to become teen mothers do not have their schooling attainment negatively impacted and increase their labor force participation. I do not find evidence of effects on firstborn children. These findings may help reconcile seemingly conflicting evidence from past studies and imply that there is potential to improve women's outcomes by reducing teen childbearing rates when opportunity costs are sufficiently high. However, these findings counter the belief that teen childbearing has been an significant source of intergenerational transmission of low socioeconomic status (*JEL* O12, I21, I25, J13, J16, J24).

*Dyson School of Applied Economics and Management, Cornell University. Email: lab372@cornell.edu. I'm grateful to Chris Barrett, Brian Dillon, John Hoddinnott, as well as seminar participants of the Development Economics Graduate Seminar at Cornell University for helpful comments and suggestions.

1 Introduction

Does teen childbearing cause economic disadvantage that persists across generations? Research across fields shows that teenage motherhood is associated with lower socioeconomic status for mothers and children. Teen mothers are more likely to have less schooling, poorer health, and lower income later in life than their counterparts who delay fertility. Their children are also more likely to experience worse outcomes throughout life. However, these associations do not necessarily imply that teen motherhood *causes* lower socioeconomic status. The reason is that women who become teen mothers tend to come from disadvantaged backgrounds, which makes their prospects less promising regardless of the timing of their first birth. That is, the associations may suffer from selection bias. Researchers have long attempted to discern whether the observed patterns arise from selection or causation, and the question remains of significant policy relevance, especially in low- and middle-income countries, where roughly a third of women become mothers during adolescence (UNFPA, 2022).

Despite much effort devoted to studying the causal effects of teen childbearing on mothers and their children in high-income countries, the evidence has been mixed. Some research supports that the negative link between socioeconomic outcomes and teen childbearing stems from selection bias. Others suggest it is primarily causal. Another set of studies falls in between, finding modest adverse effects. An overlooked explanation for these seemingly conflicting results is that the consequences of teen childbearing significantly differ across women, even after accounting for differences in socioeconomic backgrounds and other relevant pre-motherhood characteristics[1]. In the program evaluation literature, when these differences are driven by individual factors that are unobservable to the researcher, they are called "effect heterogeneity" or "essential heterogeneity[2]." If such heterogeneity is substantial, the variation in the estimated consequences of teen childbearing will be substantial as well. Different methods and samples, at best, generate different local average

[1] The alternative explanation, which has fueled the progress in the literature, is that some studies have done better than others at addressing selection bias. Diaz and Fiel (2016), however, persuasively argue that the literature's concern with selection bias has overshadowed the importance of considering effect heterogeneity.

[2] Throughout the paper, I use the terms essential heterogeneity, unobserved heterogeneity, and effect heterogeneity interchangeably.

treatment effects (LATE), leading to conflicting policy recommendations.

In this paper, I examine the consequences of teen childbearing[3] in the context of a Latin American country, testing for essential heterogeneity. In particular, I use a large sample of Ecuadorian women and firstborn children to estimate teen childbearing's marginal treatment effects (MTE) (Heckman and Vytlacil, 2005) on mothers' schooling and labor force participation and firstborn mortality, schooling, and nutritional outcomes. While minimizing selection bias, the MTE framework allows me to test for essential heterogeneity by characterizing it on a single parameter that encompasses unobservable characteristics that make women less likely to become teen mothers. It also allows me to clarify how effect parameters that are often of interest to researchers, such as the average treatment effect (ATE), the average treatment effect on the treated (ATT), and the average treatment effect on the untreated (ATU), vary in the presence of effect heterogeneity[4].

For identification, I exploit plausible exogenous variation from the difference between the observed age at menarche (AAM)[5] and the expected AAM. I estimate the expected AAM using a random forest algorithm trained on non teen mothers, using proxies of pre-menarche characteristics. While it is well established that menarcheal timing is a strong predictor of teen motherhood status (Ribar, 1994; Klepinger et al., 1999), by using the deviations from the expectation rather than the AAM itself, I better isolate its plausibly exogenous variation, purging it from environmental influences—in particular, socioeconomic background—that may correlate with outcomes. Additionally, I allow the AAM to be part of the outcome's equation. That is, I relax the exclusion restriction to correct any remaining bias that may arise if the AAM affects the outcomes through channels other than teen motherhood.

Results suggest that unobserved heterogeneity is key to understanding maternal outcomes. I find that teen motherhood significantly reduces educational attainment and labor force participation for women whose *unobservable* characteristics make them *less likely* to become teen mothers. In contrast, for women whose *unobservable* characteristics make them *more likely* to become teen

[3]This papers defines teen childbearing as having a child at ages 15 to 19. As I discuss in the section 4, I exclude from the analysis all women who had children at ages fourteen or younger.

[4]Note that if there is no essential heterogeneity, the ATE, the ATT, the ATU and any LATE are all equal.

[5]Menarche refers to the first ocurrence of menstruation.

mothers, educational attainment and labor force participation are less likely to be affected. Indeed, I find that teen childbearing does not affect educational attainment and has a *positive* effect on labor force participation for women whose unobservables make them the most likely to become teen mothers. This heterogeneity affects how we should think about the average consequences of teen childbearing and what parameters may (or may not be) policy-relevant. I present local versions of the ATE, ATT, and ATU to illustrate the consequences of essential heterogeneity[6].

In my preferred specification, women in the sample who delayed fertility to adulthood or had no children would have been 30% less likely to finish high school and would have lost 1.9 years of schooling, on average, in the counterfactual case where they had a child as teenagers—these parameters are the (local) ATUs. In sharp contrast, the average effects on the probability of high school completion and years of schooling for those who gave birth as teenagers–i.e., the (local) ATTs—are *positive* but not statistically different from zero. Consequently, the overall average effect of teen childbearing in the sample—i.e., the (local) ATEs—is between the two parameters, suggesting that teen motherhood reduces the probability of high school completion by 15% and schooling attainment by one year, on average. Overall, the marginal effects on high school completion range from +0.3 to -0.3 and from +1.46 to -2.07 for years of schooling. I find similar patterns when analyzing labor force participation.

These results should be interpreted with some caution. While the ATT, often considered a policy-relevant parameter, implies that delaying motherhood would not have led to improved outcomes *on average* for teen mothers in the sample, it does not imply that policy has no role. Indeed, 75% of the women have a corresponding negative marginal treatment effect regarding their schooling attainment. And around half of the women have a negative marginal treatment effect on their labor force participation. Nonetheless, these results do imply that policies to reduce childbearing may not improve women's outcomes in settings where opportunity costs for teen childbearing are not high enough.

[6] As discussed in section 5, these are local parameters because the propensity score of teen childbearing estimated using deviations from the expected age at menarche (the first step when estimating MTE) does not contain the full unit interval, which is necessary to calculate the ATE, ATU, and ATT. Instead, I follow Carneiro et al. (2011) and calculate local versions of these parameters for observations within the support of the propensity score.

For firstborn children, I fail to find significant detrimental impacts on measures of mortality, relative years of schooling, and heigh-for-age Z scores, suggesting that the negative associations in OLS estimates are mostly driven by selection bias. I also don't find evidence to indicate that essential heterogeneity is a concern when analyzing these outcomes. However, I find that teen mothers' firstborns, fourteen or older, are less likely to live in the same household as their mothers. In light of recent evidence suggesting that Norwegian children of teen mothers have worse medium- and long-term outcomes than their adult-born cousins (Aizer et al., 2022), the lack of significant impacts in this paper may suggest that negative effects of teen childbearing on children may be hard to detect in the short-term. They could also reflect differences in how families from different cultural backgrounds respond to a teenager having a child. The mechanism through which mothers might be negatively affected but not their firstborn warrants further research.

This paper contributes to our understanding of the consequences of early childbearing in three ways. First, it shows that when studying the societal implications of teen childbearing, women's unobserved characteristics loom large, even if they are often overlooked. This paper provides rigorous evidence suggesting that unobserved heterogeneity is a viable explanation as to why many studies find that teen childbearing significantly reduces women's educational attainment (Klepinger et al., 1999; Levine and Painter, 2003; Holmlund, 2005; Fletcher and Wolfe, 2009; Lang and Weinstein, 2015; Herrera-Almanza and Sahn, 2018), while others find no significant effects (Geronimus and Korenman, 1992; Ribar, 1994; Hotz et al., 1997) or only modest adverse effects (Hotz et al., 2005; Ashcraft et al., 2013)[7]. This paper makes a similar argument to that posed by Diaz and Fiel (2016), but uses a stronger identification strategy that allows to more credibly minimize selection bias. In addition, the MTE framework clarifies how to recover and understand the ATE, ATT, and ATU in the presence of effect heterogeneity.

Second, by studying mothers and their firstborn children, this study offers insights that challenge the common belief that teen childbearing has been an important driver of intergenerational transmission of low socioeconomic status. In particular, the findings suggest that such a belief might be

[7]There is also evidence of positive effects (Ribar, 1994; Azevedo et al., 2012).

overstated. Women more prone to becoming teen mothers are less affected by it. And even when teen mothers are adversely impacted, their children don't present worse short-term outcomes than comparable children born to adult mothers.

Third, this study contributes new findings from Ecuador, a middle-income country with high teenage pregnancy rates, illegal abortion, and low contraception use. Ecuador has some of the highest teen childbearing rates in Latin America and the Caribbean, a region whose rates are only surpassed by those in Sub-Saharan Africa (United Nations Population Division, 2022). Latin America has also exhibited the slowest decline in rates since 1990 (Santelli et al., 2017). These concerning trends have put teen pregnancy back on the agenda for policymakers, making this paper a timely contribution. Additionally, the paper adds to a handful that study low- and middle-income countries, where teen childbearing is a widespread issue (Herrera-Almanza and Sahn, 2018; Branson and Byker, 2018).

This paper also speaks to the growing literature that estimates marginal treatment effects in contexts where essential heterogeneity is important. These contexts include returns to education (Carneiro et al., 2011; Kaufmann, 2014), fertility and female labor supply (Liu et al., 2022), returns to early child care (Cornelissen et al., 2018; Felfe and Lalive, 2018), effects of foster care (Doyle Jr, 2007), among others. The present study is the first to estimate the marginal treatment effects of teen childbearing, showing that essential heterogeneity is particularly important for this topic.

The rest of the paper is organized as follows. In section 2, I reviewing the literature on the consequences of teen childbearing. In section 3, I present the identification framework. In section 4, I describe the data, the construction of the instrument, and the method followed to relax the exclusion restriction. In section 5, I present the results, and in section 6, I conclude.

2 Background

2.1 Previous Works on the Consequences of Teen Childbearing

The amount of work in the social sciences studying the causal effect of teen pregnancy on maternal and child outcomes is too large for this section to do it justice. Instead, I review some of the most influential, relatively recent studies, mainly from the economics literature, to make three points. First, the evidence on the effect of teen childbearing on educational attainment and other socioeconomic outcomes is mixed. Second, concerns about selection bias have been the primary drivers of the literature's evolution. Third, rigorous causal analysis has primarily drawn samples from high-income countries, despite the higher prevalence of teen childbearing in low- and middle-income countries. Appendix Table A1 offers a summary of this literature review.

2.1.1 Evidence from high-income countries

Researchers have advanced the literature by studying new and larger samples, using data with richer covariates, and proposing new ways to deal with selection bias. The three most widely used strategies include 'within family' estimators (family fixed effects), propensity score methods, and instrumental variables/ natural experiments.

Family Fixed effects (FFE). Comparing siblings, when one sibling had a child as a teen while the other delayed fertility, is one of the most compelling ways researchers have dealt with selection bias. The advantage of this method is that comparing sisters is thought to control for what is often regarded as the primary source of bias in cross-sectional estimates: unobserved family characteristics. Still, the evidence from studies using this method is mixed.

For instance, Geronimus and Korenman (1992) used samples from three different surveys to find that in two of them, sisters' who had a child as a teen had non-significant differences in income and educational attainment measures compared to a sister that delayed her childbearing to adulthood. Interestingly, they find little difference between cross-sectional estimates,which suggest dire adverse effects, and their within-family comparisons in their third sample. Holmlund (2005), using samples from Sweden, finds that within-family estimators overstate any adverse impacts of teen childbearing if they do not control for within-family heterogeneity. However, the author reports that teen childbearing reduces 0.59 years of attained schooling, even after accounting for pre-motherhood differences among sisters.

FFE is the leading method economists have used to study the impacts of teen childbearing on children. These studies compare the outcomes of cousins when one mother had the child as a teenager while her sister delayed her first birth. Geronimus, Koreman, and Hillemeier (1994) use a US sample to examine teen motherhood's effect on multiple child development assessment scores. They find little evidence of adverse effects; although not significant, the within-family estimates favor teen mothers in most outcomes. López Turley (2003) doubles the sample size of Geronimus, Koreman, and Hillemeier and still finds that maternal age is not a significant predictor of children's development assessment scores. Similarly, Rosenzwein and Wolpin (1995) find positive but not statistically significant differences in children's birth weights. However, they find that teen mothers have shorter gestational periods than adult mothers.

In contrast, in a recent and comprehensive study, Aizer, Devereux, and Salvanes (2022) document that much but not all of the adverse effects on teen mothers' children are explained by selection. The authors point out that prior research using within-family estimators shares one limitation: a small sample size. They use population-wide data from the Norwegian register, allowing them to have a much larger sample size and examine short, medium, and long-term outcomes. They report that children of teen mothers have lower educational attainment, IQ scores, and earnings than their adult-born cousins.

Even assuming constant treatment effects, within-family estimators have a few limitations. First, unobserved within-family heterogeneity may cause FFE estimates to be biased. Second, it is often difficult to assess the policy relevance of FFE estimates when the subsample of families with at least two sisters, only one of which was a teen mother, systematically differs from the overall population of families. Third, the stable unit treatment values assumption (SUTVA) is likely violated. This assumption is violated when a teen pregnancy event affects the probability of teen pregnancy of other individuals included in the analysis or their outcomes. FFE likely violates SUTVA in more than one way. First, suppose a sibling who is not a teen mother is adversely affected by her sister giving birth as a teenager. In that case, FFE estimates again underestimate the impact of teen childbearing on women and possibly their children too. Indeed, recent evidence suggests that younger sisters might be negatively affected when older sisters become pregnant as teenagers (Heissel, 2021). Second, there is evidence of peer effects in teen childbearing (Yakusheva and Fletcher, 2015), so it is possible that a women getting pregnant as a teen affects the likelihood of her sister getting an early pregnancy, leading to biases that are hard to sign.

Propensity Score Matching (PSM) and Inverse Probability Weighting (IPW).

These studies usually conclude that cross-sectional estimates only slightly overstate negative impacts. Their advantages are using larger sample sizes and not relying on particular population subsamples. A prime example of such studies is Levine and Painter (2003), who use a within-school PSM technique in a US sample to find that teen mothers have 22.1 percentage points higher likelihood of dropping out of school. Another example is Diaz and Fiel (2016), which is also the only study (this author is aware of) that explicitly models treatment effect heterogeneity. They use smoothing-differencing (S-D) to graphically explore effect heterogeneity and IPW to estimate the ATE, ATT, and ATU in a sample from the US. Their S-D results suggest that teens most likely to experience a pregnancy experience smaller adverse consequences regarding high school completion, which is consistent with the findings of this study. In contrast, their findings using IPW suggest that the ATE, ATT, and ATU are similar, suggesting no treatment heterogeneity.

