The Precocious Period: Impact of Early Menarche on Schooling in India

Madhulika Khanna *

February 17, 2021

Abstract

This paper studies the impact of early onset of menstruation on school enrollment in India, where the onset of menstruation (menarche) has immense socio-cultural importance, and menarche marks a child's transition into womanhood. Estimates based on a difference-in-differences model show that starting menses before age twelve decreases school enrollment by 13%. While understanding the role of menarche—a universal biological event for all women—in determining education is a critical end in itself, it also has implications for the effect of early nutrition on girls' well-being. Better-nourished girls typically reach menarche earlier, and if menarche impedes schooling, their gains due to better nutrition may be undercut. The effect of menarche on school enrollment is stronger if girls live in communities where the perceived safety among children is low or if they belong to social groups with restrictive gender norms. Conversely, dropout rates are lower in communities with higher expected wages for female-dominated professions. Jointly, these findings suggest that the interaction between cultural norms and the local economic opportunities shapes female education decisions in response to menses.

JEL Classification: J16; I24; O12

Keywords: School Enrollment; Education; Gender; Social Norms; India

*Madhulika Khanna: Georgetown University, 3700 O St NW, Washington, DC 20057, mk1469@georgetown.edu.
Acknowledgements: I am very grateful to Garance Genicot and Martin Ravallion for their guidance over the course of this project. This paper also benefited from the helpful comments from Ashish Aggarwal, Sarah Baird, Brad Barham, Anja Benshaul-Tolonen, Becka Brolinson, Caitlin Brown, David Eil, Deepti Goel, Maria Hernandez, Pamela Jakiela, Shareen Joshi, Nishtha Kochhar, Catherine Maclean, Juan Margitic, Umberto Muratori, Deniz Sanin, Dario Sansone, and Andrew Zeitlin as well as seminar participants at Georgetown University, Towson University, Colgate University, and conference participants at the Sustainability and Development Conference, the Midwest Economic Association Conference, the Washington Area Development Economics Symposium, the Midwest International Economic Development Conference, the Georgetown Center for Economic Research Conference, and the Southern Economic Association Conference. I have also benefited from fruitful discussions with several seminar speakers and co-workers at Georgetown University. The usual caveats apply.

"*It is a vicious circle. Women are deprived of rights because of their lack of education, and their lack of education results from their lack of rights.*"

-Leo Tolstoy (in Anna Karenina)

1 Introduction

Despite universal primary school enrollment in India, many adolescent girls drop out after completing eight years of compulsory schooling (ASER 2018). Consequently, a notable gender gap in school enrollment emerges and grows during teenage years (Muralidharan and Prakash 2017). The existing literature has posited many explanations for the emergence of a gender gap in schooling in this setting, including the burden of housework (Sundaram and Vanneman 2008), lower returns to female education (Chamarbagwala 2008), and the social barriers that impede access to schooling due to caste-based allocation of schooling infrastructure (Jacoby and Mansuri 2011). In this paper, I identify a biological driver of the decrease in school enrollment during adolescence: the onset of menstruation (menarche).

This paper is set in India, where menarche completely alters the norms of a girl's engagement with her family and her community (Seymour 1999). Parents mark their daughter's menarche with rituals and celebrations, and in some communities, they seclude their daughter within the house as part of these rituals (Dharmalingam 1994). Menstruating women are excluded from participating in religious practices or even entering the kitchen (Mahon and Fernandes 2010; Garg, Sharma, and Sahay 2001). Menarche also signifies the beginning of a critical period in her life when "she has acquired the capability to reproduce but she has no authority to do so" and her "vulnerability is at its peak" (Dube 1988). Once the girl reaches menarche, her family's honor is associated with its capacity to maintain their daughter's virginity. To ensure their daughter's safety, parents often restrict her mobility.

In this paper, I study if the familial and societal responses to menarche in this context affect girls' school enrollment and characterize the relationship between early menarche and schooling in a conceptual framework where the parents decide on the optimal years of education and age at marriage for their daughter. The empirical evidence is consistent with the predictions of this framework and shows that an early onset of menstruation decreases school enrollment.

The primary empirical challenge of estimating the impact of early menarche is that girls experience many transitions, such as the transition to secondary schooling and increased burden of housework, during this time period. Any association between early menarche and enrollment also reflects these confounding changes. To isolate the causal impact of early menarche, I exploit the variation in the timing of menarche of a cohort of 1,000 children in the state of Andhra Pradesh (now Andhra Pradesh and Telangana) between two survey rounds. More specifically, I use the enrollment status of those girls who had not reached menarche by age twelve as the counterfactual for those who experienced an earlier menarche along in an individual fixed effects model.[1] The main result of

1. In this sample, a little more than a quarter of the surveyed girls had reached menarche before age twelve.

this paper is that reaching menarche before age twelve causes a 14.3% decrease in school enrollment rate (equivalent to a 13% point decrease over the counterfactual enrollment rate of 90.6%).

Using the children's self-reports of their aspirations, I also show that reaching menarche before twelve increases a girl's desire to become a full-time parent or housewife.[2] While the construction of identity is critical as an outcome by itself, it can have consequences beyond the primary outcome studied in this paper. For instance, a woman's sense of self shapes her labor market outcomes (Bertrand 2011; Olivetti, Patacchini, and Zenou 2020). I find that early menarche also has effects that persist well into early adulthood. Consistent with the existing literature (Desai and Andrist 2010), I show that girls who reach menarche before twelve are 11% more likely to be married at twenty-two, about a year younger than their peers.

Finally, girls are more likely to drop out of school after menarche if they reside in communities with a lower average perception of safety among children or if they belong to castes with restrictive gender norms. By locating an increase in dropout rates at menarche, these results suggest that gender-specific social norms and concerns for safety are amplified after this one event. As such, policy responses to higher dropout rates among adolescent girls should address their safety concerns and challenge gender-based cultural norms. At the same time, I find that the negative impact of menarche on the enrollment rate is attenuated in communities with higher expected wages for female-dominated professions. By and large, the findings of the paper are consistent with the view of Jayachandran (2015) that it is the interaction between cultural norms and local economic opportunities that drives female educational attainment.

The existing evidence on the effect of menstruation in general and menarche in particular is far from conclusive. Oster and Thornton (2011) find that menstruation accounts for only 0.4 missed school days in a 180-day school year in Nepal. On the other hand, studies that use self-reported data and case studies argue that menstruation is a critical factor that keeps girls out of classrooms. For instance, in a study conducted in three states in India, over half of the interviewed adolescent girls reported that menstruation was a significant barrier to attending school (Sivakami et al. 2019). Similar conclusions are evident from the studies based in other contexts, including Northern Tanzania (Sommer 2010) and Western Kenya (Mason et al. 2013). The strong positive response in pubescent girls' enrollment to the construction of sex-separate toilets suggests that privacy and safety concerns during this age are critical factors in a girl's enrollment decision (Adukia 2017). However, none of these studies establish an explicit causal link between menarche and school enrollment, as done in this paper. In line with the evidence on the limited association between menstrual days and school attendance in Oster and Thornton (2011), I show that if enrolled at twelve, girls' menarche status does not affect the time they spend at school. Taken together, these results suggest that while menarche affects enrollment, it may have limited effect on school attendance.

When combined with different strands of literature on menarche, nutrition, and later-life outcomes, the key result of this paper implies that the path from nutritional gains in childhood to later

2. This may be the direct effect of menarche or the effect of altered norms of interacting with her family and community after menarche.

life outcomes might be more complex for girls. First, it is well-established that better-nourished girls typically reach menarche before other girls in their cohort (see Villamor and Jansen (2016) for a review on correlates of menarcheal timing). In this sample, too, taller and heavier girls reach menarche before others in their cohort. Next, there is also considerable evidence that improvements in childhood nutrition improve later life outcomes through complementarities between nutrition and learning across ages (see for instance Cunha and Heckman (2007), Attanasio and (2015), and Currie and Almond (2011) for a review). Therefore, the timing of exposure to social programs may have additional effects due to the induced changes in their biology. Relatedly, Barham, Macours, and Maluccio (2018) find that when girls joined a conditional cash transfer program in Nicaragua between eleven and fourteen, they had earlier menarche, higher fertility, and lower labor force participation than those who joined the program earlier.

Finally, this paper's findings are also relevant for the literature that uses the timing of menarche as an instrument for the age at which a woman gets married. This strategy was first used by Field and Ambrus (2008), who examine cohorts born between 1951 and 1970 in Bangladesh. Subsequently, many papers have followed this strategy to explore the effect of early marriage on a variety of policy-relevant outcomes, in (Sekhri and Debnath 2014; Chari et al. 2017; Roychowdhury and Dhamija 2018; Dhamija and Roychowdhury 2020) and outside India (Sunder 2017; Cantet 2019; Hicks and Hicks 2019) for younger cohorts. In Field and Ambrus (2008)'s sample, 70% of the interviewed women had married within two years after their menarche. Since then, the gap between the age at menarche and age at marriage has, if anything, increased over time, allowing for the possibility that menarche may affect margins other than the marriage decision. By establishing an explicit causal relationship between early menarche and education using difference-in-differences and individual fixed effects, this paper goes beyond the reduced form association between menarche and education shown in this literature.[3] Lastly, using survey data with adolescents ensures that estimates in this paper are due to recall bias in retrospective information on menarche collected from adult married women stemming from various sources, including leaving school.

The rest of the paper is organized as follows. Section 2 describes the socio-cultural norms surrounding menarche in India, and Section 3 presents the underlying conceptual framework. Section 4 describes the data followed by a discussion on the empirical strategy in Section 5. Section 6 presents the results describing the impact of early menarche on enrollment and discusses robustness checks and heterogeneity. Section 7 examines the relationship between age at menarche and other critical outcomes. Section 8 discusses the results and conclusion.

2 Context

Despite being a ubiquitous and regularly occurring experience for women, menstruation is surrounded by various myths and taboos. This is especially true in South Asia where these myths and taboos continue to define women's day-to-day activities. The origin of mythical narratives

3. In a companion note, I show that the two-stage least squares estimates under plausible violations in the exclusion restriction assumption are wide and contain zero (Khanna, 2021).

around menstruation can be traced back to *Rig Veda* which was composed between 1500 and 1200 BC (Chawla 1994). When Lord *Indra* slayed a high caste demon, *Vritra*, part of *Indra's* guilt was assumed by womankind. His guilt appears every month in the form of their menstrual flow, and menstruation is considered ritually impure.