The drawback of PSM, IPW, and similar methods is their reliance on conditional independence, an assumption too strong to make when there is reason to believe that unobserved heterogeneity may lead to selection into treatment.

Instrumental Variables (IVs) and Natural Experiments. Like FFE studies, IV studies have produced mixed evidence. For instance, Ribar (1994) and Klepinger, Lundberg and Plotnick (1999) study samples from the National Longitudinal Survey of Youth and use the age of menarche, among other instrumental variables, to reach different conclusions. The former finds slightly positive, non-statistically significant effects on the propensity of high school completion. The latter finds that teen mothers lose more than 2.5 years of schooling.

Another set of studies compares teen mothers with women who got pregnant as teenagers but suffered a miscarriage, using ordinary least squares or IV estimation strategies. These studies exploit the plausibly (conditional) exogeneity of miscarriages. Hotz, Williams and Sanders (2005) use this technique and generally document minor non-significant adverse effects on educational attainment and positive effects on cumulative work hours and earnings. Fletcher and Wolfe (2009) worry that community-level characteristics correlate with the probability of a miscarriage, so they include community-level fixed effects. They find instead statistically significant adverse results in the probability of receiving a high school diploma (-9%) and annual income. Ashcraft, Fernández-Val and Lang (2013) also carefully consider the possibility that miscarriage is not socially random in the presence of abortions. The authors derive a consistent estimator for the effects of teen pregnancy, consisting of a weighted average of the OLS and IV estimates. They find no effects of teen childbearing on the probability of having a high school diploma and only modest adverse effects on years of schooling.

IV studies have moved forward by trying to resolve biases caused by failures of the exclusion restriction: that the IV should only affect the outcome only through its effect on the treatment variable. Treatment effect heterogeneity, on the other hand, has rarely been discussed, despite that in its presence, fully saturated models estimate at best estimate a LATE, which may or may not be policy-relevant.

2.1.2 Evidence from low and middle-income countries

Only a few studies rigorously address selection bias focusing on low- and middle-income countries. Azevedo, López-Calva and Perova (2012) compared teen mothers with women who reported a miscarriage during adolescence. They find that teen mothers have 0.31 extra years of schooling and are more likely to be employed. In contrast, Herrera-Almanza and Sahn (2018) use community access to contraceptives as an IV in Madagascar and find that teen motherhood reduces women's likelihood of completing (the equivalent of) middle school by more than forty percent.

Rigorous causal evidence from low and middle-income countries on children's outcomes is even more limited. One notable exception is Branson and Byker (2018), who study both mothers and children. In a differences-in-difference framework, they examine the effects of a program that increased reproductive health knowledge and clinical access for South African teens. They find that adolescents impacted by the policy reduced their likelihood of becoming teen mothers by 11%, completed one more year of schooling, and earned 30% higher wages. They also show that the firstborn children of teens impacted by the policy reduced their likelihood of being stunted by 15% and increased the height-for-age scores by 0.8 standard deviations.

2.2 Teen childbearing in Ecuador

According to World Population Prospects' data, over 13 million girls between 15 and 19 gave birth in 2021 (United Nations Population Division, 2022). Almost all those births (96%) occurred in low- and middle-income countries, where adolescent childbearing rates are much higher. The adolescent childbearing rate in less economically developed countries is around 46 births per 1,000 women, or four times higher than the rate in more developed countries, which is estimated to be less than 12. There are also significant disparities across regions. Latin America and the Caribbean (LAC) had a rate of 53.2 births per 1,000 women, second only to sub-Saharan Africa (SSA), with the highest global rate at 101. Within Latin America, there are also considerable disparities among countries. For example, Nicaragua reports the highest rate, with 85.6 births per 1,000 women, while Chile, one of the wealthiest countries in the region, reports a significantly lower rate of 24.1. Still, even low rates in LAC are higher than the highest rates observed in North America (the US rate is 16) and Europe (the UK rate is 11). Furthermore, while all regions worldwide have experienced declining teen childbearing rates, the LAC region has had the slowest decline.

In Ecuador, the setting of this study, the number of births per 1000 women is estimated to be 63, ranking 11 out of 50 in the LAC region and within the top 25% of highest rates across countries globally, but significantly lower than the rates of over 100 observed in SSA. While knowledge of contraceptive methods is widespread, in 2018, only 43% of Ecuadorian women aged 12 to 24 declared to have used any contraceptive in their first sexual relationship. 41% of women aged 20 to 24 reported having been in a marriage or a de facto marriage before they were 18. Moreover, until 2021, women and girls could only seek a legal abortion when their pregnancy endangered their life or in the case of rape of a woman or girl with severe mental disabilities.

The rates of teen pregnancies are a cause of worry for policymakers in Ecuador. The Ecuadorian Ministry of Education data documented that 6,847 adolescents left school due to pregnancy in 2015. Also, a report commissioned by the United Nations Population Fund estimated that the cost of teen childbearing due to productivity losses and the cost of health services was close to 0.26% of GDP in 2017. This report, however, assumes causation to estimate adolescent childbearing's costs.

3 Identification Framework

The analysis of unobserved heterogeneity and the identification of MTE follows from a latent index model (Heckman and Vytlacil, 2005, 2007; Carneiro et al., 2011). Consider potential outcomes (Y_1, Y_0) linearly projected on a vector of observed covariates X, where Y_1 is an outcome for a woman/child if the child was born when the mother was between 15 and 19, and Y_0 is the outcome for the same woman/child if the child was born when the mother was not a teenager or woman had no children:

$$Y_0 = \beta_0' X + \epsilon_0 \tag{1}$$

$$Y_1 = \beta_1' X + \epsilon_1 \tag{2}$$

by definition ϵ_j is normalized to $E[e_i|X = x] = 0$, for $j = 0, 1$ (Brinch et al., 2017).

For a child or a mother, either Y_1 or Y_0 are realized so the observed outcome can be written as:

$$Y = Y_0 + (Y_1 - Y_0) \times Teen. \tag{3}$$

where $Teen$ is a dummy variable equal to one if the women or the child's mother is observed to be a teen mother. The effect of teen childbearing on Y can be written as

$$Y_1 - Y_0 = (\beta_1 - \beta_0)' X + \epsilon_1 - \epsilon_0 \tag{4}$$

The effect has two components: an observable component: $(\beta_1 - \beta_0)' X$, and an unobservable component, $\epsilon_1 - \epsilon_0$.

Selection into teen motherhood is modeled using a latent index model:

$$Teen^* = f(Z) - V, \tag{5}$$

$$Teen = 1 \quad \text{if} \quad Teen^* \geq 0, \;\; Teen = 0 \text{ otherwise}, \tag{6}$$

where $Z = (X, \tilde{Z})$. That is, Z is a vector that contains observables X, plus an instrumental variable $\tilde{Z}$. $f(Z)$ is an unspecified function that maps vector Z into a single number. V is a random variable with continuous distribution function $F_{V|X}(V)$. Since V enters equation 5 with a negative sign, it encompasses all unobservable characteristics that make teen motherhood *less likely*.

Defining the quantiles of V as a new variable $F_{V|X}(V) = U_{teen}$, normalizes U_{teen} to be continuous variable that is uniformly distributed[8], and $f(Z)$ to be the propensity score $P(Z)$[9].

The marginal treatment effect (MTE) is defined across observables and quantiles of unobservables U_{teen}:

$$MTE(x, u) = E[Y_1 - Y_0 | X = x, U_{teen} = u] \tag{7}$$

and it can be estimated estimated using local instrumental variables (LIV) (Heckman and Vytlacil, 1999, 2005):

$$\frac{\delta E(Y|X = x, P(Z) = p)}{\delta p} = MTE(x, p) \tag{8}$$

[8]This is the probability integral transform.

[9]$F_{V|X}(f(Z)) = P(V \leq f(Z)) = Pr(Teen = 1|Z) = P(Z)$

Note that marginal increases in $P(Z)$, starting at high values of $P(Z)$, identify the effects for women with high values U_{teen}, because women with low values of U_{teen} would already be teen mothers. Therefore, marginal changes on $P(Z)$ allow the recovery of effects at all margins U_{teen} within the empirical support of $P(Z)$ (Carneiro et al., 2011).

As is standard in the literature, I assume that X conditional on U_{teen} is independent of the error terms in the potential outcome equations: $(\epsilon_1, \epsilon_0) \perp\!\!\!\perp X|U_{teen}$. This assumption implies that $E[\epsilon_1 - \epsilon_0|X = x, U_{teen} = u] = E[\epsilon_1 - \epsilon_0|U_{teen} = u]$, allowing for the marginal treatment effect to be additively separable between the observed and unobserved components.

I take the expectation of Y conditional on $X = x$ and $U_{teen} = p$,

$$E[Y|X = x, P(Z) = p] = \beta_0^{'}x + (\beta_1 - \beta_0)^{'}x \times p + E[e_1 - e_0|U_{teen} = p] \times p, \tag{9}$$

where p is a specific value of the propensity score P(Z). Note that $E[e_1 - e_0|U_{teen} = p] \times p$ is only a function of the propensity score so it can be written as $K(p)$.

The marginal treatment effect function can be written as:

$$\frac{\delta E(Y|X = x, P(Z) = p)}{\delta p} = (\beta_1 - \beta_0)^{'}x + \frac{\delta K(p)}{\delta p} = MTE(x, p), \tag{10}$$

where $(\beta_1 - \beta_0)^{'}x$ is the intercept of the MTE function and depends on observed characteristics. $\frac{\delta K(p)}{\delta p}$ is the slope of the MTE function and only depends on the propensity score. Equation 10 can be recovered using semiparametric techniques or making functional form assumptions about $K(p)$ (Heckman et al., 2013).

Several common parameters of interest such as the average treatment effect (ATE), the average treatment effect on the treated (ATT), the average treatment effect on the untreated (ATU), or the LATE can be generated as different weighted averages of the MTE (Heckman and Vytlacil, 2005; Cornelissen et al., 2016). If unobserved characteristics to teen motherhood are not an important source of heterogeneity (for instance if $e_1 = e_0$) then all treatment parameters are the same, they are identifiable via Instrumental Variable (IV) methods with any valid instrument, and the MTE function is a flat line.

3.1 Estimation

I first estimate the selection equation $P(Z)$ using a probit model. Then, for my baseline results, I approximate the function $\frac{\delta K(p)}{\delta p}$ using a second-degree polynomial on p. I show that the results do not meaningfully change when approximating the function with higher-order polynomials or using a semiparametric specification. The estimating equation is the following:

$$Y = \beta_0' x + (\beta_1 - \beta_0)' x \cdot p + \sum_{l=1}^{2} \pi_l \frac{(p^l - 1) \cdot p}{l+1} \tag{11}$$

where the last term, $K(p)$, is modeled as a third degree polynomial[10].

The MTE is recovered as the derivative of equation 11:

$$MTE(x, p) = (\beta_1 - \beta_0)' x + \sum_{l=1}^{2} \pi_l (p^l - \frac{1}{l+1}) \tag{12}$$

4 Data and Estimation Issues

4.1 Data

The data for this study comes from the 2018 ENSANUT (Ecuadorian National Survey of Health and Nutrition), carried out by the Ecuadorian Institute of Statistics (INEC, 2018). This survey was designed to evaluate the health and nutritional status of the population, with over 40,000 households participating in face-to-face interviews. A majority (85%) of these interviews took place between November 2018 and January 2019, with the remaining 15% conducted from June to July 2019[11]. The survey's scope is both national and provincial, targeting all women aged 10 to 49 in each household. These interviews gathered detailed information on sexual and maternal health, onset of

[10]For this estimation, I use the default polynomial form of the Stata package *mtefe* (Andresen, 2018)

[11]The interval in data collection is likely due to budget constraints, but this should not impact the analysis of this paper, which primarily uses information from women's birth histories.

menstruation, birth histories, and child health. Additionally, the survey collected data on household conditions, assets, education, employment status, and income.

My analysis uses three overlapping subsamples of the ENSANUT data. The first, or base sample, includes 39,629 women aged 15 and above, all with complete birth histories and anthropometric data. This subset is used to determine the likelihood of teenage births using a random forest algorithm. Women who gave birth at age 14 or younger are excluded, as these cases often result from sexual violence (Casas Isaza et al., 2015). After eliminating 61 instances where the onset of menstruation was not reported, I use a second random forest algorithm to estimate the expected age at menarche using data from non-teenage mothers and women without children. Further details of these calculations are provided in subsequent subsections and Appendix B.

The second subsample consists of 24,452 women aged 25 and above. I use this subsample to study maternal outcomes. The analysis is limited to women over 25 to concentrate on those likely to have completed their education. Following the approach of Fields and Ambrus (2008), only women with menarche ages between 10 and 16 are included, as ages outside this range can indicate medical issues. The third subsample, used for examining children's outcomes, comprises all mothers from the first subsample, including those under 25. This group provides data on 27,957 firstborn children (98.33% of whom are alive) and their mothers. However, the number of observations varies depending on the specific outcome being studied. The main difference between the samples is that the women's sample excludes pregnant women or those under 25, while the firstborn sample covers all biological mothers aged 15 to 49. In the latter, mother-firstborn pairs are only included if the firstborn is alive and residing with the mother at the time of the survey, except when studying mortality. Summary statistics for these three subsamples are shown in Table 1.

4.1.1 Outcome Variables

For the women's analysis, I study high school completion, years of schooling, and labor market participation. Although data on earnings are available, studying this outcome would require an additional source of exogenous variation to address the endogeneity of labor market participation.

Without such variation, I do not analyze earnings in this paper.

The analysis of first born children encompasses three primary outcomes. The first is mortality, specifically whether the child was alive at the time of the survey. This outcome includes all firstborn children. Then, as an intermediate step, I check whether teen motherhood is causally associated with children not living in the surveyed household. As I discuss in the appendix section C, I find that children are less likely to be observed if they were born to a teen mother. To mitigate the risk of non-random selection influencing the Marginal Treatment Effect (MTE) estimates, I limit the analysis for subsequent outcomes to younger children. This restriction ensures that the probability of a child living with their mother is not influenced by the mother's teen status. The second outcome is relative years of schooling, calculated as the number of years a child has completed in school against the expected number based on their age[12]. This analysis focuses on children between 6 to 14 years of age. The third outcome examines height-for-age z-scores for children under 14, using the *zanthro* package in Stata (Vidmar et al., 2004) for calculation.

4.2 Addressing Endogeneity using variation from Menarcheal Timing

As shown in previous studies and can be directly tested, the AAM is a relevant instrument, highly predictive of teen mother status (Ribar, 1994; Klepinger et al., 1999). Its strong-predictive power comes from the correlation between early pubertal maturation and both earlier sexual debuts and riskier sexual behaviors at all socioeconomic levels (Benson and Torpy, 1995; Glynn et al., 2010; Cheong et al., 2015; Baams et al., 2015). The extended duration between the onset of sexual activity (affected by pubertal maturation) and the transition into adulthood mechanically increases the probability of teenage pregnancy.