The notion of menstruation's ritual impurity is most evident in taboos and cultural practices that confine the scope of women's daily lives. For instance, women are not supposed to enter the kitchen or the prayer room during their menses. Some dietary restrictions, such as the prohibition of sour food, are also followed during menstruation (Garg, Sharma, and Sahay 2001; Garg and Anand 2015). In some communities, young girls are instructed to stay away from water sources during their menses since they will pollute it (Mahon and Fernandes 2010).

Menarche is believed to mark a girl's transition into womanhood in South Asia. Now that she is a woman, her parents must plan her marriage. Delayed marriage signifies the failure of a girl's parents to discharge their duties (ICRW and Plan Asia 2013). According to the religious scriptures, by performing the act of *kanyadan*, or handing over his daughter to her husband, the bride's father earns good virtues. However, if the girl is perceived to be "unrestrained", it would not be easy to arrange her marriage (Dube 1988). Parents have to safeguard her virginity until the day of her wedding. Now that she is a woman, she is more vulnerable to sexual harassment and is observed more carefully by the community members (Singh and Vennam 2016). Her mobility and the extent and the scope of her interactions are restricted (Seymour 1999). To further unpack the relevance of the socio-cultural role of menarche in explaining the key results of this paper, the rest of this section discusses evidence from Vennam and Komanduri (2009), a report based on qualitative data collected from the cohort under consideration. Vennam and Komanduri (2009) use extensive qualitative data collected from children from every sampling site to discuss the critical transitions experienced by children between the ages of eight and fifteen.[4]

Families celebrate their daughters' menarche by hosting a meal for their relatives and friends within the community. Girls are kept at home for five to seven days after menarche. They are not allowed to touch any person during this period. Some families continue secluding their daughters during her menses period, even after menarche. After menarche, girls in Muslim communities may wear a *burqa*, and in Hindu communities, long skirts and *saris* instead of their usual clothes. They are no longer allowed to go out alone, even for short distances. Mothers explain that if their daughters do not adhere to these restrictions, the community might disapprove, and this will affect their marriage prospects. Importantly, parents are preoccupied with preparing their daughters for marriage and with approval from the local community.

Restrictions on mobility as a result of menarche are entwined with the schooling decision. One girl explains that now she wears a *burqa* while going to school to avoid unwanted attention. Girls report being *eve-teased*, a colloquial term for public sexual harassment, on their way to school. In sum, menarche characterizes tremendous changes in a girl's life that go well beyond the biological changes.

4. See Section 4 for a detailed discussion of the sampling strategy.

3 Conceptual Framework

Drawing on the discussion on the socio-cultural role of menarche in a girl's life in Section 2, this section describes a simple framework of how the onset of menstruation can affect her education and age at marriage. This framework has two goals. First, it describes a setup under which menarche affects schooling. Second, it derives testable implications for the heterogeneous impact of menarche on schooling. This framework has to be qualified in several respects. The choice problem is static and conceptualized in a unitary household setup.

Formally, I model parents' choice of length of schooling, and consequently, the age at dropout (e) and age at marriage (a). The age at menarche (m) is a component of this decision-making process because menarche increases the cost of attending school for the girl. This formulation draws from the discussion in Section 2 and captures the safety and the social cost of having an unmarried daughter who she goes to school.

Parents' utility maximization problem is as follows:

$$\max_{\{e,a\}} \quad \omega b(e) + v(a) - c(e, max\{a - m, 0\})$$

$$\text{subject to}$$

$$(1) \quad a \geq m$$

$$(2) \quad a \geq e$$

where $\omega b(e)$ is the direct benefit to the parents from their daughter's education and her time at her natal home before marriage, which is strictly increasing ($b_1 > 0$) and concave ($b_{11} < 0$). ω captures the future monetary returns the benefits from education. Next, $v(a)$ denotes the gain from delaying the marriage, which is increasing ($v_1 > 0$) up to some age, $\bar{a}$, and then starts decreasing; delaying marriage beyond some age would entail negative returns.[5] Further, $v(a)$ is concave: ($v_{11} < 0$). Education is costly, so is delaying marriage after attaining menarche (m): $c(e, max\{a-m, 0\})$. The cost function is increasing ($c_1 > 0$ and $c_2 > 0$) and convex ($c_{11} > 0$ and $c_{22} > 0$). One can think of $max\{a-m, 0\}$ as the *reputation*, or the *bad reputation*, that the girl may build once she crosses the threshold of womanhood and still lives in her natal home. As discussed above, traveling to school and socializing with boys (arguably at school) amplifies the reputation cost after menarche: $c_{12} > 0$.

Menarche is a lower bound on the age at marriage as described by the constraint $a \geq m$. In addition, a girl cannot continue her education after marriage; this feature of the choice problem is described by the second constraint, that is, $a \geq e$.

5. Rao (1993) discusses that there is a pressure on women to get married within an acceptable age range, and it becomes increasingly difficult to find a match once they are past a certain age. In this setup, after $\bar{a}$, the gain from delaying the marriage is only by allowing the daughter to remain in school and she will be married off as soon as she drops out.

Consider the interior solution where $a > m$ and $a > e$.[6] [7] The relationship between years of education and age at menarche is:[8]

$$\frac{\partial e(m)}{\partial m} = \frac{-v_{11}c_{12}}{(\omega b_{11} - c_{11})(v_{11} - c_{12}) - c_{12}^2} > 0. \tag{3.1}$$

Prediction 1: Completed years of schooling is increasing in the age at menarche. Therefore, girls are more likely to drop out of school after menarche.

Note that $\frac{\partial e(m)}{\partial m}$ is higher if c_{12} increases. More intuitively, the safety concerns and reputation costs are higher for those girls who reach menarche before their peers since they enter they cross the believed threshold of womanhood sooner when still not married. Although the primary contribution of the paper is to estimate the effect of early menarche on school enrollment, I also discuss suggestive evidence on the effect of later menarche.

Prediction 2: The effect of the onset of menstruation on schooling, if any, is smaller for girls who reach menarche later than their peers.

This framework can also be used to derive implications for heterogeneity in the impact of early menarche on school enrollment along some policy-relevant dimensions. Consider two communities A and B such that the complementarity between education and delay in marriage beyond menarche is stronger in community A. In other words, the safety and social costs of attending school after menarche are higher in community A. Note that $\partial e(m)/\partial m|_A > \partial e(m)\partial m|_B$; the relationship between menarche and education is stronger in community A than in community B.

Prediction 3: The decrease in enrollment due to early menarche is more pronounced if the safety or reputation costs are more pertinent.

Next, consider the returns to education realized by parents. In this setup, ω captures the monetary returns to daughter's education. Consider ω_A and ω_B such that $\omega_A > \omega_B$; higher ω denotes a higher returns to education, including the returns realized in the labor market. Note that $\partial e(m; \omega_A)/\partial m < \partial e(m; \omega_B)/\partial m$, that is, the relationship between the age at menarche and education is weaker when the returns to education are higher. Thus, the prediction for the heterogeneity

6. Consider the case wherein the first constraint ($a \geq m$) is binding. In this case, as the age at menarche increases, the binding constraint on the age of marriage is tightened, increasing the age at marriage. If the girl is already out of school, her education is not affected. However, if she is currently in school, she will drop out since $a \geq e$. Indeed, this is the case discussed in Field and Ambrus (2008). However, for the state of Andhra Pradesh, the median length of time between when the girl reaches menarche and when she gets married is five years, suggesting that the constraint is not binding (see Section 4 for details of data used in the study.).

7. Consider when the second constraint ($a \geq e$) is binding. In this case, the optimal level of education is more than the age of marriage. Higher age at menarche will increase the age at marriage, and therefore education; this is an additional channel (beyond safety and privacy) through which menarche could affect education. Predictions on the effect of menarche that are described below will still hold.

8. Second-order conditions associated with the interior solution are: (1) $\omega b_{11} - c_{11} < 0$, (2) $v_{11} - c_{22} < 0$, and (3) $(\omega b_{11} - c_{11})(v_{11} - c_{22}) - (c_{12})^2 > 0$. These conditions hold under the assumptions of the framework described above.

in the impact of early menarche along returns to education is as follows:

Prediction 4: The decrease in enrollment due to early menarche is abated if the potential returns to education are higher.

Lastly, the relationship between age at marriage and age at menarche is:

$$\frac{\partial a(m)}{\partial m} = \frac{-c_{22}(\omega b_{11} - c_{11}) - c_{12}^2}{(\omega b_{11} - c_{11})(v_{11} - c_{12}) - c_{12}^2} > 0. \tag{3.2}$$

To summarize, this framework shows that as long as a girl experiences substantial safety and reputation costs of attending school after menarche, earlier menarche decreases school enrollment. The effect of a later onset of menstruation, if any, will be smaller. Therefore, this framework provides an intuitive and parsimonious setup to interpret the main results of this paper.

4 Data

This paper uses data collected by the Young Lives longitudinal study that follows two cohorts of children in Andhra Pradesh over the course of their childhood, and focuses on the older cohort that was born in 1994-95.[9] Within each sampling site, households were screened to compile a list of eligible children (aged eight in 2001). From this sampling site-level list, a sample of 50 children born between January 1994 and June 1995 was randomly selected.

The Young Lives study tracks children in the same cohort over their childhood, adolescence, and early adulthood. Therefore, it is ideal for implementing the identification strategy used in this paper given that it exploits the variation in the timing of menarche.[10] The Young Lives study also collected detailed socioeconomic and demographic data from household, child and community surveys. In particular, primary care-givers were asked about consumption expenditure, economic shocks and their perceptions of children's health status. Children took standardized tests that measure their cognitive ability, and were also interviewed about important aspects of their experiences like schooling, time use, aspirations, social networks and their perceptions of the social milieu. Children were also asked about their marital outcomes in later survey rounds. Community surveys were conducted through focus-group discussions and collected information on topics like local wages and public services. The Young Lives study also collected data on children's physical development through on-site anthropometric measurements. Importantly for the purpose of this paper, the second survey round, during which children were about twelve years old, asked girls if they had started

9. Andhra Pradesh was split into Andhra Pradesh and Telangana in 2014. Appendix A describes the administrative and the socio-cultural context of the state of Andhra Pradesh.

10. The other longitudinal surveys in India (such as the Rural Economic and Demographic Survey and the India Human Development Survey (IHDS)) do not collect information on the timing of the menarche of the school-going girls in the households. Data on age at menarche, whenever collected (for instance, in the IHDS and the National Family Health Survey), is retrospective and endogenous to other life events (including leaving school).

menstruating yet, and, if so, at what age.[11]

The first round of the study was conducted in 2002 when these children were eight years old. The next four rounds of data were collected in 2006, 2009, 2013, and 2016. Attrition was low: the first four rounds tracked 94.4% of the children, and all the five rounds tracked 92.4% of the children. The rate of attrition did not differ across girls and boys, or across girls by their menarche status.

There are about 1,000 children in the cohort studied in the paper. Male and female are nearly equally represented in the sample: 49% of the children are boys, and 51% are girls. About one-fourth of the data are collected from urban areas, and the remaining from rural areas. Backward castes constitute about 47% of the sample, while scheduled castes are 21%, scheduled tribes, 11%, and Muslims, 6% of the sample. The remaining sample represents the upper caste households in the state. About 35% of the children are first-borns.

Although the primary variable of interest is the enrollment status of the children, I also use their test scores for supplementary analysis. Peabody Picture Vocabulary Tests (PPVT) and Mathematics tests were administered in the second and third rounds of the survey.[12] [13] For the analysis purpose, I normalize children's raw test scores to the distribution of scores in each survey round.

5 Empirical Specification

To identify the effect of early menarche on enrollment, I exploit the variation in the timing of menarche within a cohort. The information on the timing of menarche is taken from the second survey round when the typical cohort age was twelve. A little more than a quarter of girls had started menstruating by this time.[14] For the rest of the paper, these girls are referred as the "the early menarche group", and the remaining girls are referred as "the late menarche group".[15]

I use a difference-in-differences strategy that compares two groups of children across the first and the second survey rounds, that is, when they were eight and twelve, respectively.[16] The specification for this comparison is:

11. Given that menstruation is a taboo in India, this variable may not be accurate. There are two things to note here. First, the Young Lives study extensively piloted the survey and amended those questions that fieldworkers thought were culturally inappropriate (Young Lives 2017). Second, the proportion of girls who reach menarche before age twelve in these data is very similar to the number when I use a larger data set representative of the state of Andhra Pradesh (India Human Development Survey, 2011-12), where 25% of the women reached menarche before age twelve, as opposed to 26% in the study sample (Figure B.1).

12. The PPVT test is designed to measure receptive vocabulary. In a PPVT test, an examiner shows a set of four images simultaneously and asks the children to select the image that best represents the word spoken by the examiner.

13. These tests were administered to all the children at their homes during the interview, and not only to those children who were in school at the time of the survey.

14. Recall data for age at menarche is missing for a small sub-sample of girls. Further, for some girls, the recall age of when they started menstruating is higher than the precisely measured current age. Therefore, I use an indicator variable to define early menarche, and therefore, the treatment group.

15. I use the second round of the nationally-representative India Human Development Survey to confirmed that starting menses before twelve indeed denotes early menarche: only 25% of the interviewed women in Andhra Pradesh reached menarche by this age.

16. There are three relevant groups of children: girls in the early menarche group, girls in the late menarche group and boys.

$$y_{it} = \beta_0 + \beta_2 Round2_{it} + \beta_2 EarlyMenarche_i Round2_{it} + \gamma X_{it} + \delta_i + \epsilon_{it}, \tag{5.1}$$

where y_{it} is the enrollment status of child i in survey round t. $EarlyMenarche_i$ is an indicator equal to one if child i reached menarche before twelve, and $Round2_{it}$ is an indicator equal to one when we consider data from the second survey round. X_{it} is a vector of child-specific time-varying characteristics, and δ_i denote individual fixed effects. ϵ_{it} is a conditionally-mean-zero error term. Since a sampling site is also the primary sampling unit, standard errors are clustered at the sampling site-level. To estimate the impact of early menarche, I compare girls in the early menarche groups to girls in the late menarche group and β_2 is the difference-in-differences estimator.

Individual fixed effects capture the local region's unobserved characteristics, including cultural practices and the quality of schooling. Including individual fixed effects also alleviates the concerns of the effect of climate on age at menarche (Sohn 2016). Causal identification based on difference-in-differences strategy requires that the timing of menarche is conditionally independent of those time-varying factors that influence enrollment. If the timing of menarche solely depended on random genetic variation, the identification based on a difference-in-differences strategy is valid. Kaprio et al. (1995) show that the correlation between the ages at menarche is 0.75 for monozygotic (identical) twin pairs but only 0.31 for dizygotic (fraternal) twin pairs. Using a seventy-four years-long panel study from the United States, Towne et al. (2005) show that almost half of the variation in the timing of menarche is due to genetic factors. However, the pre-eminence of genetic factors in explaining the variation in age at menarche is an established fact only for developed countries. In South Asia, skeletal height, weight, physical activity, and socioeconomic status are correlated with age at menarche (Bagga, Kulkarni, et al. 2000; Chowdhury et al. 2000; Dambhare, Wagh, and Dudhe 2012; Rah et al. 2009). Therefore, in addition to individual fixed effects, I include household size, wealth index, sex of the head of the household, number of older siblings, number of younger siblings, number of older brothers, and number of older sisters as child-level time-varying controls.

Table 1 summarizes children's characteristics at age eight. There are some statistically significant differences across girls in the early and late menarche groups. First, girls in the early menarche group are taller and have higher BMI than girls in the late menarche group and boys. This difference corroborates the evidence on the relationship between age at menarche and nutritional status. Note that better nutrition is typically associated with higher school enrollment (Nandi et al. 2015). Therefore, if anything, estimates from a difference-in-differences model without individual fixed effects will underestimate the true effect of early menarche. Second, in line with the son-biased fertility stopping rule documented in Jayachandran and Pande (2017), boys have fewer younger siblings and more older sisters. Girls in the late menarche group also have more older sisters than girls in the early menarche group. While other differences across girls in the early and the late menarche group are not statistically significant, overall, girls in the early menarche group come from smaller, richer, and less vulnerable households, all characteristics associated with a higher likelihood of remaining in school.

For further exploratory analysis, I estimate the effect of early menarche for different sub-groups of children. The specification for such analysis is:

$$y_{it} = \beta_0 + \beta_1 Round2_{it} + \sum_h \beta_{2h} EarlyMenarche_{ic} Round2_{ict} H_i^h + \gamma X_{it} + \delta_i + \epsilon_{it}, \tag{5.2}$$

where all variables are defined as in equation 5.1, and H_i^h is an indicator equal to one if child i belongs to the sub-group being considered, where h is sub-group. Here, β_{2h} is the effect of early menarche on sub-group h.

The first survey round did not collect data on some important aspects of children's experiences. For these outcomes (such as safety perceptions, time use, etc.), I estimate a contemporaneous regression while augmenting the set of child-level control variables and including sampling-site fixed effects.[17] For instance, when time use of girls in the early and the late menarche groups are compared at age twelve, the specification is:

$$y_{ic} = \beta_0 + \beta_1 EarlyMenarche_{ic} + \gamma X_{ic} + \delta_c + \epsilon_{ic}, \tag{5.3}$$

and β_1 is the intent-to-treat estimator of the impact of menarche if the timing of menarche is as good as random once observables are controlled for. In case this assumption does not hold, β_1 is still a useful description that complements the understanding of menarche's effects on children's experiences. Here, X_{ic} is a vector of child specific characteristics, δ_c is a set of sampling site dummies and ϵ_{ic} is a conditionally-mean-zero error term.

6 Results on Enrollment

6.1 Descriptive Evidence

Figure 1 plots school enrollment by gender. A gap in school enrollment across genders emerges when children are twelve and widens as they grow older. At nineteen, the gender gap in the enrollment rate is 25%. Figure 2 plots these trends again, but children are now classified into three groups: boys, girls in the early menarche group, and girls in the late menarche group. There is no difference in enrollment rates across the three groups when children are about eight. Girls in the early menarche group and boys were still equally likely to be in school at twelve. The enrollment rate for girls in the early menarche group, however, falls visibly between eight and twelve; while the enrollment rate for girls in the late menarche group shows a sharp drop between twelve and fifteen.[18]

The modal reason for dropping out of school between the first and the second survey rounds for boys is "truancy" followed by "needed for work at home". For girls in the late menarche group, the top reasons for dropping out of school are "needed for work at home" and "truancy". For girls in the early menarche group, the modal reason is "other reasons", that is, not one of the reasons

17. Additional control variables include mother's literacy status, an indicator if the child's mother tongue is the local language (Telugu), caste, test scores at age eight, and height and BMI at age eight.

18. This pattern is also replicated with exact age data (Figure B.2, Appendix B).

listed among the extensive prespecified options in the survey. The second most often stated reason is "needed for work at home". Clearly, girls in the early menarche group drop out of school due to reasons that are very different from other children in their cohort.

There are two key patterns in these data that suggest that early menarche induces girls to drop out of school. First, there is a substantial drop in school enrollment rate for girls after they start menstruating. This sharp decrease is not experienced by other children in their cohort. Second, after girls start menstruating, the reasons that they cite for dropping out of school are different from the reasons cited by other children in their cohort.

6.2 Key Results

Results from various specifications of Equation 5.1 are shown in Table 2. Columns (1)-(4) compare girls in the early menarche group to girls in the late menarche group, while columns (5)-(8) compare girls in the early menarche group to boys.

Column (1) considers Equation 5.1 while not including any controls and fixed effects: the estimated impact is -12.3% point. This estimate changes little when child specific controls and sampling site fixed effects are included. The most conservative specification yields an estimate of -13% point, a substantial drop in enrollment rate (Table 2). This drop represents 14.3% of the average enrollment rate for girls in the late menarche group at twelve. The estimated effect of the onset of menstruation when girls in the early menarche group is compared to boys in the most conservative specification is -10.9% point, that is, 12% of the enrollment rate for boys at twelve.

If the gender gap in enrollment at twelve is due to early menarche, enrollment rates for boys and girls in the late menarche group should be similar until the latter group start menstruating. Indeed, in a comparison of the enrollment rates between boys and girls in the late menarche group, the difference-in-differences estimate is small (1.4% point), and not significantly different from zero (Table 3).