In addition, some have argued that the AAM may be as good as randomly assigned. For instance, Field and Ambrus (2008), who used the AAM to instrument for early marriage in Bangladesh, argued that the high heritability of the AAM indicates "a high degree of genetic determinism and

[12] In Ecuador, the education system includes two mandatory stages: "basic general education" (ten years) and "unified general bachillerato" (three years), summing up to 13 years for high school completion. This structure is comparable to the U.S. system in terms of duration, with the distinction that kindergarten is considered the first grade in Ecuador.

hence a minimal role of environmental influences on maturation," citing genetics research (Kaprio et al., 1995). More contemporary studies looking at the correlation between AAMs among family members and between monozygotic twins as opposed to dizygotic twins still suggest that genetic factors heavily influence variation in menarcheal timing (Towne et al., 2005; Anderson et al., 2007; Morris et al., 2011). Overall, the evidence suggests that genetic determinants account for 50 to 80% of the variation in menarcheal age (Gajdos et al., 2009). In this paper, I exploit the highly complex polygenic nature of the AAM, which among homogenous populations, produces quasi-random variation uncorrelated with family socioeconomic backgrounds and other confounders of teen childbearing.

However, contemporary medical and genetics research also emphasizes the non-negligible role of environmental and gene-by-environment influences affecting the AAM. There is evidence of associations between the timing of menarche and nutritional status (Jansen et al., 2015; Villamor et al., 2017), socioeconomic conditions (Chavarro et al., 2004; Wronka and Pawlińska-Chmara, 2005; Jansen et al., 2015; Marván et al., 2020), psychosocial stress (Romans et al., 2003; Tither and Ellis, 2008), and exposure to endocrine-disrupting chemicals (Parent et al., 2005; Buttke et al., 2012). An observed secular decline of the average AAM also points to environmental effects that still need to be better understood (Parent et al., 2003; Marván et al., 2020). If environmental elements independently impact the onset of puberty and adolescent childbearing, employing AAM as an instrumental variable may result in biased estimates. For this study, it is fortunate to note that despite changes in environments, improvements in living conditions, and secular trends, a significant individual variation in the physiological timing of menarche, ranging between 4-5 years, is consistently observed (Parent et al., 2005; Gajdos et al., 2009). In the following subsection, I detail the instrument's construction, which intends to isolate the as-good-as-random variation of the AAM.

4.3 The Instrument: the difference between expected and observed AAM

Let the AAM be the sum of two components: the expected AAM, a function of observables Or_i, which reflects the average age at menarche for a woman born in a particular year, month, region, and with a specific set of socioeconomic conditions, plus an exogenous $\tilde{Z}$ that captures the quasi-random deviations from the expectation that stems from normal genetic processes:

$$AAM_i = E[AAM(Or_i)] + \tilde{Z}_i \tag{13}$$

Then, simply subtracting the observed AAM from the expected AAM isolates the quasi-exogenous component:

$$\tilde{Z}_i = AAM_i - E[AAM(Or_i)] \tag{14}$$

This idea relates somewhat to a recent paper by Borusyak and Hull (2021), who refers to a similar procedure as "recentering" an IV.

Note, however, that the object $E[AAM(O_i)]$ first needs to be estimated. Ideally, the vector Or_i would include pre-menarche measures of family socioeconomic and nutrition status, geographical information on where the individual i grew up, ethnicity, and birth year. Unfortunately, the survey I use does not include retrospective information about socioeconomic status.

To circumvent the lack of retrospective information, I modify equation 14 to depend on current information:

$$\tilde{Z}_i = AAM_i - E[AAM(Op_i)|Teen = 0] \tag{15}$$

where $E[AAM(Opi)|Teen = 0]$ is the expected AAM for individual i given post-menarche observables Op in the counterfactual case woman i did not become a teen mother. In practice, I calculate the expected AAM without assuming a functional form by estimating a random forest model while excluding all teen mothers. This exclusion allows me to use variables that could be affected by teen motherhood. In particular, I use information on whether the woman lives in a rural area, the region where the woman lives (as a proxy for where the woman grew up), the year of birth and month of birth, her race/ethnicity, height (as a proxy for pre-menarche nutritional status), and a series of variables that reflect asset ownership, access to public services, and household conditions. In addition, I include the probability of teen motherhood, also calculated with a random forest that uses similar inputs but excludes the variables more likely affected by teen motherhood: asset ownership, access to services, and household conditions. I train the model on (a randomly selected) 70% of 27,466 non-teen mothers and leave the 30% to test predictions and ensure the model does not overfit the data. More details of these calculations, including all variables used, are in appendix B.

After training the model, I predict the AAM for teen and non-teen mothers. Note that because some of the variables in the vector Op may be affected by teen motherhood, the estimation of the expected AAM might be biased for teen mothers. This bias is unlikely to be a cause of concern for the purpose of this paper, however, because of the inverse relationship between socioeconomic status and AAM in the context I study. Specifically, if the bias in the prediction were non-negligible, the instrument could lose its relevance[13]. Nonetheless, the result section shows that the instrument strongly predicts teen motherhood status.

[13]To see this, consider an example: assume that teen motherhood adversely affects asset ownership, household conditions, and access to public services for an individual j who was a teen mother with $AAM_j < E[AAM(Or_j)]$ (that is, with menarche lower than the average given her pre-menarche observables). For this example, also assume that the vector Op contains only one variable that reflects socioeconomic status. Since I only observe post-menarche observables Op_j and socioeconomic conditions are negatively correlated with AAM, the prediction of the expected AAM for woman j, i.e., $E[AAM(\hat{Op_j})|Teen = 0]$, will be upwardly biased; the reason for this bias is that non-teen mothers with similar post-menarche characteristics $Op_i = Op_j$ likely had worse socioeconomic conditions pre-menarche ($Or_i < Op_j$). Thus, the expected AAM calculated with pre-menarche observables would be higher for non-teen mothers $E[AAM(Or_i)] > E[AAM(Or_j)]$, even though the prediction using post-menarche observables (from a model trained excluding teen mothers) is the same $E[AA\hat{M}(Op_i)] = E[AA\hat{M}(Op_j)]$. This difference in estimated expected values implies that Z calculated using post-menarche observables is likely closer to zero or even of the opposite sign than that calculated with pre-menarche observables. So if the bias on the prediction were to be of importance, Z could lose relevance. That is, it would no longer be correlated with teen motherhood status. I attempt to minimize the bias by including in the prediction of AAM women who have never been pregnant and by including as an input in the model the probability of teenage motherhood given only by plausibly exogenous variables. Note that the bias is also lower the greater the number of women who, regardless of their timing of menarche, would not have been teen mothers (the group commonly referred to as "never-takers" in the IV literature). Note that another option would be to include teen mothers in the estimation. This inclusion would also produce biased estimates of the expectation. However, in that case, the resulting bias would imply that some of the variation of the AAM associated with socioeconomic status would be assigned to instrument $\tilde{Z}$, which would be a cause of concern. Another option would be only to include variables less likely affected by teen motherhood when estimating the expected AAM. However, the procedure may then be less successful in adequately accounting for differences in socioeconomic status, which previous research has found to be a critical confounder of teen childbearing.

Figure 1 presents the distributions of the age at menarche as it is observed, the AAM as it is predicted, and the the differences between the two. Figure 2 and figure 3 illustrate how instrument $\tilde{Z}$ helps isolate the "good" variation of the AAM. Figure 2, panel A shows the association between the AAM and the probability of a teen birth after adding relevant controls. As expected, higher AAM is associated with a lower probability of teen birth. Panel B shows that instrument $\tilde{Z}$ maintains the expected association: a larger $\tilde{Z}$ is associated with a lower probability of teen birth. Figure 3 shows why $\tilde{Z}$ is more likely to satisfy the exogeneity assumption than the AAM. In both Panel A and B, the graph includes only non-teen mothers and it plots an asset index on the y-axis as a measure of current socioeconomic status. A higher value of this index indicates more access to services, better household conditions, and higher asset ownership. The regression line in panel A shows that the AAM is inversely related to current socioeconomic status. In contrast, the regression line of panel B shows that the difference between the observed and expected AAM is mostly uncorrelated with the asset index. Thus, while being similarly relevant as a predictor of teen motherhood as the AAM, deviations from the expected AAM are arguably exogenous to socioeconomic conditions that might confound its association with teen motherhood.

4.4 Relaxing the exclusion restriction

While the constructed instrument isolates some of AAM's “good" variation, there is no guarantee that it does so entirely. Moreover, even full exogeneity does not necessarily imply compliance with the exclusion restriction, especially for some children's outcomes. The reason is that the AAM may share some genetic bases with other traits. For instance, Elks et al. (2010) found evidence suggesting that several BMI and height-related genes were also associated with AAM. Therefore, since children and mothers share genetic material, it is possible that some sources of exogenous variation influencing maternal AAM also influence some children's traits, such as their height, through channels unrelated to teen motherhood.

Additionally, the AAM, which marks the beginning of women's reproductive life, may influence some women's outcomes beyond its influence on teen motherhood status. For example, the same mechanism that influences the probability of teen childbearing may influence early adult motherhood.

To deal with potential failures of the exclusion restriction in both maternal and children's outcomes, I allow the AAM to have a direct effect on outcome Y. Equation 9 becomes:

$$\tilde{Y} = \beta_0' x + \gamma \times \text{AAM} + (\beta_1 - \beta_0)' x \cdot p + K(p) \tag{16}$$

As described in section 3, $K(p)$ is equal to $E[e_1 - e_0 | U_{teen} = p] \times p$. γ is the direct effect of the AAM on the outcome Y, assumed for simplicity to be the same for all individuals. Note that the exclusion restriction amounts to setting $\gamma = 0$. Let $Y = \tilde{Y} - \gamma \times \text{AAM}$, then equation we again obtain equation 9:

$$Y = \beta_0' x + (\beta_1 - \beta_0)' xp + K(p),$$

so the MTEs are similarly defined. Nevertheless, to account for the direct effect of the AAM on outcome $\tilde{Y}$, it is first necessary to get an estimate of γ. Assuming constant effects that do not vary by teen motherhood status allows me to write the potential outcome equation as follows:

$$\tilde{Y} = \beta_0' x + \gamma \times \text{AAM} + e_0 \tag{17}$$

Therefore, estimating γ using only information on non-teen mothers is possible. Ideally, this estimation would be done in a group unaffected by the instrument because, for some women, their high AAM is the reason they did not become teen mothers (that is the a part of the group commonly referred to as "compliers" in the IV literature). For these individuals, some of the influence of the AAM on outcome $\tilde{Y}$ works through their teen motherhood status (or rather, their non-teen mother status). In practice, I attempt to minimize the presence of "compliers" by estimating equation 17 on a sample of non-teen mothers whose probability [14] of teen motherhood given plausibly exogenous observables is lower than 50% [15]. I estimate γ and adjust each outcome prior to estimating the

[14]This is the same probability that was calculated using a random forest model and is described in the previous section and appendix B.

[15]Note that the assumption of constant direct effects of the AAM allows me to recover gamma using only a subsample

MTE.

5 Results

5.1 Selection Equation

Table 2 reports the average marginal effects evaluated at covariate means for various selection-into-teen-motherhood models for the mothers' and firstborns' samples. The dependent variable is equal to one if the woman had a birth at ages 15 to 19 and zero otherwise in the mothers' sample, and is equal to one if the mother was a teen when she gave birth to her firstborn and zero otherwise in the firstborns' sample. Columns (1) to (3) contain the estimates for the mothers' sample. Column (1) includes only the excluded instrument— the difference between the observed AAM and the expected AAM— while controlling for the expected AAM. Using the expected AAM as a control is not necessary, but it will help remove some residual variation of the outcomes in the following estimations (Borusyak and Hull, 2021)[16]. Column (2) omits the expected AAM as a control and instead includes variables that proxy pre-menarche characteristics: whether the individual resides in a rural area, region of residence, self-reported ethnicity, dummies for a total of eight height quantiles, and year and month of birth dummies[17]. My preferred estimates are from a more flexible model and are presented in column (3). It uses as excluded instruments the difference between the observed AAM and the expected AAM and its square, and as controls the expected AAM and its square, as well as polynomials of the propensity of teen motherhood $P(teen|Or)$ obtained via a random forest model using the proxies for the abovementioned pre-menarche characteristics as described above and appendix B. It also includes all proxies of pre-menarche characteristics.

of adult mothers.

[16]Note that adjusting the model using the expected AAM as a control makes the coefficient of the excluded instrument equivalent to that of the observed AAM obtained if it was used as an instrument while also controlling for the expected AAM.

[17]Current residence in a rural area is taken as a proxy for whether the individual grew up in a rural area. I use region of residence instead of more disaggregated information, such as the province of residence, because while migration to more populated cities is common, migration between regions is less likely. Finally, I used a set of 8 dummy variables reflecting eight adult height quantiles as a proxy early-life environment (Currie and Vogl, 2013). I use dummies for quantiles instead of a continuous measure of height to minimize the possibility of bias stemming from the direct relationship between the onset of puberty and height.

Including pre-menarche characteristics despite already including $P(teen|Or)$ further reduces the outcomes' variation in the following estimations.

Columns (1)-(3) show three things. First, the likelihood of becoming a teen mother decreases between 3% to 4% for each year of difference between the observed and expected AAM, regardless of the controls used. Second, the excluded instrument is strongly correlated with the probability of teen motherhood, as confirmed by the high chi-squared statistics and their corresponding p-values. Third, as suggested by the Pseudo R squared, including the propensity of teen motherhood given by pre-menarche characteristics and its polynomials helps explain more variation in teen motherhood status, which should increase the precision of the estimates in the following results. Hence, the model from column (3) is my preferred selection-into-teen-motherhood model for the mothers' sample.

In regards to variables used, the only difference between columns (4)-(6) and (1)-(3), other than the larger sample, is that instead of controlling for maternal year and month of birth, I use children's year and month of birth, since it is impossible to control for both the mother's and children's dates of birth while estimating a probit model because of the high multicollinearity between them and teen motherhood status. For the same reasons outlined for the mothers' sample, for the firstborn sample, the preferred estimation is the more flexible model presented in column (6).

Figure 4 plots the predicted probabilities of columns (3) and (6) models by teen motherhood status. For the mothers' sample, the predicted probabilities of only including the difference between the observed and expected AAM and its squared range from 0.159 to 0.405 while including all column (3)'s controls increases the coverage of the range, which then goes from .00034 to 0.933. Similarly, for the firstborns' sample, including all controls expands the range of support from 0.221-0.509 to 0.008 - 0.999. To aggregate the MTE to produce the ATE, ATT and ATU, is necessary that the instrument generates variation in the full unit interval (i.e., from 0 to 1) (Heckman and Vytlacil, 2005). As is common in the applied MTE literature, I don't have Full support. However, the available support is enough to test for essential heterogeneity. Also, I follow Carneiro et. al (Carneiro et al., 2011) and in the next section present local version of the ATE, ATT and ATU, calculated within the empirical support of $P(Z)$.