I use the simulation exercise described in Athey and Imbens (2017) to study the likelihood that a result of this magnitude could have occurred by chance by generating randomness in early menarche. I re-assign the *treatment status* to a randomly selected set of girls and re-estimate the effect of the placebo *treatment* on school enrollment.[19] Figure B.3 in Appendix B shows the distribution of β_3, the placebo impact of early menarche on school enrollment when girls in the early and the late menarche groups are compared. Reassuringly, the placebo estimates are centered around zero, and only 3% replications estimates reach the estimated result of 13% point. [20]

6.3 Parallel Trends

The validity of the difference-in-differences identification strategy hinges on the assumption of parallel trends: the average change in the control group reflects the counterfactual change in the

19. Following Young (2019), I use 2000 replications.

20. In a comparison between girls in the early menarche group and boys, five replications reach the estimated effect of 10.9% point, confirming that results described in Section 6.2 are not spurious (Figure B.4, Appendix B).

treatment group if there were no treatment. In this context, the assumption of parallel trends implies that the enrollment trends for any two groups will be indistinguishable if their menstruation status does not change across two consecutive survey rounds. While the nature of the Young Lives study does not allow a test of pre-treatment parallel trends, note that three groups of children are equally likely and universally to be in school at eight (Table 1). Further, the children across the three groups start school at the same age, at around five years (Table 1). The parallel trends assumption holds when I compare girls' enrollment rates in the early and late menarche groups across the third and fourth survey rounds (Table 4). Although boys are not a good counterfactual for girls in the early menarche group, trends in enrollment are also statistically indistinguishable for comparisons between girls in the early menarche group and boys across (1) the second and third survey rounds and (2) the third and fourth survey rounds.

Kahn-Lang and Lang (2019) discuss the need to carefully consider assumptions inherent in difference-in-differences analysis by (1) addressing the differences in the initial levels, (2) exploring if the parallel trends assumption requires a justification of the chosen functional form, and (3) acknowledging that comparing pre-treatment trends does not establish parallel trends during the treatment period.

I addressed concerns (1) and (3) in the discussion on parallel trends earlier in this subsection. Average enrollment rates for children across groups do not differ when the children are eight, that is, in the pre-treatment stage. The comparisons across the third and the fourth survey rounds demonstrate that post-treatment parallel trends hold. Thus, it can be claimed with some credibility that parallel trends would have been the counterfactual trends.

If the initial distributions of the outcome (enrollment rate in this case) are different across comparison groups, it is difficult to justify the parallel trends assumption using pretesting with linear trends. This is the motivation for the Kahn-Lang and Lang (2019)'s second concern. They agree that it is hard to make a case for one model over the other in the absence of a theory and recommend testing parallel trends for different functional form assumptions. In addition to the tests using a linear probability model described above, tests using logit functional form also indicate that the counterfactual trends in enrollment are parallel across comparison groups (Table B.1).[21]

6.4 Robustness to Alternative Hypotheses Spuriously Linking Menarche to School Enrollment

Association of early menarche with nutrition. Menarche is correlated with better nutritional status in South Asia. However, better nutritional status is associated with higher educational attainments in the same population (Nandi et al. 2015). Since the girls who reach menarche before twelve are healthier, it is unlikely that they dropped out of school on account of poor nutritional status. Not only do girls in the early menarche group have better nutritional status at age eight, their height and BMI also increase at a faster rate between ages eight and twelve (Table B.2). Of

21. Instead of individual fixed effects, which can be demanding for a conditional-logit specification, I use sampling-site fixed effects in this specification.

course, including individual fixed effects mitigates the concerns of bias induced by the correlation between better nutritional status and earlier menarche.

Reallocation of resources to siblings. The fact that girls in the early menarche group had better nutritional status at age eight, however, raises another potential concern. If equity considerations dictate parents' decision to allocate resources to children within the household, they may reduce complementary investments in their healthier daughter, affecting her schooling. While I do not observe child-specific investments between ages eight and twelve, non-food expenditure and dietary inputs do not differ across girls in the early and the late menarche groups at twelve (Table B.3).

Taller children undertaking wage labor. One may also be worried that parents pull their children out of school once they reach a certain height to have them undertake wage labor. Since girls in the early menarche group were taller than their counterparts at age eight, the decrease in enrollment at age twelve might be such a phenomenon. Three pieces of evidence point that this is unlikely to be the case. First, enrollment levels never converge across the two groups of girls (Figure 2). Second, while girls in the early menarche group spend more time on paid work, this effect is entirely driven by those who dropped out of school (Table 5). Finally, for both the younger siblings and the elder siblings, their time use patterns do not differ by the timing of their sisters' menarche (Table B.4; Table B.5, Appendix B), suggesting that it is unlikely that girls were pulled out of school to help the family with housework.[22]

The link between stress, early menarche, and schooling. Since high levels of cortisol due to stress leads to earlier menarche, unexpected negative economic shocks that induce mental stress can also lead to early menarche (Karapanou and Papadimitriou 2010). Any association between earlier menarche and school enrollment might be capturing the effect of external stressors on enrollment. To test for the robustness to this possible source of bias, I use health emergencies and deaths in the family as proxies for external stressors.[23] The incidence of health emergencies, deaths, and economic shock due to droughts are similar for girls in the early and late menarche groups (Table 1). Another test to confirm that this alternative hypothesis is not driving the key result of this paper is to check if co-resident siblings of girls in the early menarche group drop out of school at a higher rate than siblings of other surveyed children. Table B.6 in Appendix B presents these results and confirms that siblings of girls in the early menarche group are not more likely to drop out of school between the first and the second survey rounds.

6.5 The Impact of Attaining Menarche Between Twelve and Fifteen

The third survey round of the Young Lives study did not collect information about menarche status. To explore if reaching menarche after twelve affects school enrollment, I compare girls in the late menarche group with boys across the second and third survey rounds. The underlying

22. The Young Lives study collects information on the time use of all the children between the ages of five and eighteen in the household for a typical day in the last week. An important caveat is that the time use data for siblings is missing for 491 siblings out of the total 1,847 siblings of the Young Lives children.

23. Health emergencies and deaths in the family are the most significant drivers of a family's descent into poverty in developing countries (Krishna 2010).

assumption for this analysis is that most of the girls in the late menarche group start menstruating between the second and third survey rounds. In the second round of the India Human Development Survey (IHDS), which was conducted in 2011-12, 99% of the interviewed women in Andhra Pradesh had reached menarche before they turned fifteen, and 90% of the entire sample of interviewed women had reached menarche before age fifteen. Since age at menarche among Indian women is steadily declining (Pathak, Tripathi, and Subramanian 2014), patterns from the IHDS support the assumption that girls in the late menarche group reach menarche before fifteen.

When girls in the late menarche group are compared with boys, although statistically insignificant, the impact estimator is -5.1% point or 7.6% of the average enrollment rate for boys at fifteen (p-value: 0.17; Table 6). When I compare girls in the early and the late menarche groups, β_3 is not significantly different from zero.

In line with Prediction 2 in Section 3, the estimated effect of late menarche on school enrollment is smaller than the estimated effect of early menarche. Further, enrollment rates do not converge, and girls in the early menarche group have lower enrollment rates than other girls in their cohort persistently.

6.6 Heterogeneity in the Impact on Enrollment

Overall, the results show that the decrease in enrollment caused by menarche drives the gender gap in enrollment that emerges between the ages of eight and twelve. This subsection tests Predictions 3 and 4 and explores how the strength of the relationship between early menarche and enrollment varies with safety and reputation costs and with local labor market conditions. This section focuses on the comparison between girls in the early and the late menarche groups across the first and the second survey rounds.

Reputation and Safety Costs

In South Asia, the notion of "family honor" is inextricably linked to women's behavior in the family. As discussed in Section 2, the fear of being labeled as a girl of bad character who may bring dishonor upon the family intensifies as a girl transitions into womanhood after menarche. Further, these concerns are more pertinent among upper castes than lower castes (Eswaran, Ramaswami, and Wadhwa 2013). In Andhra Pradesh, the evidence from the second round of the IHDS shows that female mobility is least restricted among women from the scheduled castes.[24] Further, women from the scheduled castes are less like to practice *purdah*, a social practice wherein women are physically segregated from men or veil their faces (Table B.7). The decrease in school enrollment due to menarche is 3.5% point larger for girls from non-scheduled caste families, suggesting that the cultural norms that restrict female mobility could be critical drivers of a higher dropout rate after menarche (Table 7).

The discussion in Section 2 suggests that the onset of menstruation might alter a girl's perceived needs and experiences of safety in her community. Indeed, a comparison of girls' responses

24. Scheduled caste is the official designation for the historically disadvantaged caste group in India.

to "whether they feel safe in the community" at twelve shows that girls in the early menarche group were 6.8% point less likely to strongly agree that they feel safe in their community. The corresponding average for girls in the late menarche group was 81%. An immediate implication of this change is that the effect of early menarche on enrollment could be pronounced in relatively unsafe communities. I use the community-level average of boys' safety perceptions to encapsulate children's concerns and experiences of the safety of their neighborhood.[25] Indeed, the decrease in enrollment is larger in communities with a lower average perceived safety (Table 7). While the decline in enrollment for girls in unsafe communities is 17.7% points, it is only 4.2% points in safer communities.[26]

Local Labor Market Conditions

Substantial potential gains from education can mitigate the negative impact of menarche on enrollment. Since parents base their expectations of their daughter's earning potential on the female wage profile in their neighborhood (Chamarbagwala 2008; Jensen 2012), average salaries for female-dominated professions, teaching and nursing, are used as proxies for local wages for educated women. The Young Lives study collected this information through community-level focus-group discussions. I bifurcate the sample by the indicator that the average wages for these professions are less than the median for the whole sample. Results based on both these proxies suggest that the negative impact of menarche on enrollment is smaller in communities with higher wages for women (Table 8).

7 Other Related Outcomes

The evidence on parallel trends and a battery of robustness checks described in Section 6 confirm that early menarche reduces the school enrollment rate. In this section, I explore if outcomes other than school enrollment also vary by a girl's menarche status.

7.1 Test Scores

Girls in the early menarche group outperform girls in the late menarche group when children are twelve, although differences in PPVT test scores are not statistically significant (Table B.8). While boys outperform girls in the late menarche group, there is no difference in the average test scores for girls in the early menarche group and boys at age twelve. However, if I consider the trends in these test scores across the second and the third survey (when the children were twelve and fifteen, respectively), girls in the early menarche group lose the advantage they had at age twelve. The difference-in-differences estimator for a comparison between girls in the early menarche group and boys is negative. Moreover, the gender gaps in the test scores for both PPVT and Mathematics tests widen (Table 9). Taken together with the fact that girls in the early menarche group are healthier

25. I only use data for boys because girls' perceptions of safety are affected by menarche.

26. Since heterogeneous effects are explored in a specification with individual fixed effects, the effects on different sub-groups, although qualitatively large, are not statistically significant.

at age eight, these patterns suggest that specific social contexts can attenuate the gains from early nutrition and drive the gender gaps in not only school enrollment but also learning.

7.2 Time Use

When a girl drops out of school, the most immediate change is the reallocation of the time she spent on education-related activities. The Young Lives survey asked children about their time for a typical day in the last week. The decrease in time spent at school and on studying at home translates into an additional twenty-nine minutes on paid work and seventeen minutes on household chores. Changes in time use patterns of other activities are not statistically significant. In the sub-sample of those girls still enrolled in school, time spent at school and on studying does not vary with the girl's menarche status, suggesting that early menarche does not affect the intensive margin of schooling (Table 5). This is in line with Oster and Thornton (2011) that menstruation explains only 0.4 missed days in a 180 day school-year.

7.3 Aspirations

Besides changes in time use, by restricting mobility, menarche also redefines the scope of a girl's social interactions (Seymour, 1999). Do changes in parents' and society's interactions with girls in the early menarche group affect how they see themselves? Early menarche increases a girl's desire to become a full-time parent/housewife by 9.7% point. Notably, at twelve, only 10.7% of girls in the late menarche group report that they want to become full-time parents or housewives when they grow up (Table 10). However, if I restrict the analysis to the sample of girls still in school, the effect size is smaller (-1.3% point) and statistically insignificant. While girls in the early and the late menarche groups are equally likely to aspire to receive a college education, girls in the early menarche group are less likely to believe that they would reach their desired level of education.[27] It is an open question whether it is the change in aspirations after menarche that induces a girl to drop out or her withdrawal from formal education that alters how she sees herself.

7.4 Marital Outcomes

At nineteen, 45% of the girls in the early menarche group are married, while 31% of girls in the late menarche group are married. The marriage rate increases for both groups between nineteen and twenty-two. At twenty-two, 65% of the girls in the early menarche group are married, while 54% of girls in the late menarche group are married (Figure 4).[28] Once child-specific characteristics and sampling site fixed effects are included, girls in the early menarche group are 11% point (p-value: 0.12) more likely to be married at nineteen and also 11% point (p-value: 0.16) more likely to be

27. College refers to post-high school education that typically starts around the age of 18.

28. The Young Lives study started collecting data on marital outcomes only from the fourth survey round when the typical cohort age was nineteen.

married at twenty-two.[29] Rates at which girls in the early and the late menarche groups get married between nineteen and twenty-two are statistically indistinguishable (Table 11).

Figure 5 shows the distribution of age at marriage for girls in the early and late menarche groups. The distribution of age at marriage for girls in the late menarche group is shifted to the right of the distribution for girls in the early menarche group: girls in the late menarche group marry later, on average, and at all quantiles of age at marriage. For the restricted sample of ever-married girls, the average age at which girls in the early menarche group get married is 0.81 years lower than the average age at which girls in the late menarche group get married (Table 11).

8 Conclusion

In this paper, I examine the effect of menarche on schooling in India, where menarche is associated with substantial changes in a girl's life. Using a dataset that follows the same cohort of children during childhood, adolescence, and early adulthood, I show that reaching menarche before age twelve causes a 14.3% decrease in school enrollment. This result is robust to a variety of alternate explanations that could spuriously link early menarche with enrollment. The magnitude of the estimated effect of early menarche is within the range of effect sizes of many education-related policy reforms and schemes in India, including school sanitation (Adukia 2017), free school lunch to primary school students (Kaur 2017), a large-scale school construction program (Azam and Saing 2017), and providing bicycles to school-going girls (Muralidharan and Prakash 2017).

At age twelve, girls in the early menarche group perform as well as boys in tests designed to measure their mathematical ability and vocabulary. However, boys outperform girls in the early menarche group when the children are fifteen. Therefore, in settings where menstruation is associated with ritual impurity and job prospects for women are limited, the onset of menstruation can trigger transitions that can minimize the long-term gains through the link between better nutrition and schooling. This result also contributes to our understanding of the complex process of human capital formation by showing adolescence is another sensitive period for girls (Barham, Macours, and Maluccio 2018). Thus, if certain milieu foster gender-restrictive norms such that early menarche impedes schooling, gains from nutritional interventions may be undermined for girls. Early menarche becomes another channel that fosters and sustains gender gaps in human development.

Given the sizable impact of menarche on education, improving access to menses management technology may seem the most appropriate policy response. While menstrual hygiene management is a growing sector within international development, the existing evidence on the impact of these interventions on schooling is not very promising across contexts, including Nepal (Oster and Thornton 2011), Kenya (Stopford 2011; (Benshaul-Tolonen et al. 2019), and Ghana (Montgomery et al. 2012).

29. There are two key differences in the empirical strategy adopted in this subsection. First, the child-specific controls correspond to her natal household instead of the household she currently resides in because marital outcomes would have depended on her natal home's characteristics since, after marriage, she must live with her husband's family. Second, instead of an asset index, land-holding of the natal household when children were fifteen is used to capture their natal household's wealth. Since the turnover rates in land markets are low (Basu Roy and Ghosh Dastidar 2018), land-holdings capture the natal household's wealth at the time of the wedding.

These evaluations study the efficacy of menstrual hygiene management technology among the selected group of girls who are still in school after menarche. If menarche affects school enrollment, as I show in this paper for India, an absence of any impact of menstrual hygiene management on attendance is not a confirmation that menstruation does not affect schooling.

Early menarche has implications for a girl's life beyond those implied by the change in her physiology. A better understanding of these changes has to be at the center of any policy response. The first step in that direction is an analysis of the interaction between menarche and the decision around school enrollment. Results from exploring the heterogeneity in the impacts show that higher dropout rates due to early menarche are located in relatively unsafe communities. Indeed, interventions that address the safety concerns of young girls, such as providing bicycles (Muralidharan and Prakash 2017) and separate-sex toilets (Adukia 2017), have been successful in encouraging higher female enrollment. Furthermore, in urban India, the fear of street harassment affects women's college choices (Borker 2017).

The impact of menarche is more pronounced among girls from non-scheduled caste households, where restrictive gender norms are more pertinent. In such conservative communities, while families personally reject restrictive norms, they incorrectly estimate that other people accept them and continue to follow them (Bursztyn, González, and Yanagizawa-Drott 2020). These perceived cultural norms are not immutable. Bursztyn, González, and Yanagizawa-Drott (2020) show that correcting false beliefs does alter behavior. Persuasion and discussion can also change gender attitudes: a social campaign that engaged children in discussions about gender equality shifted gender attitudes to be more progressive in India (Dhar, Jain, and Jayachandran 2018), and an intervention that fosters grit closes the gender gap in competitiveness among school children in Turkey (Alan and Ertac 2019). The appropriate policy responses to higher dropout rates among adolescent girls should address girls' safety concerns in particular and gender-based cultural norms in general.

The factors that counter the effect of early menarche on enrollment are equally important. Girls are less likely to drop out of school after menarche if they live in neighborhoods with higher wages for female-dominated professions. Socio-cultural norms restrict girls' mobility and society's progress towards gender parity, but increased economic opportunities act as a vital counterweight to their impact on girls' education.

References

Adukia, Anjali. 2017. "Sanitation and education." *American Economic Journal: Applied Economics* 9 (2): 23–59.

Alan, Sule, and Seda Ertac. 2019. "Mitigating the gender gap in the willingness to compete: Evidence from a randomized field experiment." *Journal of the European Economic Association* 17 (4): 1147–1185.

ASER. 2018. *'Beyond Basics' 2017.* Technical report.

Athey, Susan, and Guido W Imbens. 2017. "The state of applied econometrics: Causality and policy evaluation." *Journal of Economic Perspectives* 31 (2): 3–32.

Attanasio, Orazio P. 2015. "The determinants of human capital formation during the early years of life: Theory, measurement, and policies." *Journal of the European Economic Association* 13 (6): 949–997.

Azam, Mehtabul, and Chan Hang Saing. 2017. "Assessing the impact of district primary education program in India." *Review of Development Economics* 21 (4): 1113–1131.

Bagga, Amrita, S Kulkarni, et al. 2000. "Age at menarche and secular trend in Maharashtrian (Indian) girls." *Acta Biologica Szegediensis* 44 (1-4): 53–57.

Barham, Tania, Karen Macours, and John A Maluccio. 2018. "Experimental evidence of exposure to a conditional cash transfer during early teenage years: Young women's fertility and labor market outcomes."

Basu Roy, Sharanya, and Sayantan Ghosh Dastidar. 2018. "Why do men rape? Understanding the determinants of rapes in India." *Third World Quarterly:* 1–23.

Benshaul-Tolonen, Anja, Garazi Zulaika, Elizabeth Nyothach, Clifford Oduor, Linda Mason, David Obor, Kelly T Alexander, Kayla F Laserson, and Penelope A Phillips-Howard. 2019. "Pupil absenteeism, measurement, and menstruation: Evidence from western Kenya." *CDEP-CGEG WP,* no. 74.

Bertrand, Marianne. 2011. "New perspectives on gender." In *Handbook of labor economics,* 4:1543–1590. Elsevier.

Borker, Girija. 2017. "Safety first: Perceived risk of street harassment and educational choices of women." *Job Market Paper, Department of Economics, Brown University.*

Bursztyn, Leonardo, Alessandra L González, and David Yanagizawa-Drott. 2020. "Misperceived social norms: Women working outside the home in saudi arabia." *American economic review* 110 (10): 2997–3029.

Cantet, Natalia. 2019. "The Effect of Teenage Pregnancy on Schooling and Labor Force Participation: Evidence From Urban South Africa."

Chamarbagwala, Rubiana. 2008. "Regional returns to education, child labour and schooling in India." *The Journal of Development Studies* 44 (2): 233–257.

Chari, AV, Rachel Heath, Annemie Maertens, and Freeha Fatima. 2017. "The causal effect of maternal age at marriage on child wellbeing: Evidence from India." *Journal of Development Economics* 127:42–55.

Chawla, Janet. 1994. "Mythic origins of menstrual taboo in Rig Veda." *Economic and Political Weekly:* 2817–2827.

Chowdhury, Sameena, AK Mohammed Shahabuddin, Anil J Seal, Kurshid K Talukder, Quamrul Hassan, Rowshan A Begum, Q Rahman, A Tomkins, A Costello, and MQK Talukder. 2000. "Nutritional status and age at menarche in a rural area of Bangladesh." *Annals of Human Biology* 27 (3): 249–256.

Cunha, Flavio, and James Heckman. 2007. "The technology of skill formation." *American Economic Review* 97 (2): 31–47.