5.2 Two stage least square estimates

Before presenting the estimated MTE, it is helpful to illustrate some of the challenges of using instrumental variables, especially if treatment effect heterogeneity is a distinct possibility. To do so, I specify the following outcome equation:

$$Y = \alpha + \beta \times \text{Teen} + \delta' X + e, \tag{18}$$

where Y denotes the outcome for either mother or firstborn, where the direct effect of the AAM has already been subtracted, and X is the same vector of observables characteristics used in the estimation of $P(Z)$. As most studies do, equation 18 implicitly incorporates the assumption of constant treatment effects of teen childbearing. The first stage equation is given by:

$$\text{Teen} = a + b'Z + \delta' X + u,$$

where Z is the vector of excluded instruments.

I specify the instrument in four different ways. First, I use deviation from the expected age at menarche. Second, a dummy variable equal to one if the observed AAM is below the expected AAM. Third, I use a propensity score calculated with dummy variables for each age at menarche form 11 to 16 (10 is the excluded category. The fourth instrument is my preferred specification of the propensity score. Under constant effects, the specification of the instruments should produce similar coefficients. Under effect heterogeneity, each instrument may calculate a different LATE, or may reflect different weighted averages treatment effects.[18]

[18] Heckman, Urzua and Vytlacil (2013) show that only using the propensity score as the instrument guarantees non-negative weights.

5.2.1 Mothers' sample

Table 3 shows the 2SLS estimates of the effect of teen childbearing for mothers on their years of schooling, the likelihood of completing high school, and their likelihood of participating in the labor force under different ways of specifying the excluded instruments. For reference, it also includes the effects suggested by an OLS regression.

Years of schooling._ In Panel A, column (1), the OLS estimate suggests that a woman who becomes pregnant from ages 15-19 loses 2.2 years of completed schooling. Column (2) uses the difference between the observed AAM and the expected AAM as the excluded instrument in a 2SLS linear specification and suggests that, on average, women lose 0.35 years of schooling. However, this estimate is not statistically different from zero. Column (3) uses as an excluded instrument a dummy variable equal to one if the observed AAM is below the expected AAM. It suggests that women only lose an average of 0.18 years of attained schooling. The effect is also not statistically different from zero.

Columns (4) and (5) use predicted probabilities as instruments[19]. The first stage of the model from column (4) uses dummy variables for each year of observed age at menarche as the excluded instruments, and the first stage from column (5) uses the difference between the observed and expected AAM and its square as excluded instruments. The estimates from these two models, both statistically different from zero, suggest that teen childbearing causes women to lose -1.4 and 0.96 years of attained schooling, respectively. However, a test of equality of these coefficients is rejected at the 1% level.

Two key insights emerge from the data in table 3. Firstly, all instrumental variable (IV) estimates fall between zero and the ordinary least squares (OLS) regression estimate. This suggests that the observed correlation between teen motherhood and lower socioeconomic status may be influenced by selection bias. Secondly, different instruments, even when using similar sources of exogenous

[19] **?** shows that IV estimates from a vector Z of excluded instruments may be interpreted as a weighted average of LATEs or MTEs, and argue that these weights are not guaranteed to be positive, even when all LATE or MTE are positive. They show, however, that using the propensity score P(Z) as an instrument always produces non-negative weights. Furthermore, **?** suggests that using the predicted probabilities as instruments is robust to the misspecification of the propensity score, and **?** shows that this procedure outperforms the linear 2SLS in terms of efficiency.

variation, can yield divergent conclusions. Panel A. of figure 3 plots these coefficients and their 95% confidence intervals[20]. Looking at estimates from columns (1) and (2), a reasonable conclusion is that there is no evidence that teen childbearing reduces schooling attainment. In contrast, the estimates from columns (3) and (4) suggest that teen childbearing causes significant losses in schooling attainment. And even for these two estimates, the conclusion is not quite the same, as the estimate from column (3) is 50% larger in absolute value than that of column (4).

The pattern of IV estimates is consistent with unobserved heterogeneity, as the different ways of defining the instrument may generate different LATEs. However, heterogeneity is not the only possible explanation consistent with the observed pattern.

High school completion._ Panel B of table 3 reports results from OLS and IV estimates for the effects on high school completion. The OLS estimate suggests that adolescent mothers are 22% less likely to have completed high school than their adult counterparts. As with years of schooling attained, the IV estimates suggest mixed results. The estimates from columns (1) and (2) suggest that women who gave birth as adolescents are 7.5% and 3.9% less likely to finish high school, respectively. These estimates are not statistically different from each other, and neither is significantly different from zero.

In contrast, the estimates from columns (3) and (4) suggest that teen childbearing is causally related to an 18% and 15% lower likelihood of completing high school. As with years of schooling attainment, these estimates are statistically different from zero, as well as from estimates from columns (1) and (2), and each other at the 5% level. Panel B of figure 5 plots these estimates, which are, not surprisingly, also consistent with unobserved heterogeneity.

Labor force participation._ Panel C of table 3 reports results from OLS and IV estimates for the effects on labor force participation. The OLS estimate suggests that adolescent mothers are 3% less likely to participate in the labor force. Interestingly, three IV estimates are larger in absolute value

[20] The p-value for a test of equality of all coefficients is rejected at the 1% level. The estimates from columns (1) and (2) are not statistically different. Individual tests for equality of coefficients between the estimate from column (3) with the rest of the coefficients are all rejected at the 1% level. The estimate from column (4) is statistically different from that of column (2) at the 10% level.

than the OLS estimate, but none of the four statistically differs from zero. Furthermore, individual test equality of coefficients suggests that none of the estimates are statistically different from each other. Panel C of figure 5 plots these estimates. In this case, all IV estimates lead to the same interpretation: we fail to reject the null.

5.2.2 Firstborns' sample

Table 4 shows the 2SLS estimates of the effect of teen childbearing for firstborn children on the likelihood of being alive at the time of the survey, relative years of schooling for children of ages 6-14, and height for age z scores for children 14 or younger under the same ways of specifying the excluded instruments as for the previous sample, as well as an OLS estimate.

Mortality._ The OLS estimate in Column 1 of panel A suggests that teen motherhood is associated with a 0.6% lower likelihood of being alive at the time of the survey. The estimate is statistically significant, implying that teen motherhood is associated with six extra deaths for every thousand firstborn children in my sample, a 36% larger than the overall mean. The 2SLS estimates of the following columns are also negative, but none are significantly different from zero. The estimate from column (2) is the same up to 3 decimal points compared to the OLS estimate, while the estimate from column (3) is significantly larger in absolute value. Nevertheless, a joint test for all the IV coefficients cannot reject the null that these are all nondifferent from zero. Likewise, individual equality tests for coefficients cannot be rejected between any of the coefficients. Since the outcome is close to one, linear specifications might cause concern. Yet, the coefficient from IV Poisson regression is -0.003 (not presented), or half the OLS estimate in absolute value but still not significantly different from zero.

Figure 6 plots the coefficients with their corresponding 95% confidence intervals. It shows that all confidence intervals include the OLS estimate and zero. Thus, these estimates are not precise enough to suggest that the OLS estimate is biased nor that teen childbearing negatively impacts the likelihood of the firstborn being alive at the time of the survey. For the same reason, they are of little use as evidence of heterogeneity, even if they range from positive to more negative than the OLS estimate.

Relative Years of Education._ Panel B of table 6 reports the OLS and IV estimates for relative years of schooling, defined for children of ages 6 to 14, as the number of years of schooling attained divided by the number of years children should have attained given their age. Column (1) contains the OLS estimate and shows a negative correlation. In particular, it suggests that firstborn children of teen mothers attain 0.016 years less per year of age, which implies that by age 14, children of teen mothers attained 0.14 fewer years of schooling—a small effect.

All the 2SLS estimates are negative but are not statistically significant. Estimates from columns (1) and (2) are larger in absolute value than the OLS, while the estimates from (3) and (4) are very close to the OLS estimate. As before, none of the estimates is statistically significantly different than the OLS estimate. Individual tests of equality between the coefficients also reject that any of them is different than any other. Similarly to the outcome of mortality, the effects, if any, are too small to detect given the instruments. Once again, from these estimates, it is not possible to say much about whether the OLS estimate is biased or whether there is a causal link between teen motherhood and lower relative schooling attainment. These estimates also are not suggestive of treatment heterogeneity for this outcome.

Height-for-age Z scores._ Panel C of table 4 presents the results from the 2SLS specifications on the effects of teen childbearing on height-for-age z scores for firstborn children of ages 14 or younger. Column 1, the OLS estimate, suggests a negative correlation. Firstborn children have, on average, 0.16 lower z scores.

All the IV estimates are positive, with effect sizes that range from 0.07 to 0.16. However, the estimates are imprecise. Panel C of figure 6 plots these coefficients, and like with the two previous outcomes, their 95% confidence intervals include in the OLS estimate and zero. Individual equality tests among each coefficient also do not reject the null that the estimates are statistically identical.

5.3 Marginal Treatment Effects

5.3.1 Mothers' Sample

Figure 7 presents plots for the MTE curve from equation 12 for each outcome. Full support in $P(z)$ is not achieved. Additionally, little support within values of $P(Z)$ may produce unrealistic estimates. Thus, for baseline results, I trim all values below the one percentile and above the 99 percentile of the $P(z)$ distribution, dropping 1% of the sample (245 observations) when estimating MTE on years of schooling and high school completion. I trim 5% of the sample (1,224 observations) when estimating the MTE curve for labor force participation.

Following Carneiro et al. (2011), I rescale the weights within the support so that they sum up to one and present a local (i.e., within the support) version of the ATE, ATT, and ATU for each outcome in table 5.

Years of schooling._ Panel A of figure 7 shows the MTE estimates for years of schooling at values of $P(z)$ that go from 0.08 to 0.63. The shape of the curve suggests heterogeneity in treatment effect. In particular, treatment effects are positive at lower values of U_{teen} but decrease continuously with higher values of U_{teen}. Treatment effects range from a positive 1.46 years of schooling for women at the 9% quantile of resistance to teen childbearing to a negative 2.07 for women at the 63% quantile of resistance to teen childbearing. Positive treatment effects correspond to values where u is less than 0.2, at which point they cross the zero threshold and become negative. However, the 95% confidence intervals from 150 bootstrap replication shown in panel A of figure 7 always include zero, suggesting that these estimates are imprecise.

Column (1) of table 5 presents treatment effect parameters. While these parameters are local estimates because I lack full support in P(z), they still suggest heterogeneity in treatment effects. In particular, the ATE for individuals with P(Z) values from 0.08 to 0.63 equals -1.1, significant at the 10% level. The ATU is -1.88, also significant at the 10% level. In contrast, the ATT equals 0.32, but it is not statistically different from zero. However, the p-value from a joint significance test for the π_l parameters presented at the bottom of table 5 that can be interpreted as a test of essential heterogeneity suggests that these parameters are not statistically different from zero. Given the MTE curve's shape and the treatment effect parameters, the failure to reject the null might be related to a power issue rather than an absence of economically relevant heterogeneity in treatment effects.

High school completion._ Panel B of figure 7 shows the MTE estimates for the probability of completing high school at values of $P(z)$ that go from 0.08 to 0.63. The shape of the curve suggests heterogeneity in treatment effects. Treatment effects are positive at lower values of U_{teen} but decrease continuously with higher values of U_{teen}. Effects range from a 0.35 higher likelihood of completing high school for women at the 9% quantile of resistance to teen childbearing to a 0.33 lower likelihood of completing high school for women at the 63% quantile of resistance to teen childbearing. Positive treatment effects correspond to values where u is less than 0.21, at which point they cross the zero threshold and become negative. The 95% confidence intervals from 150 bootstrap replication shown in panel B of figure 7 include zero for the positive section of the MTE curve but significantly differ from zero around values of u>0.3.

Column (2) of table 5 presents local treatment effect parameters –i.e., parameters estimated within the available support. Again, these parameters suggest heterogeneity in treatment effects. In particular, the ATE for individuals with P(Z) values from 0.08 to 0.63 equals -0.17, significant at the 1% level. The ATU is -0.3, also significant at the 1% level. In contrast, the ATT equals 0.05 but is not statistically different from zero. The p-value for the test of essential heterogeneity presented at the bottom of table 5 is significant, also suggesting that treatment effects vary across resistant-to-treatment quantiles for this outcome.

Labor force participation._ Panel C of figure 7 shows the MTE estimates for the probability of participating in the labor force at values of $P(z)$ that go from 0.14 to 0.51. The shape of the curve suggests heterogeneity in treatment effects. Treatment effects are positive at lower values of U_{teen} but decrease continuously with higher values of U_{teen}. Effects range from a 0.04 higher likelihood of participating for women at the 14% quantile of resistance to teen childbearing to a 0.26 lower likelihood of participating in the labor force at the 51% quantile of resistance to teen childbearing. Positive treatment effects correspond to values where u is less than 0.22, at which point they cross the zero threshold and become negative. The 95% confidence intervals from 150 bootstrap replication shown in panel B of figure 7 include zero for the positive section of the MTE curve but significantly differ from zero at the negative section of the MTE curve around values of $u > 0.4$.

Column (3) of table 5 presents local treatment effect parameters –i.e., parameters estimated within the available support. Again, these parameters suggest heterogeneity in treatment effects. The ATE for individuals with P(Z) values from 0.014 to 0.51 equals -0.08, significant at the 10% level. The ATU is -0.18, significant at the 1% level. In contrast, the ATT equals 0.1 and is not statistically different from zero at the 5% level. The p-value for the test of essential heterogeneity presented at the bottom of table 5 is significant, also suggesting that treatment effects vary across resistant-to-treatment quantiles for this outcome.

5.3.2 Firstborn children sample

Figure 8 presents plots for the MTE curve from equation 12 for each outcome. For the mortality outcome and relative years of education, $P(z)$ ranges from 0 to 1, allowing for direct calculation of treatment parameters. For height-for-age Z scores, I trim a total of 4% of the sample from the tails of the $P(z)$ distribution for this outcome to avoid little support from producing unrealistic treatment effects estimates at the extremes of the MTE curve. After trimming, the propensity score ranges from 0.16 to 0.62

Mortality._ Panel A of figure 8 shows the MTE estimates for the probability of being alive at the time of the survey for values of $P(z)$ that go from 0 to 1. The shape of the curve does not support the hypothesis of heterogeneity in treatment effects. However, effects range from a 0.008 higher likelihood of being alive at the time of the survey for the 1% quantile of resistance to teen childbearing to a 0.03 lower likelihood for the 99% quantile of resistance to teen childbearing. Both average and heterogeneous effects may be too small to be detected. The 95% confidence intervals from 150 bootstrap replication always include zero.

Column (1) of table 6 presents treatment effect parameters. The ATE for individuals with P(Z) values from 0.01 to 0.99 equals -0.01 but is not significantly different than zero. The ATUT is -0.02, which is also not statistically significant. The ATT equals -0.002 and is also not statistically different from zero. The p-value for the test of essential heterogeneity presented at the bottom of table 6 is close to one, also providing no evidence that treatment effects vary across resistant-to-treatment quantiles for this outcome.

Relative years of education_ Panel B of figure 8 shows the MTE estimates for the probability of being alive at values of $P(z)$ that go from 0 to 1. The shape of the curve does not support the hypothesis of heterogeneity in treatment effects. The effects range from 0.02 fewer years of schooling attained per year of age for the 1% quantile of resistance to teen childbearing to 0.08 fewer years of schooling attained for the 99% quantile of resistance to teen childbearing. The magnitude of effects are small across the quantiles of resistance, and the 95% confidence intervals from 150 bootstrap replication always include zero.