Currie, Janet, and Douglas Almond. 2011. "Human capital development before age five." In *Handbook of labor economics,* 4:1315–1486. Elsevier.

Dambhare, Dharampal G, Sanjay V Wagh, and Jayesh Y Dudhe. 2012. "Age at menarche and menstrual cycle pattern among school adolescent girls in Central India." *Global Journal of Health Science* 4 (1): 105.

Desai, Sonalde, and Lester Andrist. 2010. "Gender scripts and age at marriage in India." *Demography* 47 (3): 667–687.

Dhamija, Gaurav, and Punarjit Roychowdhury. 2020. "Age at Marriage and Women's Labour Market Outcomes in India." *Journal of International Development* 32 (3): 342–374.

Dhar, Diva, Tarun Jain, and Seema Jayachandran. 2018. *Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India.* Technical report. National Bureau of Economic Research.

Dharmalingam, A. 1994. "The implications of menarche and wedding ceremonies for the status of women in a South Indian village." *Indian Anthropologist* 24 (1): 31–43.

Dube, Leela. 1988. "On the construction of gender: Hindu girls in patrilineal India." *Economic and Political Weekly:* WS11–WS19.

Eswaran, Mukesh, Bharat Ramaswami, and Wilima Wadhwa. 2013. “Status, caste, and the time allocation of women in rural India.” *Economic Development and Cultural Change* 61 (2): 311–333.

Field, Erica, and Attila Ambrus. 2008. “Early marriage, age of menarche, and female schooling attainment in Bangladesh.” *Journal of Political Economy* 116 (5): 881–930.

Garg, Suneela, and Tanu Anand. 2015. “Menstruation related myths in India: strategies for combating it.” *Journal of Family Medicine and Primary Care* 4 (2): 184.

Garg, Suneela, Nandini Sharma, and Ragini Sahay. 2001. “Socio-cultural aspects of menstruation in an urban slum in Delhi, India.” *Reproductive Health Matters* 9 (17): 16–25.

Hicks, Joan, and Daniel L Hicks. 2019. “Lucky Late Bloomers? Consequences of Delayed Marriage for Women in Rural Western Kenya.” *Consequences of Delayed Marriage for Women in Rural Western Kenya (June 2019).*

ICRW and Plan Asia. 2013. *Plan Asia Child Marriage Initiative: Summary of Research in Bangladesh, India and Nepal.* Technical report. Plan Asia Regional Office.

Jacoby, Hanan G, and Ghazala Mansuri. 2011. *Crossing boundaries: gender, caste and schooling in rural Pakistan.* The World Bank.

Jayachandran, Seema. 2015. “The roots of gender inequality in developing countries.” *Annual Review of Economics* 7 (1): 63–88.

Jayachandran, Seema, and Rohini Pande. 2017. “Why are Indian children so short? The role of birth order and son preference.” *American Economic Review* 107 (9): 2600–2629.

Jensen, Robert. 2012. “Do labor market opportunities affect young women’s work and family decisions? Experimental evidence from India.” *The Quarterly Journal of Economics* 127 (2): 753–792.

Kahn-Lang, Ariella, and Kevin Lang. 2019. “The promise and pitfalls of differences-in-differences: Reflections on 16 and pregnant and other applications.” *Journal of Business & Economic Statistics:* 1–14.

Kaprio, Jaakko, Arja Rimpelä, Torsten Winter, Richard J Viken, Matti Rimpelä, and Richard J Rose. 1995. “Common genetic influences on BMI and age at menarche.” *Human Biology:* 739–753.

Karapanou, Olga, and Anastasios Papadimitriou. 2010. “Determinants of menarche.” *Reproductive Biology and Endocrinology* 8 (1): 115.

Kaur, Randeep. 2017. “Essays in human capital development.” PhD diss.

Krishna, Anirudh. 2010. “Who became poor, who escaped poverty, and why? Developing and using a retrospective methodology in five countries.” *Journal of Policy Analysis and Management* 29 (2): 351–372.

Mahon, Thérèse, and Maria Fernandes. 2010. “Menstrual hygiene in South Asia: a neglected issue for WASH (water, sanitation and hygiene) programmes.” *Gender & Development* 18 (1): 99–113.

Mason, Linda, Elizabeth Nyothach, Kelly Alexander, Frank O Odhiambo, Alie Eleveld, John Vulule, Richard Rheingans, Kayla F Laserson, Aisha Mohammed, and Penelope A Phillips-Howard. 2013. “‘We keep it secret so no one should know’–A qualitative study to explore young schoolgirls attitudes and experiences with menstruation in rural Western Kenya.” *PloS one* 8 (11).

Montgomery, Paul, Caitlin R Ryus, Catherine S Dolan, Sue Dopson, and Linda M Scott. 2012. “Sanitary pad interventions for girls’ education in Ghana: a pilot study.” *PloS one* 7 (10): e48274.

Muralidharan, Karthik, and Nishith Prakash. 2017. “Cycling to school: increasing secondary school enrollment for girls in India.” *American Economic Journal: Applied Economics* 9 (3): 321–50.

Nandi, Arindam, Ashvin Ashok, Sanjay Kinra, Jere R Behrman, and Ramanan Laxminarayan. 2015. “Early Childhood Nutrition Is Positively Associated with Adolescent Educational Outcomes: Evidence from the Andhra Pradesh Child and Parents Study (APCAPS)–3.” *The Journal of Nutrition* 146 (4): 806–813.

Olivetti, Claudia, Eleonora Patacchini, and Yves Zenou. 2020. “Mothers, peers, and gender-role identity.” *Journal of the European Economic Association* 18 (1): 266–301.

Oster, Emily, and Rebecca Thornton. 2011. “Menstruation, sanitary products, and school attendance: Evidence from a randomized evaluation.” *American Economic Journal: Applied Economics* 3 (1): 91–100.

Pathak, Praveen Kumar, Niharika Tripathi, and SV Subramanian. 2014. “Secular trends in menarcheal age in India-evidence from the Indian human development survey.” *PLoS One* 9 (11): e111027.

Rah, Jee H, Abu Ahmed Shamim, Ummeh T Arju, Alain B Labrique, Mahbubur Rashid, and Parul Christian. 2009. “Age of onset, nutritional determinants, and seasonal variations in menarche in rural Bangladesh.” *Journal of Health, Population, and Nutrition* 27 (6): 802.

Rao, Vijayendra. 1993. “Dowry ‘inflation’in rural India: A statistical investigation.” *Population Studies* 47 (2): 283–293.

Roychowdhury, Punarjit, and Gaurav Dhamija. 2018. “The causal impact of women’s age at marriage on domestic violence in India.” *Feminist Economics.*

Sekhri, Sheetal, and Sisir Debnath. 2014. "Intergenerational consequences of early age marriages of girls: Effect on children's human capital." *The Journal of Development Studies* 50 (12): 1670–1686.

Seymour, Susan C. 1999. *Women, family, and child care in India: A world in transition.* Cambridge University Press.

Singh, Renu, and Uma Vennam. 2016. "Factors shaping trajectories to child and early marriage: Evidence from Young Lives in India."

Sivakami, Muthusamy, Anna Maria van Eijk, Harshad Thakur, Narendra Kakade, Chetan Patil, Sharayu Shinde, Nikita Surani, Ashley Bauman, Garazi Zulaika, Yusuf Kabir, et al. 2019. "Effect of menstruation on girls and their schooling, and facilitators of menstrual hygiene management in schools: surveys in government schools in three states in India, 2015." *Journal of Global Health* 9 (1).

Sohn, Kitae. 2016. "The influence of climate on age at menarche: Augmented with the influence of ancestry." *HOMO-Journal of Comparative Human Biology* 67 (4): 328–336.

Sommer, Marni. 2010. "Where the education system and women's bodies collide: The social and health impact of girls' experiences of menstruation and schooling in Tanzania." *Journal of Adolescence* 33 (4): 521–529.

Stopford, Amy Louise. 2011. "The Effect of Commercial Sanitary Pad Use on School Attendance and Health of Adolescents in Western Kenya." PhD diss.

Sundaram, Aparna, and Reeve Vanneman. 2008. "Gender differentials in literacy in India: The intriguing relationship with women's labor force participation." *World Development* 36 (1): 128–143.

Sunder, Naveen. 2017. "Marriage Age, Social Status and Intergenerational Effects in Uganda."

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

Vennam, Uma, and Anuradha Komanduri. 2009. "Young Lives qualitative research: round 1–India."

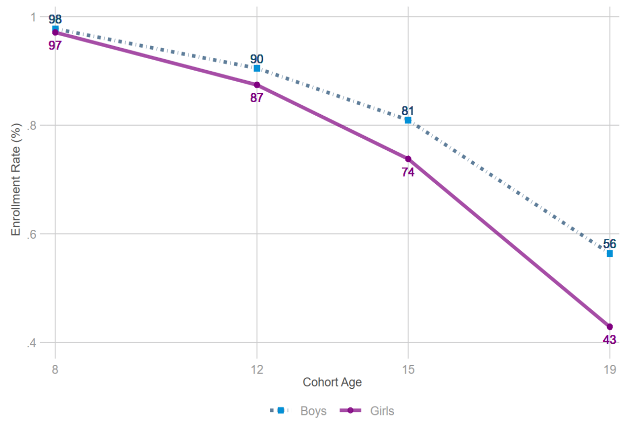
Villamor, Eduardo, and Erica C Jansen. 2016. "Nutritional determinants of the timing of puberty." *Annual Review of Public Health* 37:33–46.

Young, Alwyn. 2019. "Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results." *The Quarterly Journal of Economics* 134 (2): 557–598.

Young Lives. 2017. *A Guide to Young Lives Research.* Technical report. Young Lives.

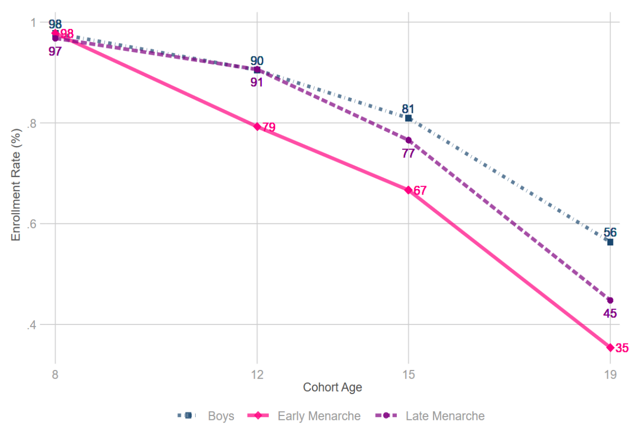
Figures

Figure 1. Enrollment Rate by Gender



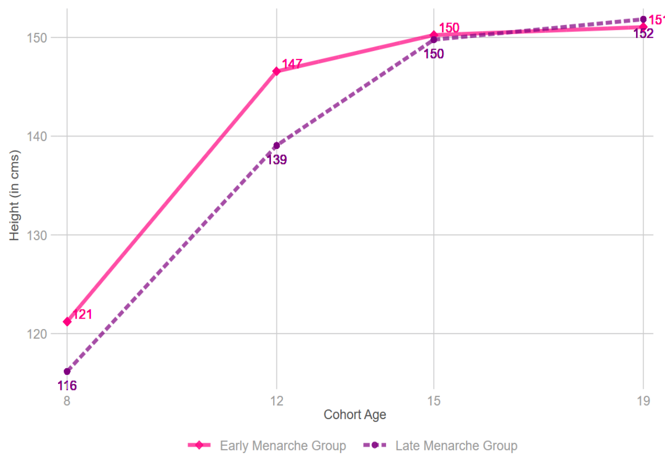
Note: Data are taken from the Young Lives study. The solid (blue) line represents the trends in enrollment rate for boys and the dashed (purple) line represents the trends in enrollment rate for girls.

Figure 2. Enrollment Rate by Gender and Menarche Status



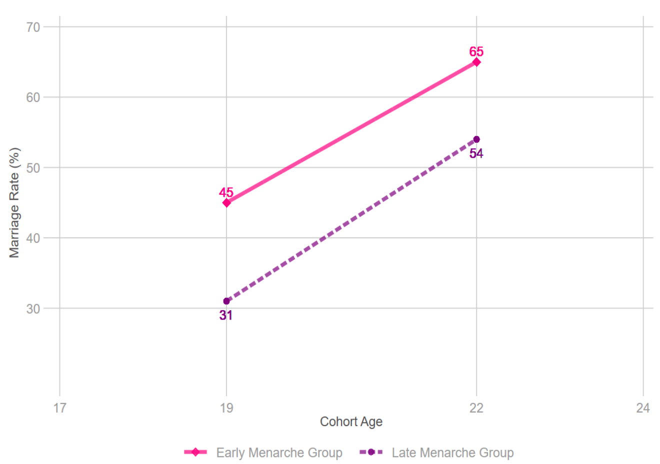
Note: Data are taken from the Young Lives study. The dash-dotted (blue) line represents the trends in enrollment rate for boys, the dashed (purple) line represents the trends in enrollment rate for girls in the late menarche group and the solid (pink) line represents the trends in enrollment rate for girls in early menarche group.

Figure 3. Height by Menarche Status



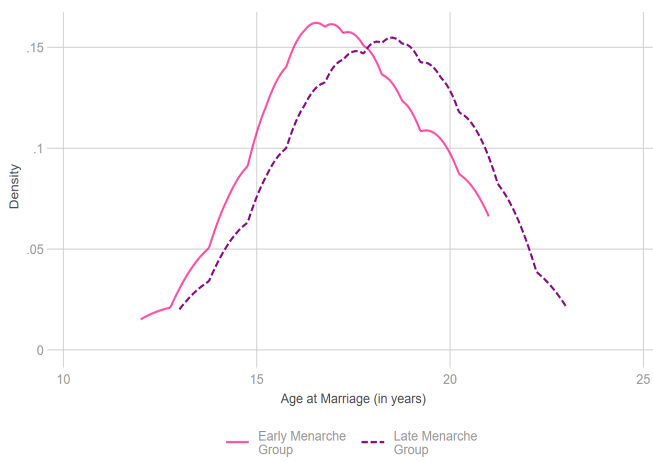
Note: Data are taken from the Young Lives study. The solid (pink) line represents the trends in height (in centimeters) for girls in the early menarche group and the dashed (purple) line represents the trends in height (in centimeters) for girls in the late menarche group.

Figure 4. Marriage Rate by Menarche Status



Note: Data are taken from the Young Lives study. The solid (pink) line represents the trends in marriage rate for girls in the early menarche group and the dashed (purple) line represents the trends in marriage rate for girls in the late menarche group. Both the lines describe trends the ages of nineteen and twenty-two.

Figure 5. Age at Marriage by Menarche Status



Note: Data are taken from the Young Lives study. The solid (pink) line represents the distribution of age at marriage for girls in the early menarche group and the dashed (purple) line represents the trends in age at marriage for girls in the late menarche group.

Tables

Table 1. Sample Characteristics in Round 1

	Early Menarche	Late Menarche	Boys	Difference	
	(1)	(2)	(3)	(1) - (2)	(1) - (3)
Household size	5.32	5.62	5.56	0.30	-0.24
Wealth index	0.43	0.40	0.41	0.03	0.02
Access to drinking water	0.80	0.82	0.84	-0.02	-0.04
Electricity connection	0.84	0.82	0.82	0.02	0.03
Area of land owned (in ha.)	1.24	0.68	0.89	0.56	0.35
Literate mother	0.27	0.21	0.21	0.06	0.06*
Age of the household head	39.84	40.51	40.00	-0.67	-0.15
Schedule Caste	0.22	0.195	0.22	-0.67	0.00
Experience of health shocks	0.27	0.30	0.34	-0.03	-0.07
Experience of drought	0.23	0.28	0.30	-0.05	-0.07*
Speaks Telugu	0.86	0.85	0.83	0.01	0.03
Number of older siblings	0.91	1.07	1.16	-0.16	-0.25**
Number of younger siblings	0.69	0.80	0.67	-0.11	0.02
Number of older brothers	0.66	0.61	0.72	0.05	-0.06
Number of older sisters	0.42	0.67	0.72	-0.24**	-0.30**
Raven's test score at age 8	23.37	22.83	22.97	0.54	0.40
BMI at age 8 (kg/m^2)	14.21	13.74	13.99	0.48***	0.22
Height at age 8 (cm)	121.21	116.16	118.56	5.05***	2.65***
Weight at age 8 (kg)	20.96	18.57	19.70	2.39***	1.26***
Age in first year at school (in years)	4.98	5.05	5.04	-0.07	-0.06
Enrollment rate at age eight	97.78	96.79	98.14	0.99	-0.36

Note: This table reports simple differences in means across the three groups of children.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

Table 2. Enrolment Across Rounds 1 and 2 for Girls in the Early Menarche Group

	vs Late Menrache Girls				vs Boys			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Early menarche	0.010	0.018	0.026		-0.004	-0.003	-0.005	
	(0.011)	(0.019)	(0.022)		(0.014)	(0.014)	(0.014)	
Round 2	-0.062**	-0.035	-0.034	-0.009	-0.078***	-0.043**	-0.048**	-0.030
	(0.022)	(0.031)	(0.030)	(0.037)	(0.014)	(0.017)	(0.018)	(0.036)
Early menarche X Round 2	-0.123**	-0.121***	-0.122***	-0.130**	-0.107**	-0.111**	-0.111**	-0.109**
	(0.032)	(0.031)	(0.031)	(0.045)	(0.038)	(0.037)	(0.037)	(0.052)
Observations	1017	1013	1013	1013	1240	1234	1234	1234
R-squared	0.0494	0.106	0.162	0.604	0.0555	0.100	0.128	0.599
Control average	0.906	0.906	0.906	0.906	0.903	0.903	0.903	0.903
Cluster Fixed effects	No	No	Yes		No	No	Yes	
Individual Fixed effects	No	No	No	Yes	No	No	No	Yes
Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes

Note: This table reports the results from a difference-in-differences specification (see Equation 5.1) for comparisons that span the first and the second survey rounds. Controls in columns (1)-(3) and (5)-(7) include household size, caste, sex of the household head, wealth index, mom's literacy status, age of the head of the household, indicator of whether the child speaks the majority language, number of younger siblings, number of older siblings, number of older sisters, number of older brothers, test scores at age eight, weight at age eight, and BMI at age eight. Columns (4) and (8) only control for time-varying variables (household size, sex of the household head, wealth index, number of younger siblings, number of older siblings, number of older sisters, and number of older brothers). Sampling-site fixed effects are included in columns (1)-(3) and (5)-(7), and individual fixed effects are included in columns (4) and (8). Robust standard errors clustered at the level of the sampling site are reported.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

Table 3. Enrolment Across Rounds 1 and 2 for Girls in the Late Menarche Group

	vs Boys			
	(1)	(2)	(3)	(4)
Late menarche	-0.014	-0.014	-0.020	0.000
	(0.009)	(0.010)	(0.012)	(.)
Round 2	-0.078***	-0.061**	-0.063**	-0.056**
	(0.014)	(0.017)	(0.018)	(0.023)
Late menarche X Round 2	0.017	0.012	0.012	0.014
	(0.023)	(0.024)	(0.024)	(0.033)
Observations	1717	1707	1707	1707
R-squared	0.0228	0.0611	0.0857	0.587
Control average	0.903	0.903	0.903	0.903
Cluster Fixed effects	No	No	Yes	
Individual Fixed effects	No	No	No	Yes
Controls	No	Yes	Yes	Yes

Note: This table reports the results from a difference-in-differences specification (see Equation 5.1) for comparisons that span the first and the second survey rounds. Controls in columns (1)-(3) include household size, caste, sex of the household head, wealth index, mom's literacy status, age of the head of the household, indicator of whether the child speaks the majority language, number of younger siblings, number of older siblings, number of older sisters, number of older brothers, test scores at age eight, weight at age eight, and BMI at age eight. Column (4) only controls for time-varying variables (household size, sex of the household head, wealth index, number of younger siblings, number of older siblings, number of older sisters, and number of older brothers). Sampling-site fixed effects are included in columns (1)-(3) and individual fixed effects are included in column (4). Robust standard errors clustered at the level of the sampling site are reported.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

Table 4. Tests for Parallel Trends: OLS

	Girls in the Early Menarche Group vs Boys		Girls in the Early vs Late Menarche Groups
	(1)	(2)	(3)
Early menarche X Round 3	-0.035		
	(0.057)		
Early menarche X Round 4		-0.062	-0.006
		(0.081)	(0.075)
Observations	1216	1186	962
R-squared	0.819	0.780	0.774
Control average	0.808	0.561	0.448
Individual fixed effect	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Note: This table reports the results from a difference-in-differences specification (see Equation 5.1) for comparisons that span the second and the third survey rounds (column 2), and the third and the fourth survey rounds (columns 3 and 4). Controls include household size, sex of the household head, wealth index, number of younger siblings, number of older siblings, number of older sisters, and number of older brothers. Individual fixed effects are included. Robust standard errors clustered at the level of the sampling site are reported.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