Column (2) of table 6 presents treatment effect parameters. None of the treatment parameters is economically or statistically significant, and the p-value for the test of essential heterogeneity presented at the bottom of table 6 is close to one, also providing no evidence that treatment effects vary across resistant-to-treatment quantiles for this outcome.

Height-for-age Z scores._ Panel C of figure 8 shows the MTE estimates for the probability of being alive at values of $P(z)$ that go from 0.16 to 0.62. The shape of the curve does not support the hypothesis of heterogeneity in treatment effects. However, the effects range from 0.06 lower standard deviations of height-for-age for the 16% quantile of resistance to teen childbearing to 0.28 higher standard deviation for the 99=62% quantile of resistance to teen childbearing. Still, the 95% dconfidence intervals from 150 bootstrap replication always include zero.

Column (3) of table 6 presents treatment effect parameters. None of the treatment parameters is statistically significant, and the p-value for the test of essential heterogeneity presented at the bottom of table 6 is 0.12, also providing no evidence that treatment effects vary across resistant-to-treatment quantiles for this outcome.

6 Conclusion

One in every three women becomes a mother as a teenager in low- and middle-income countries. Despite the severity of this statistic, research about the consequences of teen childbearing is lacking in these countries. The necessity to close this gap becomes more pressing as the studies from high-income countries provide mixed results, making it hard to draw clear conclusions and policy recommendations.

In this paper, I estimate marginal treatment effects to examine whether the consequences of being a teen mom differ across women in ways that are unobservable, using data from a middle-income country with high teen pregnancy rates. I show that effects vary for maternal years of schooling, the probability of high school completion, and the likelihood of labor force participation based on a parameter that encompasses unobservable characteristics that make women less likely to become teen mothers. Treatment effects seem more harmful the higher the unobserved resistance to treatment. However, when looking at the short-term effects on the children of teen moms, I fail to find significant detrimental impacts. Even non-casual associations suggest small effects. Nonetheless, my estimates may be too sensitive to modeling choices; thus, further research is required.

These findings have implications for policy and future research. The main policy implication that my finding suggests is that policies aiming to reduce teen births can potentially improve women's educational and labor outcomes, especially in settings where the cost of opportunity of losing educational attainment is sufficiently high. For future research, these findings suggest that the best way to move forward is to estimate policy-relevant parameters using variation directly from policies, as natural experiments that fail to consider heterogeneity will likely produce local average treatment effects that are not economically relevant.

Table 1. Summary Statistics

	N= 39,629	N= 24,452	N= 27,957
		Women Characteristics	
Age	30.22	35.81	32.962
Year of birth	1988.291	1982.695	1985.339
Age$>25=1$	.646	1	.794
Teen mother=1	.289	.300	.385
Age at menarche	12.873	12.940	12.894
Black=1	.0253	.024	.026
Indigenous=1	.130	.121	.125
Lives in rural area=1	.368	.352	.369
Lives in costal region=1	.390	.392	.385
Lives in highland region=1	.355	.354	.361
Lives in Amazonia region=1	.219	.213	.220
Lives in Galapagos region=1	.034	.040	.033
Years of schooling	11.50	11.435	11.339
High school completed=1	.5156	.538	.529
Primary school completed=1	.928	.916	.921
In labor force=1	.504	.621	.559
Monthly Labor Income ($)	693.29	748.16	699.53
			Child Characteristics
Alive			.983
Female			.488
Age			11.86
Years of schooling (age>6)			7.75
Relative grade attainment (ages 6-14)			1.028
Z score, height for age (age<15)			-.897

Source: ENSANUT Ecuador 2018. This table presents summary statistics for three overlapping samples used in the analysis. The first column corresponds to the sample used for the random forest models. The second column corresponds to the sample used for women's analysis. The third column corresponds to the sample used for the firstborn analysis.

Table 2. Selection Equation
Average Marginal Effects

	Mothers Sample			Children Sample		
	(1)	(2)	(3)	(4)	(5)	(6)
AAM - E[AAM]	-0.037***	-0.037***	-0.045***	-0.040***	-0.041***	-0.034***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
E[AAM]	0.074***		0.340***	-0.057***		-0.289***
	(0.011)		(0.020)	(0.011)		(0.018)
$P(teen\|Or)$			0.028***			0.035***
			(0.001)			(0.001)
Rural area=1		0.071***			0.091***	
		(0.007)			(0.008)	
Coastal Region=1		0.076***			0.079***	
		(0.008)			(0.009)	
Amazon Region=1		0.074***			0.075***	
		(0.009)			(0.009)	
Galapagos=1		-0.023			-0.061***	
		(0.015)			(0.016)	
Black=1		0.114***			0.206	
		(0.020)			(0.195)	
Indigenous=1		-0.002			0.079	
		(0.010)			(0.193)	
Height percentile (12.6-25) = 1		0.025			0.047***	
		(0.014)			(0.014)	
Height percentile (25.01-37.5) = 1		0.020			0.059***	
		(0.014)			(0.014)	
Height percentile (37.6-50) = 1		0.002			0.052***	
		(0.014)			(0.014)	
Height percentile (51-62.5) = 1		-0.005			0.044***	
		(0.013)			(0.013)	
Height percentile (62.6-75) = 1		-0.014			0.046***	
		(0.013)			(0.014)	
Height percentile (75.01-87.5) = 1		-0.031*			0.030*	
		(0.014)			(0.014)	
Height percentile (87.6-100) = 1		-0.075***			0.001	
		(0.013)			(0.014)	
Observations	24,452	24,452	24,452	28,148	27,957	27,957
Pseudo R^2	0.010	0.029	0.086	0.010	0.036	0.133
X^2, excluded instruments	262.676	253.919	302.982	301.927	308.604	190.016
P value, excluded instruments = 0	0.000	0.000	0.000	0.000	0.000	0.000

The table presents average marginal effects at covariate means from probit selection models in which the dependent variable is equal to one if a woman had a child at ages 15 to 19. Columns (1), (2) and (3) are for the mother's sample which excludes women 24 or younger. Columns (2) and (3) control for mother's month and year of birth. Column (3)'s model includes also all controls of column (2), but they are not presented as coefficients have no clear interpretation when including $P(teen|Or)$. Columns (4), (5) and (6) are for the first born sample, thus include mothers from 15 to 49. Columns (4) and (6) control for children's month and year of birth, all other controls use maternal information. Column (6)'s model includes also all controls of column (5), but are not included for the same reason as above. Standard errors were clustered at the unit of sampling. Stars represent significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3. 2SLS estimates of teen childbearing effects on mothers

	(1)	(2)	(3)	(4)	(5)
	OLS	IV: $\_Z$	IV: $1\{\_Z < 0\}$	IV: P(AAM dum)	IV: $P(\_Z, \_Z^2)$
Panel A.	Years of Schooling				
Teen birth=1	-2.225***	-0.359	-0.184	-1.440***	-0.964**
	(0.055)	(0.484)	(0.625)	(0.375)	(0.393)
Y mean	11.435				
Panel B.	Highschool complete=1				
Teen birth=1	-0.242***	-0.075	-0.039	-0.187***	-0.153***
	(0.007)	(0.055)	(0.072)	(0.043)	(0.045)
Y mean	0.538				
Panel C.	Labor Force participation				
Teen birth=1	-0.033***	-0.073	-0.059	-0.036	-0.028
	(0.007)	(0.058)	(0.074)	(0.046)	(0.048)
Y mean	0.622				
Obs	24,452				
Effective F stat		355.793	205.402	572.519	513.849

This table presents 2SLS estimates for women's outcomes using four different ways of specifying the instrument. All specifications include month and year of birth, $E[AAM]_i$ and its squared, polynomials of $P(teen|O)$, region of residence, whether the individual resides in a rural area, ethnicity, and dummies for one of 8 height quantiles as controls. Effective F statistics corresponds to the weak instrument test of Olea and Pflueger (2013). Column 1 and 2 are standard 2SLS linear models, columns 3 and 4 use predicted probabilities as instruments. Standard errors clustered at the unit of sampling; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. $\_Z = AAM_i - E[AAM]_i$. AAM_dum: dummy variables for each observed AAM.

Table 4. IV estimates of teen childbearing effects on first born children

	(1)	(2)	(3)	(4)	(5)
	OLS	IV: _Z	IV: $1\{\_Z < 0\}$	IV: P(AAM dum)	IV: $P(\_Z, \_Z^2)$
Panel A.	Mortality, child is alive=1				
Teen birth=1	-0.006***	-0.006	-0.029	-0.003	-0.008
	(0.002)	(0.020)	(0.029)	(0.011)	(0.012)
Effective F stat		203.238	97.387	639.671	596.536
Y mean	0.983				
Obs	27,957				
Panel B.	Relative Years of Schooling, ages 6-14				
Teen birth=1	-0.016***	-0.052	-0.096	-0.017	-0.020
	(0.006)	(0.061)	(0.092)	(0.046)	(0.044)
Effective F stat		80.757	36.440	145.636	155.459
Y mean	1.028				
Obs	8,219				
Panel C.	Height-for-age Z scores, age<15				
Teen birth=1	-0.163***	0.142	0.179	0.074	0.162
	(0.022)	(0.314)	(0.534)	(0.165)	(0.165)
Effective F stat		88.234	31.571	311.189	280.725
Y mean	-0.897				
Obs	16,865				

This table presents 2SLS estimates for first born children using four different ways of specifying the instrument. All specifications include month and year of birth, $E[AAM]_i$ and its squared, polynomials of $P(teen|O)$, region of residence, whether the individual resides in a rural area, ethnicity, and dummies for one of 8 height quantiles as controls. Effective F statistics corresponds to the weak instrument test of Olea and Pflueger (2013). Column 1 and 2 are standard 2SLS linear models, columns 3 and 4 use predicted probabilities as instruments. Standard errors clustered at the unit of sampling; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. $\_Z = AAM_i - E[AAM]_i$. AAM_dum: dummy variables for each observed AAM.

Table 5. Treatment effect parameters for maternal outcomes

	Years of Schooling	High school completion	Labor Force Participation
P(z) support	$[0.08, 0.63]$	$[0.08, 0.63]$	$[0.14, 0.51]$
ATE	-1.097*	-0.171***	-0.089*
	(0.580)	(0.061)	(0.047)
ATT	0.315	0.059	0.104**
	(0.611)	(0.078)	(0.051)
ATUT	-1.878*	-0.304***	-0.182***
	(1.00)	(0.105)	(0.062)
LATE	-0.708	-0.158***	0.006
	(0.529)	(0.057)	(0.044)
P value for test of essential heterogeneity	0.342	0.029	0.013
Obs	24,190	24,190	23,211

This table presents local treatment effect parameters obtained as weighted averages of the MTE, rescaling the weights so they sum up to one. Standard error estimated from 150 bootstrapped replications. ATE: Average Treatment Effect; ATT: Average treatment on the treated; ATU: Average treatment on the untreated. The test of essential heterogeneity is a joint test on the coefficients p.

Table 6. Treatment effect parameters for firstborn children outcomes

	Mortality	Relative Years of Schooling Ages 6-14	HAZ scores ages<15
P(z) support	$[0.01, 0.99]$	$[0.01, 0.99]$	$[0.16, 0.62]$
ATE	-0.013	-0.060	-0.027
	(0.034)	(0.090)	(0.151)
ATT	-0.002	-.025	-0.186
	(0.040)	(0.063)	(0.2282)
ATUT	-0.020	-0.081	0.064
	(0.0623)	(0.133)	0.230
LATE	-0.004	-0.083	-0.159
	(0.024)	(0.055)	(0.187)
P value for test of essential heterogeneity	0.971	0.978	0.1168
Obs	27,957	8,217	16,021

This table presents treatment effect parameters obtained as weighted averages of the MTE, rescaling the weights so they sum up to one for HAZ scores. Standard error estimated from 150 bootstrapped replications. ATE: Average Treatment Effect; ATT: Average treatment on the treated; ATU: Average treatment on the untreated. The test of essential heterogeneity is a joint test on the coefficients p.

Figure 1. Age at Menarche, Observed, Expected and Deviations

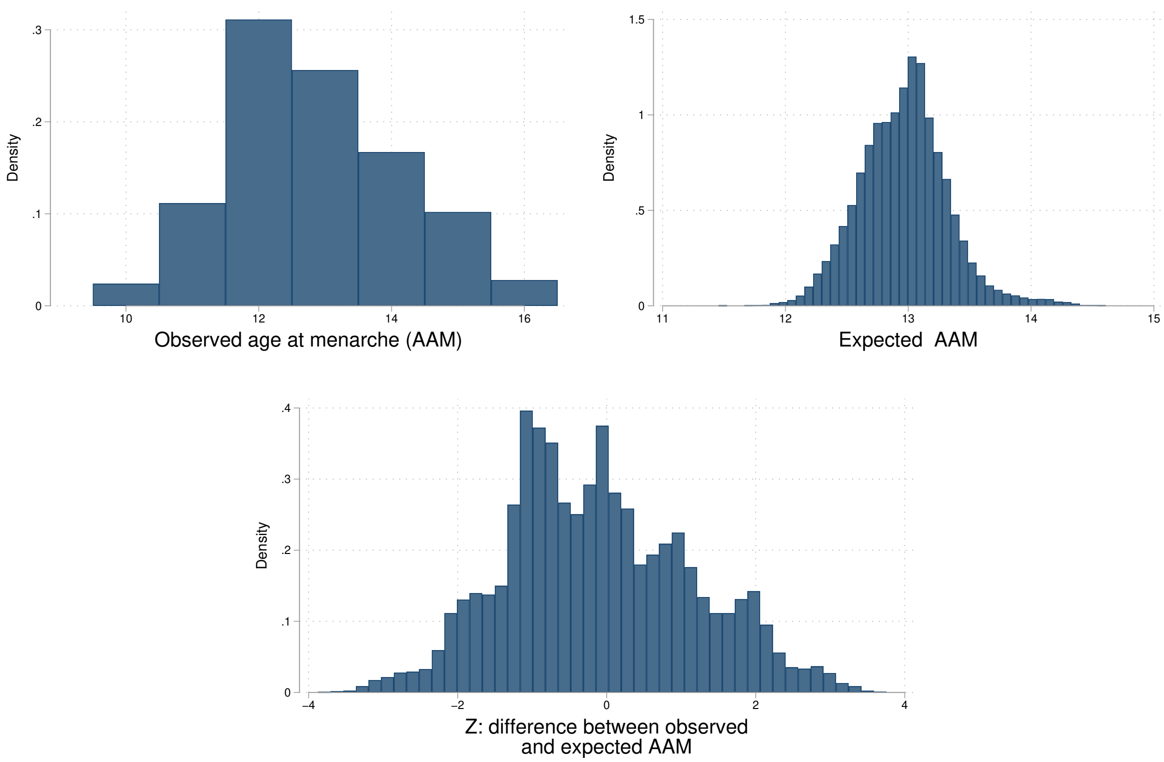


Figure 2. Deviations from the AAM still predict Teen Births

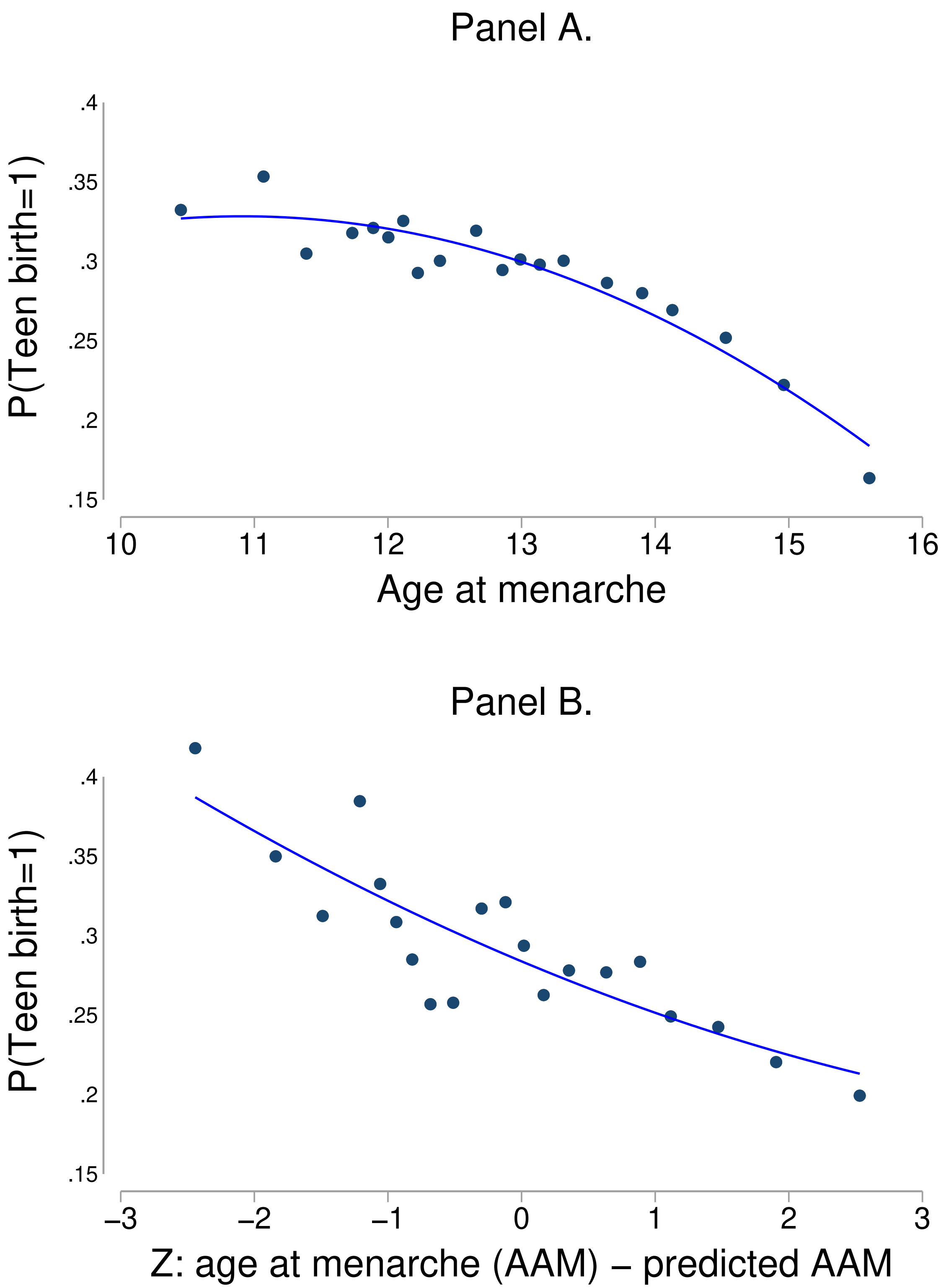


Figure 3. Isolating plausibly exogenous variation

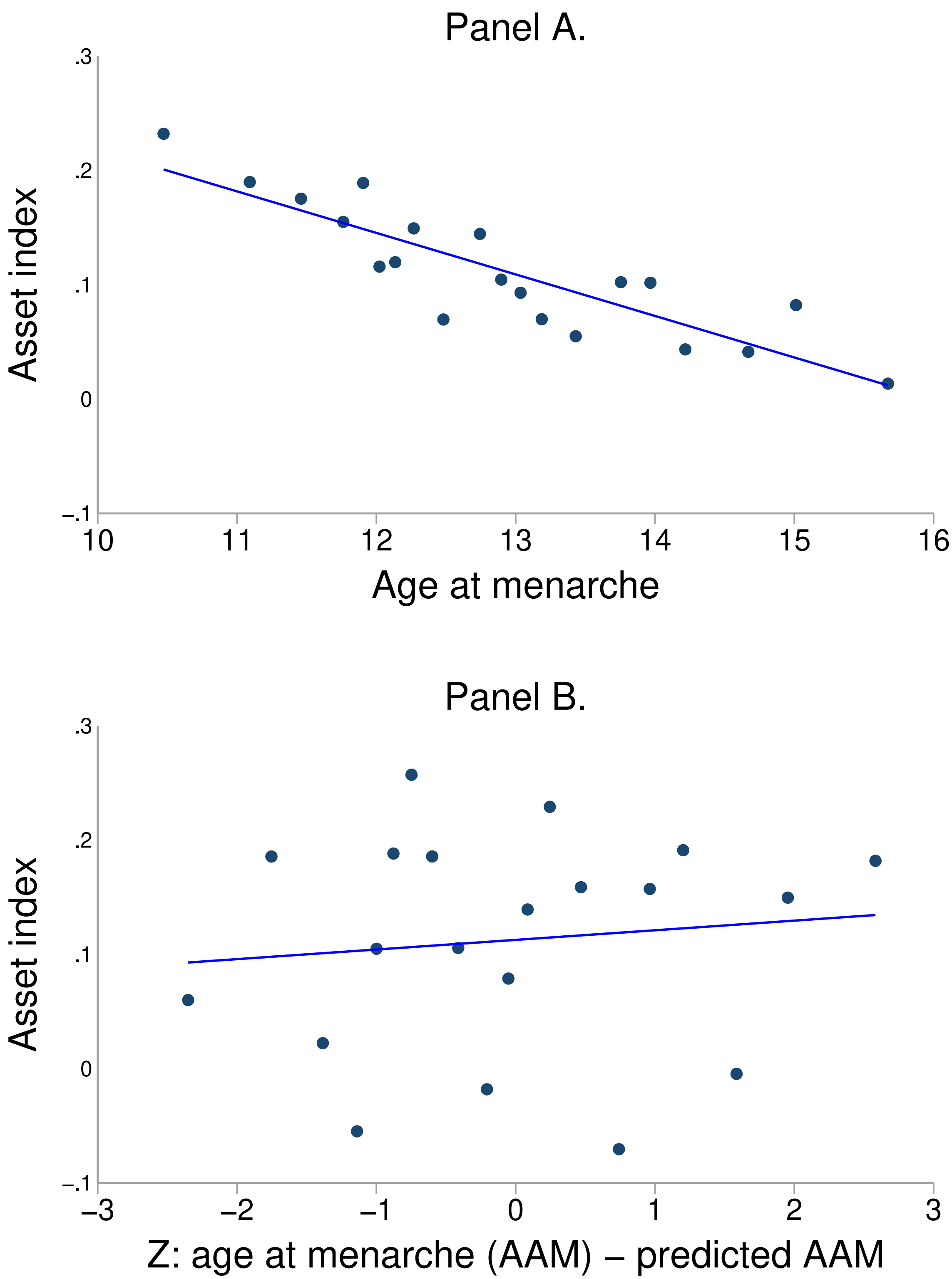


Figure 4. Common Support

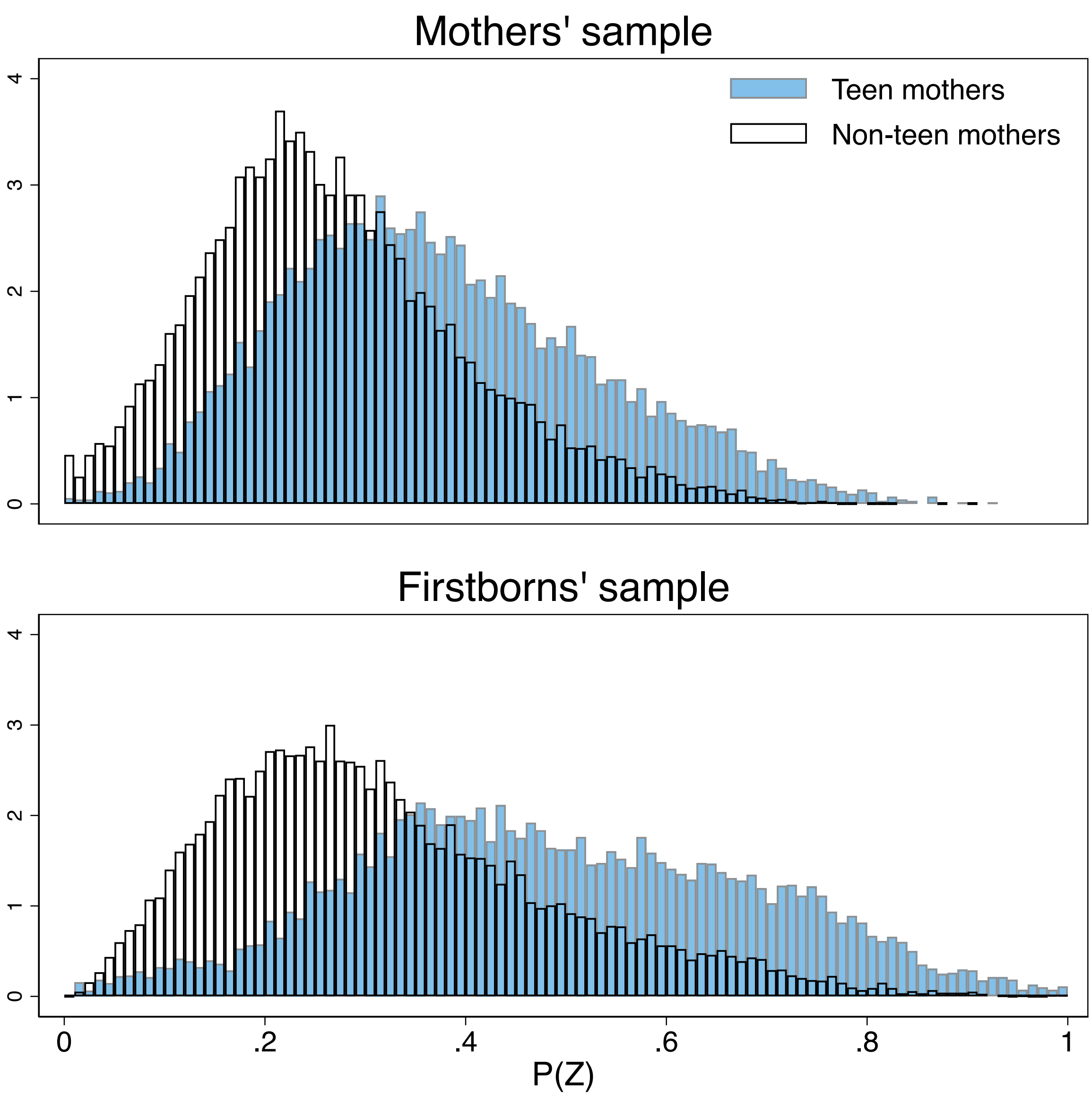


Figure 5. IV estimates of teen childbearing effects on maternal outcomes

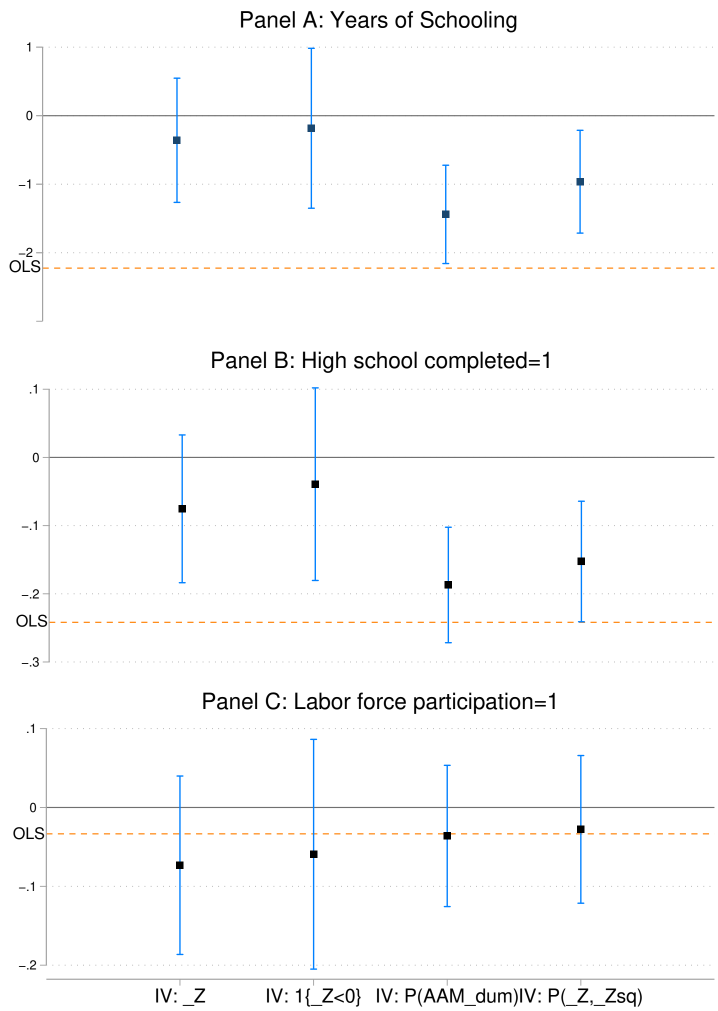


Figure 6. IV estimates of teen childbearing effects on first born outcomes

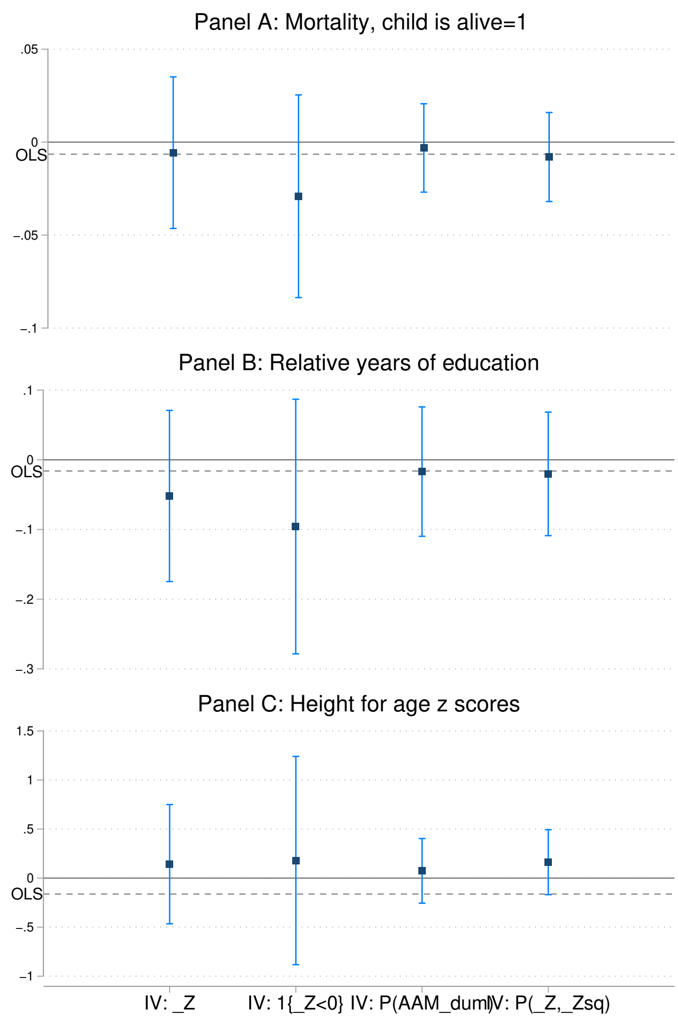


Figure 7. Marginal Treatment Effects on mother's outcomes

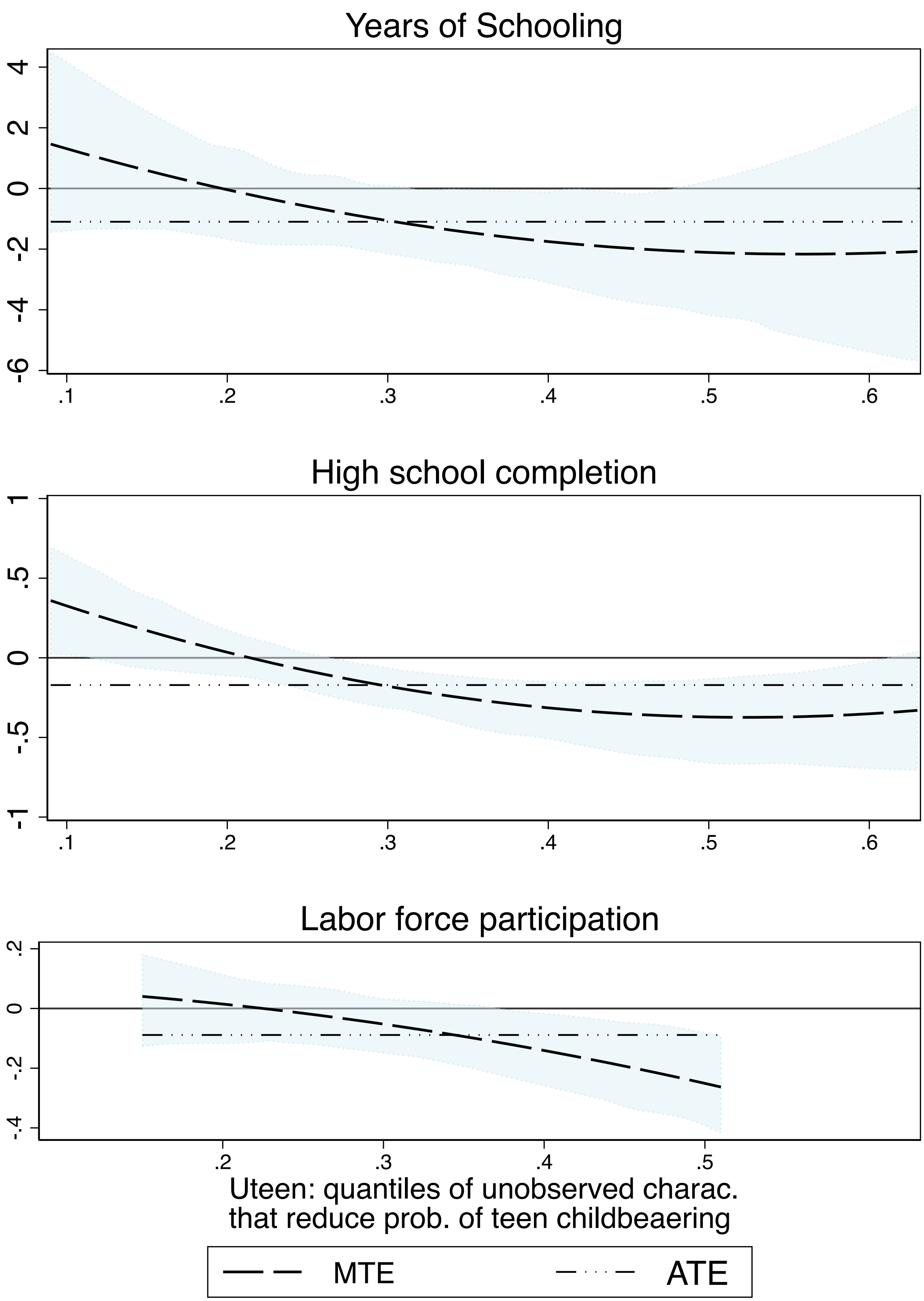
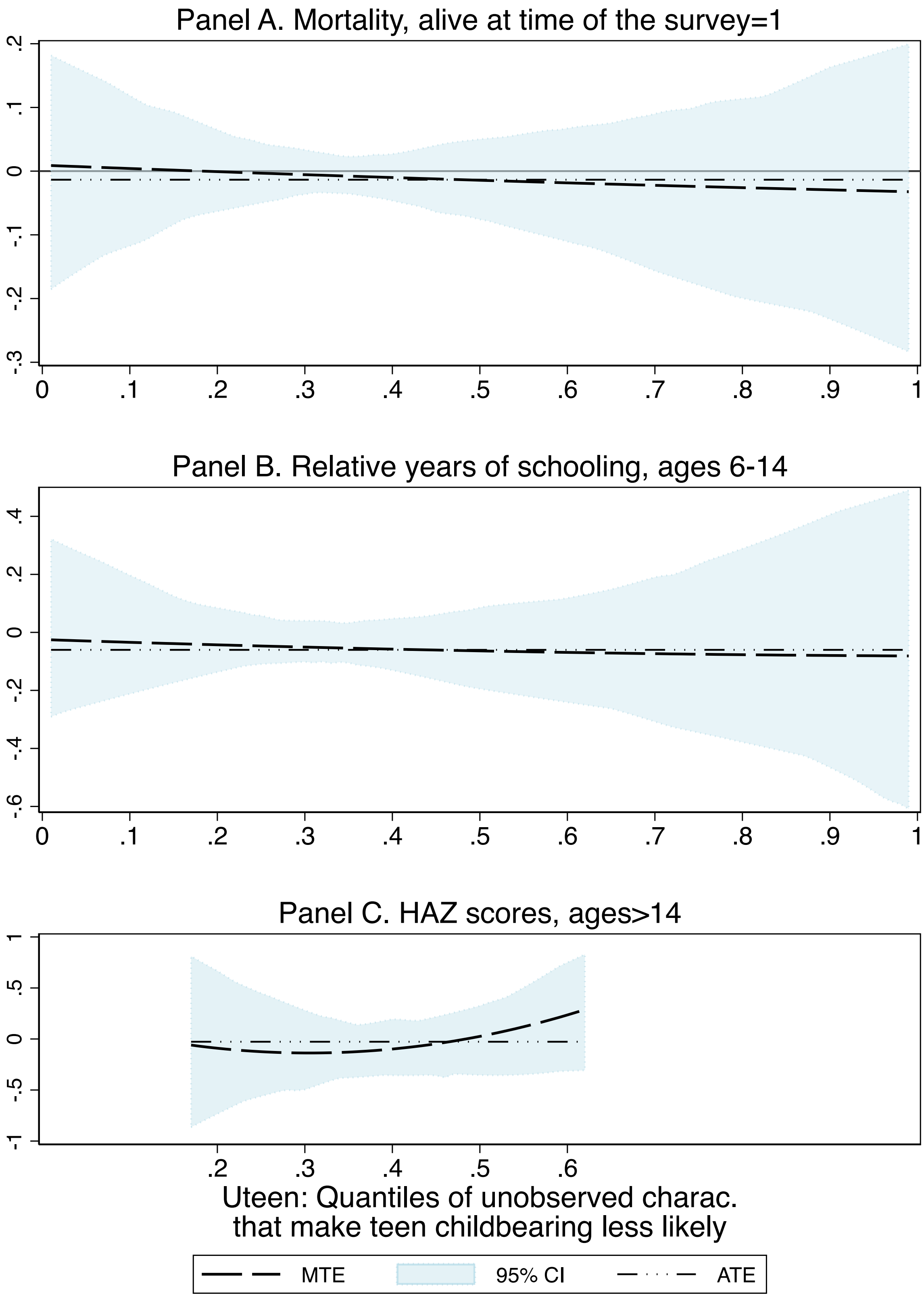


Figure 8. Marginal Treatment Effects on firstborn children outcomes



References

Aizer, Anna, Paul Devereux, and Kjell Salvanes, "Grandparents, Moms, or Dads? Why Children of Teen Mothers Do Worse in Life," *Journal of Human Resources*, 2022, *57* (6), 2012–2047.

Anderson, Carl A., David L. Duffy, Nicholas G. Martin, and Peter M. Visscher, "Estimation of Variance Components for Age at Menarche in Twin Families," *Behavior Genetics*, September 2007, *37* (5), 668–677.

Andresen, Martin Eckhoff, "MTEFE: Stata module to estimate marginal treatment effects"," 2018. This Verison 20190513.

Ashcraft, Adam, Iván Fernández-Val, and Kevin Lang, "The Consequences of Teenage Childbearing: Consistent Estimates When Abortion Makes Miscarriage Non-random," *The Economic Journal*, 2013, *123* (571), 875–905.

Azevedo, Joao Pedro, Luis F. Lopez-Calva, and Elizaveta Perova, "Is the Baby to Blame? An Inquiry into the Consequences of Early Childbearing," Technical Report, World Bank, Washington, DC 2012.

Baams, Laura, Judith Semon Dubas, Geertjan Overbeek, and Marcel A. G. van Aken, "Transitions in Body and Behavior: A Meta-Analytic Study on the Relationship Between Pubertal Development and Adolescent Sexual Behavior," *Journal of Adolescent Health*, 2015, *56* (6), 586–598.

Benson, M. D. and E. J. Torpy, "Sexual behavior in junior high school students," *Obstetrics and Gynecology*, 1995, *85* (2), 279–284.

Borusyak, Kirill and Peter Hull, "Non-random exposure to exogenous shocks," *Econometrica*, 2021.

Branson, Nicola and Tanya Byker, "Causes and consequences of teen childbearing: Evidence from a reproductive health intervention in South Africa," *Journal of Health Economics*, 2018, *57*, 221–235.

Brinch, Christian N, Magne Mogstad, and Matthew Wiswall, "Beyond LATE with a discrete instrument," *Journal of Political Economy*, 2017, *125* (4), 985–1039.

Buttke, Danielle E., Kanta Sircar, and Colleen Martin, "Exposures to Endocrine-Disrupting Chemicals and Age of Menarche in Adolescent Girls in NHANES (2003–2008)," *Environmental Health Perspectives*, November 2012, *120* (11), 1613–1618.

Carneiro, Pedro, James J Heckman, and Edward J Vytlacil, "Estimating Marginal Returns to Education," *American Economic Review*, 2011, *101* (6), 2754–2781.

Chavarro, J., E. Villamor, J. Narváez, and A. Hoyos, "Socio-demographic predictors of age at menarche in a group of Colombian university women," *Annals of Human Biology*, 2004, *31* (2), 245–257.

Cheong, Ji In, Chang Hun Lee, Jae Hong Park, Byeong Jin Ye, Kyoung Ah Kwon, Young Seok Lee, and Jae Ho Yoo, "The effect of early menarche on the sexual behaviors of Korean female adolescents," *Annals of Pediatric Endocrinology & Metabolism*, 2015, *20* (3), 130–135.

Cornelissen, Thomas, Christian Dustmann, Anna Raute, and Uta Schönberg, "From LATE to MTE: Alternative methods for the evaluation of policy interventions," *Labour Economics*, 2016, *41*, 47–60.

_ , _ , _ , and _ , "Who benefits from universal child care? Estimating marginal returns to early child care attendance," *Journal of Political Economy*, 2018, *126* (6), 2356–2409.

Currie, Janet and Tom Vogl, "Early-life health and adult circumstance in developing countries," *Annu. Rev. Econ.*, 2013, *5* (1), 1–36.

Diaz, Christina J. and Jeremy E. Fiel, "The Effect(s) of Teen Pregnancy: Reconciling Theory, Methods, and Findings," *Demography*, 2016, *53* (1), 85–116.

Elks, Cathy E., John R. B. Perry, Patrick Sulem, Daniel I. Chasman, Nora Franceschini, Chunyan He, Kathryn L. Lunetta, and et al., "Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies," *Nature Genetics*, 2010, *42* (12), 1077–1085.

Felfe, Christina and Rafael Lalive, "Does early child care affect children's development?," *Journal of Public Economics*, 2018, *159*, 33–53.

Field, Erica and Attila Ambrus, "Early Marriage, Age of Menarche, and Female Schooling Attainment in Bangladesh," *Journal of Political Economy*, 2008.

Fletcher, Jason M. and Barbara L. Wolfe, "Education and Labor Market Consequences of Teenage Childbearing Evidence Using the Timing of Pregnancy Outcomes and Community Fixed Effects," *Journal of Human Resources*, 2009, *44* (2), 303–325.

Gajdos, Zofia KZ, Joel N. Hirschhorn, and Mark R. Palmert, "What controls the timing of puberty? An update on progress from genetic investigation," *Current Opinion in Endocrinology, Diabetes and Obesity*, 2009, *16* (1), 16.

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

_ , _ , and Marianne M. Hillemeier, "Does Young Maternal Age Adversely Affect Child Development? Evidence from Cousin Comparisons in the United States," *Population and Development Review*, 1994, *20* (3), 585–609.

Glynn, Judith R., Ndoliwe Kayuni, Sian Floyd, Emmanuel Banda, Monica Francis-Chizororo, Clare Tanton, Anna Molesworth, Joanne Hemmings, Amelia C. Crampin, and Neil French, "Age at Menarche, Schooling, and Sexual Debut in Northern Malawi," *PLOS ONE*, 2010, *5* (12), e15334.

Heckman, James J. and Edward J. Vytlacil, "Local instrumental variables and latent variable models for identifying and bounding treatment effects," *Proceedings of the National Academy of Sciences*, April 1999, *96* (8), 4730–4734.

_ **and** _ , "Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation," in James J. Heckman and Edward E. Leamer, eds., *Handbook of Econometrics*, Vol. 6, Elsevier, 2007, pp. 4779–4874.

_ **and Edward Vytlacil**, "Structural Equations, Treatment Effects, and Econometric Policy Evaluation," *Econometrica*, 2005, *73* (3), 669–738.

Heckman, James, Rodrigo Pinto, and Peter Savelyev, "Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes," *The American economic review*, 2013, *103* (6), 2052–2086.

Heissel, Jennifer A., "Teen Fertility and Siblings' Outcomes Evidence of Family Spillovers Using Matched Samples," *Journal of Human Resources*, 2021, *56* (1), 40–72.

Herrera-Almanza, Catalina and David E. Sahn, "Early Childbearing, School Attainment, and Cognitive Skills: Evidence From Madagascar," *Demography*, 2018, *55* (2), 643–668.

Holmlund, Helena, "Estimating Long-Term Consequences of Teenage Childbearing An Examination of the Siblings Approach," *Journal of Human Resources*, 2005, *XL* (3), 716–743.

Hotz, V Joseph, Charles H Mullin, and Seth G Sanders, "Bounding causal effects using data from a contaminated natural experiment: Analysing the effects of teenage childbearing," *The Review of Economic Studies*, 1997, *64* (4), 575–603.

Hotz, V. Joseph, Susan Williams McElroy, and Seth G. Sanders, "Teenage Childbearing and Its Life Cycle Consequences Exploiting a Natural Experiment," *Journal of Human Resources*, July 2005, *XL* (3), 683–715.

Isaza, Ximena Casas, Oscar Cabrera, Rebecca Reingold, and Daniel Grossman, *Stolen Lives: A Multi-Country Study on the Health Effects of Forced Motherhood on Girls 9–14 Years Old* 2015.

Jansen, Erica C., Oscar F. Herrán, and Eduardo Villamor, "Trends and correlates of age at menarche in Colombia: Results from a nationally representative survey," *Economics and Human Biology*, 2015, *19*, 138–144.

Jr, Joseph J Doyle, "Child protection and child outcomes: Measuring the effects of foster care," *American Economic Review*, 2007, *97* (5), 1583–1610.

Kaprio, J., A. Rimpelä, T. Winter, R. J. Viken, M. Rimpelä, and R. J. Rose, "Common genetic influences on BMI and age at menarche," *Human Biology*, 1995, *67* (5), 739–753.

Kaufmann, Katja Maria, "Understanding the income gradient in college attendance in Mexico: The role of heterogeneity in expected returns," *Quantitative Economics*, 2014, *5* (3), 583–630.

Klepinger, Daniel, Shelly Lundberg, and Robert Plotnick, "How Does Adolescent Fertility Affect the Human Capital and Wages of Young Women?," *The Journal of Human Resources*, 1999, *34* (3), 421–448.

Lang, Kevin and Russell Weinstein, "The Consequences of Teenage Childbearing before Roe v. Wade," *American Economic Journal: Applied Economics*, 2015, *7* (4), 169–197.

Levine, David I. and Gary Painter, "The Schooling Costs of Teenage Out-of-Wedlock Childbearing: Analysis with a Within-School Propensity-Score-Matching Estimator," *The Review of Economics and Statistics*, 2003, *85* (4), 884–900.

Liu, Shenglong, Tianyu Jin, Meng Li, and Shaojie Zhou, "Fertility Policy Adjustments and Female Labor Supply: Estimation of Marginal Treatment Effect Using Chinese Census Data," *Journal of Human Resources*, 2022, p. 0620.

Marván, María Luisa, Rosa Lilia Castillo-López, Diana Donají del Callejo-Canal, Margarita Edith Canal-Martínez, and Alejandra Núñez de la Mora, "Secular trends in age at menarche in 20th century Mexico: Differences by ethnicity, area of residency, and socioeconomic status," *American Journal of Human Biology*, 2020, *32* (6), e23404.

Morris, Danielle H., Michael E. Jones, Minouk J. Schoemaker, Alan Ashworth, and Anthony J. Swerdlow, "Familial concordance for age at menarche: analyses from the Breakthrough Generations Study," *Paediatric and Perinatal Epidemiology*, 2011, *25* (3), 306–311.

Olea, José Luis Montiel and Carolin Pflueger, "A Robust Test for Weak Instruments," *Journal of Business & Economic Statistics*, July 2013, *31* (3), 358–369.

Parent, Anne-Simone, Gregory Rasier, Arlette Gerard, Sabine Heger, Christian Roth, Claudio Mastronardi, Heike Jung, Sergio R. Ojeda, and Jean-Pierre Bourguignon, "Early onset of puberty: tracking genetic and environmental factors," *Hormone Research*, 2005, *64 Suppl 2*, 41–47.

_ , Grete Teilmann, Anders Juul, Niels E. Skakkebaek, Jorma Toppari, and Jean-Pierre Bourguignon, "The timing of normal puberty and the age limits of sexual precocity: variations around the world, secular trends, and changes after migration," *Endocrine Reviews*, 2003, *24* (5), 668–693.

Ribar, David C., "Teenage Fertility and High School Completion," *The Review of Economics and Statistics*, 1994, *76* (3), 413–424.

Romans, S. E., J. M. Martin, K. Gendall, and G. P. Herbison, "Age of menarche: the role of some psychosocial factors," *Psychological Medicine*, July 2003, *33* (5), 933–939.

Rosenzweig, Mark R. and Kenneth I. Wolpin, "Sisters, Siblings, and Mothers: The Effect of Teen-Age Childbearing on Birth Outcomes in a Dynamic Family Context," *Econometrica*, 1995, *63* (2), 303–326.

Santelli, John S., Xiaoyu Song, Samantha Garbers, Vinit Sharma, and Russell M. Viner, "Global Trends in Adolescent Fertility, 1990-2012, in Relation to National Wealth, Income Inequalities, and Educational Expenditures," *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine*, 2017, *60* (2), 161–168.

Tither, Jacqueline M. and Bruce J. Ellis, "Impact of fathers on daughters' age at menarche: a genetically and environmentally controlled sibling study," *Developmental Psychology*, 2008, *44* (5), 1409–1420.

Towne, Bradford, Stefan A. Czerwinski, Ellen W. Demerath, John Blangero, Alex F. Roche, and Roger M. Siervogel, "Heritability of age at menarche in girls from the Fels Longitudinal Study," *American Journal of Physical Anthropology*, 2005, *128* (1), 210–219.

Turley, Ruth N. López, "Are Children of Young Mothers Disadvantaged Because of Their Mother's Age or Family Background?," *Child Development*, 2003, *74* (2), 465–474.

UNFPA, "Motherhood in Childhood: The Untold Story," 2022.

United Nations Population Division, "World Population Prospects, Data file," 2022.

Vidmar, Suzanna, John Carlin, Kylie Hesketh, and Tim Cole, "Standardizing anthropometric measures in children and adolescents with new functions for egen," *The Stata Journal*, 2004, *4* (1), 50–55.

Villamor, Eduardo, Constanza Marín, Mercedes Mora-Plazas, and Henry Oliveros, "Micronutrient status in middle childhood and age at menarche: results from the Bogotá School Children Cohort," *The British Journal of Nutrition*, 2017, *118* (12), 1097–1105.

Wronka, Iwona and Romana Pawlińska-Chmara, "Menarcheal age and socio-economic factors in Poland," *Annals of Human Biology*, 2005, *32* (5), 630–638.

Yakusheva, Olga and Jason Fletcher, "Learning from teen childbearing experiences of close friends: Evidence using miscarriages as a natural experiment," *Review of Economics and Statistics*, 2015, *97* (1), 29–43.

A Relevant Literature

Appendix Table A1. Relevant Literature

Study	Identification Method	Focus	Results (+ / - /null)
	High-income countries		
Ribar (1994)	IV (age of menarche)	Mothers	null & +
Klepinger et al. (1999)	IV (age of menarche, community indicators)	Mothers	-
Hotz et al. (2005)	IV (Miscarriages)	Mothers	small -
Fletcher and Wolfe (2009)	IV (Miscarriage, Community FE), OLS	Mothers	-
Ashcraft et al. (2013)	IV (Miscarriage), OLS	Mothers	small -
Geronimus and Korenman (1992)	Family FE (sisters comparison)	Mothers	null
Holmlund (2005)	Family FE (sisters comparison)	Mothers	-
Levine and Painter (2003)	Within-school PSM	Mothers	-
Geronimus et al. (1994)	Family FE (cousins comparisons)	Children	null
Rosenzweig and Wolpin (1995)	Family FE and FE-IV (cousins comparisons)	Children	null
López Turley (2003)	Family FE (cousins comparisons)	Children	null
Aizer et al. (2022)	Family FE (cousins comparisons)	Children	small -
	Low and middle-income countries		
Azevedo et al. (2012)	OLS (miscarriage sample)	Mothers	+
Herrera-Almanza and Sahn (2018)	IV (Contraception access)	Mothers	-
Branson and Byker (2018)	Diff-n-Diff (Policy variation)	Mothers, Children	-

B Predicting the age at menarche

For each woman in the study, I initially predict the likelihood of teenage births (occurrences between ages 15 to 19) using 38,770 observations. This sample comprises women aged 15 and older who provided complete information on several key variables: whether they live in rural or urban areas, their region of residence (Coastal, Highlands, Amazon, or the Galapagos Islands), birth date (year and month), race/ethnicity, and height, categorized into eight groups. These variables are presumed to be unaffected by a woman's teen motherhood status. The calculated probability of teen birth is used in three ways. First, it serves as an input for another random forest model predicting the age of menarche. Second, it acts as a flexible control in estimating both the propensity score and marginal treatments effect. Third, it aids in assessing the direct impact of age at menarche on the outcomes. Specifically, this involves analyzing the direct effect of menarche age in non-teen mothers or women without children with a predicted probability of teen birth under 50%, based on the model. Focusing on women with a lower likelihood of teen birth, aims to minimize any bias arising from evaluating the direct effect of age at menarche on women potentially induced out of teen motherhood due to her age at menarche.

I use a second random forest model to predict age at menarche, limiting the scope to 27,466 women who are either non-teen mothers or have not children. This restriction allows me to incorporate variables crucial for the analysis but potentially influenced by teen motherhood. This model includes the same variables as mentioned earlier, plus additional variables reflecting asset ownership, access to public services, and household conditions, as detailed in table C1.

Both random forest models follow standard estimation procedures. Each sample is randomly divided, allocating 70% for training and 30% for validation to ensure model accuracy and guard against overfitting. During the training phase, a randomized grid search and k-fold cross-validation are employed for optimal hyper-parameter selection.

Appendix Table B1. **Extra variables used to age at menarche prediction**

Classification	**Variables**
Household Conditions	type of access road, type of house, roof material, wall material, floor material, roof condition, wall condition, floor condition, number of people per room
Access to Services	Fuel used to cook, source of electricity, type of garbage management, shower availability/exclusivity, toilet availability, source of water, access to tubed water, internet availability, cable Tv availability
Household assets, whether the household owns at least on	fridge, computer, washing machine, blender, microwave, iron, TV, water heater, car
Other	asset index obtained as the first component of all the variables above

C Does Teen Childbearing affect the probability of observing a firstborn?

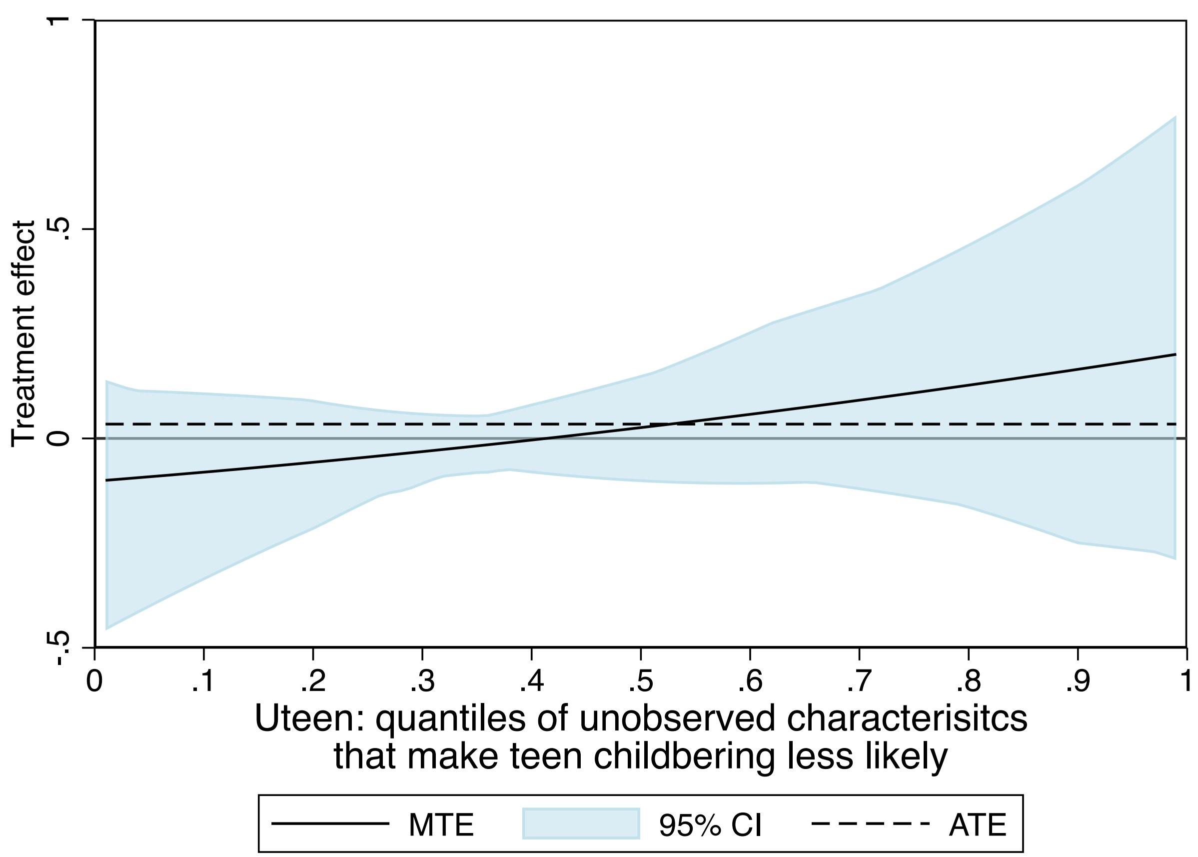
In this section, I use the same models described in the main body of the paper to check if teen childbearing affects the probability of observing the firstborn in the same household as their mother. Table C1 presents OLS and 2SLS estimates of the estimated effects of teen childbearing on a dummy variable equal to one if the firstborn reported in the mother's birth history lives at the same household as the mother at the time of the survey. The first two columns show the estimates for the entire firstborn sample. The OLS estimate in the first column suggest that the first born children of teen mothers are 5% more likely to not be in the same household as their mother at the time of the survey. The 2SLS estimate suggest that firstborns are percent more likely to not live in the same houshold as the mother.

Columns 3 and 4 include only children 14 or younger. The OLS estimate of the third column suggest that that the first born children of teen mothers are 2% more likely to not be in the same household as their mother. That estimate is half the size of the first column. The 2SLS estimate suggest that the first born children of teen mothers are 1.7% more likely to not be in the same household as their mother But this estimate is not significantly different from zero. Figure C1 shows there unobserved heterogeneity is not a concern for this outcome. The aggregated treatment effect parameters (not presented) were also not statistically significant.

Appendix Table C1. **Effects of Teen childbearing on prob. of observing firstborn**

	OLS All firstborn	2SLS All firstborn	OLS Firstborn $age < 15$	2SLS Firstborn $age < 15$
teen mother=1	0.050***	0.080**	0.024***	0.017
	(0.004)	(0.03)	(0.003)	(0.022)
Observations	28148	27957	18060	18060

Appendix Figure C1. Marginal Treatment Effects prob. of observing firstborn age<14



D Robustness Checks

Appendix Figure D1. Alternative Specifications