Table 5. Time Use Patterns, Proportion of Time in the Day

	(1) Work	(2) Work (if enrolled)	(3) Household Chores	(4) Care	(5) Sleep	(6) Play	(7) School	(8) Study	(9) School (if enrolled)
Early menarche	0.019**	0.002	0.012	0.001	0.006	0.000	-0.028*	-0.014**	-0.002
	(0.009)	(0.002)	(0.008)	(0.003)	(0.006)	(0.011)	(0.015)	(0.005)	(0.009)
Observations	483	432	483	483	483	483	496	483	432
R^2	0.125	0.110	0.202	0.159	0.342	0.192	0.257	0.298	0.440
Control average	0.0111	0	0.0488	0.0114	0.396	0.166	0.269	0.0844	0.0844
Cluster FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports the results from a single-difference specification (see Equation 5.3) for data from the second survey round when children were twelve. Controls include household size, caste, sex of the household head, wealth index,mom's literacy status, age of the head of the household, indicator of whether the child speaks the majority language, number of younger siblings, number of older siblings, number of older brothers, number of older sisters, test scores at age eight, weight at age eight, and BMI at age eight. Sampling-site fixed effects are included. Robust standard errors clustered at the level of the sampling site are reported.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

Table 6. Enrolment Across Rounds 2 and 3 for Girls in the Late Menarche Group

	vs Early Menrache Girls				vs Boys			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Late menarche	0.114**	0.128***	0.119**		0.003	0.004	-0.008	
	(0.032)	(0.027)	(0.033)		(0.027)	(0.027)	(0.030)	
Round 3	-0.126**	-0.135**	-0.137**	-0.122**	-0.095***	-0.106***	-0.108***	-0.082**
	(0.039)	(0.042)	(0.043)	(0.054)	(0.016)	(0.018)	(0.019)	(0.026)
Late menarche X Round 3	-0.014	-0.017	-0.015	-0.018	-0.045*	-0.048**	-0.048**	-0.051
	(0.043)	(0.046)	(0.046)	(0.065)	(0.022)	(0.022)	(0.022)	(0.036)
Observations	1003	995	995	995	1700	1681	1681	1681
R-squared	0.0445	0.142	0.236	0.804	0.0271	0.0936	0.141	0.787
Control average	0.667	0.667	0.667	0.667	0.808	0.667	0.667	0.667
Cluster Fixed effects	No	No	Yes		No	No	Yes	
Individual Fixed effects	No	No	No	Yes	No	No	No	Yes
Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes

Note: This table reports the results from a difference-in-differences specification (see Equation 5.1) for comparisons that span the first and the second survey rounds. Controls in columns (1)-(3) and (5)-(7) include household size, caste, sex of the household head, wealth index, mom's literacy status, age of the head of the household, indicator of whether the child speaks the majority language, number of younger siblings, number of older siblings, number of older sisters, number of older brothers, test scores at age eight, weight at age eight, and BMI at age eight. Columns (4) and (8) only control for time-varying variables (household size, sex of the household head, wealth index, number of younger siblings, number of older siblings, number of older sisters, and number of older brothers). Sampling-site fixed effects are included in columns (1)-(3) and (5)-(7), and individual fixed effects are included in columns (4) and (8). Robust standard errors clustered at the level of the sampling site are reported.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

Table 7. Impact of Early Menarche of Enrollment, by Caste and Local Safety Environment

	(1) Scheduled Castes	(2) Average Perception of Safety
Round 2	-0.009	-0.035
	(0.037)	(0.026)
Early Menarche X Round 2 X SC	-0.105	
	(0.098)	
Early Menarche X Round 2 X Non-SC	-0.137**	
	(0.061)	
Early Menarche X Round 2 X Lower Perceived Safety		-0.177*
		(0.087)
Early Menarche X Round 2 X Higher Perceived Safety		-0.042
		(0.074)
Observations	1013	1977
R-squared	0.604	0.595
Control average	0.906	0.906
Individual fixed effect	Yes	Yes
Controls	Yes	Yes
P-val (coefficient diff)	0.80	0.27

Note: This table reports the results from specification that estimates the effect of early menarche across sub-groups (see Equation 5.2) for comparisons that span the first and the second survey rounds. Column (1) reports results by the caste of the family and Column (2) reports results by the indicator that the average community level perceived safety among boys is less than the median. Controls include household size, sex of the household head, wealth index, number of younger siblings, number of older siblings, number of older sisters, and number of older brothers. Individual fixed effects are included. Robust standard errors clustered at the level of the sampling site are reported.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

Table 8. Impact of Early Menarche of Enrollment, by Average Local Wages for Female Dominated Professions

	(1) Teachers' Wages	(2) Nurses' Wages
Round 2	-0.012	-0.013
	(0.039)	(0.039)
Early Menarche X Round 2 X Lower Teacher Wages	-0.191**	
	(0.081)	
Early Menarche X Round 2 X Higher Teacher Wages	-0.067	
	(0.042)	
Early Menarche X Round 2 X Lower Nurse Wages		-0.185**
		(0.063)
Early Menarche X Round 2 X Higher Nurse Wages		-0.031
		(0.056)
Observations	1013	1013
R-squared	0.608	0.609
Control average	0.906	0.906
Individual fixed effect	Yes	Yes
Controls	Yes	Yes
P-val (coefficient diff)	0.11	0.22

Note: This table reports the results from specification that estimates the effect of early menarche across sub-groups (see Equation 5.2) for comparisons that span the first and the second survey rounds. Column (1) reports results by the indicator that the typical wages for teachers in the community is less than that for the median community and Column (2) reports results by the indicator that the typical wages for nurses in the community is less than that for the median community. Controls include household size, sex of the household head, wealth index, number of younger siblings, number of older siblings, number of older sisters, and number of older brothers. Individual fixed effects are included. Robust standard errors clustered at the level of the sampling site are reported.

* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

Table 9. Normalized Test Scores of the Girls in the Early Menarche Group versus Other Children Across Rounds 2 and 3

	PPVT			Mathematics		
	(1) vs Late Menarche Group	(2) vs Boys	(3) vs Boys	(4) vs Late Menarche Group	(5) vs Boys	(6) vs Boys
Round 3	-0.106	0.174	0.185	-0.110	0.198*	0.190
	(0.197)	(0.110)	(0.131)	(0.132)	(0.104)	(0.111)
Early menarche X Round 3	-0.038	-0.325**		0.040	-0.246	
	(0.116)	(0.103)		(0.125)	(0.103)	
Girls X Round 3			-0.307**			-0.278**
			(0.114)			(0.130)
Observations	934	1133	1821	987	1209	1933
R-squared	0.796	0.825	0.814	0.796	0.815	0.811
Control average	-0.253	0.230	0.230	-0.220	0.208	0.208
Individual fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports the results from a difference-in-differences specification (see Equation 5.1). Controls include household size, sex of the household head, wealth index, number of younger siblings, number of older siblings, number of older sisters, and number of older brothers. Individual fixed effects are included. Robust standard errors clustered at the level of the sampling site are reported.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

Table 10. Aspirations and Beliefs

	(1)	(2)	(3)	(4)
	Aspiration: Fulltime parent or housewife	Aspiration: Fulltime parent or housewife (if enrolled)	Aspiration: Attend college	Aspirtation likely to be fulfilled
Early menarche			-0.043	-0.064*
			(0.045)	(0.038)
Round 2	0.020	-0.006		
	(0.038)	(0.033)		
Early menarche X Round 2	0.097*	-0.013		
	(0.055)	(0.049)		
Observations	1014	935	507	507
R-squared	0.548	0.544	0.115	0.230
Control average	0.107	0.107	0.738	0.861
Cluster Fixed effects			Yes	Yes
Individual Fixed effects	Yes	Yes	No	No
Controls	Yes	Yes	Yes	Yes

Note: Columns (1) and (2) reports the results from a difference-in-differences specification (see Equation 5.1) for comparisons that span the first and the second survey rounds. Controls include household size, sex of the household head, wealth index, number of younger siblings, number of older siblings, number of older sisters, and number of older brothers. Individual fixed effects are included. Individual fixed effects are included. Columns (3) and (4) reports the results from a single-difference specification (see Equation 5.3). Additional controls include caste, mom's literacy status, indicator of whether the child speaks the majority language, test scores at age eight, weight at age eight, and BMI at age eight. Sampling-site fixed effects are included. Robust standard errors clustered at the level of the sampling site are reported in all specifications.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

Table 11. Marriage Timing

	(1) Ever Married at 19	(2) Ever Married at 22	(3) Age	(4) Gap Between Ages at Dropout and Marriage	(5) Ever Married
Early menarche	0.114	0.112	-0.813**	-0.795*	0.128*
	(0.072)	(0.076)	(0.347)	(0.383)	(0.065)
Round 5					0.201***
					(0.020)
Early menarche X Round 5					-0.029
					(0.042)
Observations	495	495	272	264	953
R-squared	0.203	0.218	0.246	0.272	0.251
Control average	0.325	0.539	18.18	2.461	0.461
Cluster Fixed effects	Yes	Yes	Yes	Yes	
Individual Fixed effects	No	No	No	No	Yes
Controls	Yes	Yes	Yes	Yes	Yes

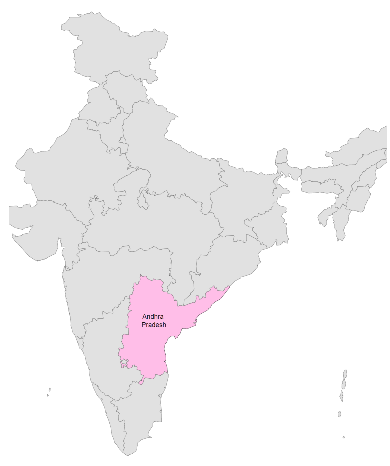
Note: Columns (1)-(4) of this table reports the results from a single-difference specification (see Equation 5.3) and column (5) reports results from a difference-in-differences specification with individual fixed effects (see Equation 5.2). All controls include girl's natal home's household size, caste, sex of the household head, landholding in round 3, mom's literacy status, age of the head of the household, indicator of whether the girl speaks the majority language, number of younger siblings, number of older siblings, number of older brothers, number of older sisters, test scores at age eight, weight at age eight, and BMI at age eight. Robust standard errors clustered at the level of the sampling site are reported.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

Appendices (not for publication)

A Context: State of Andhra Pradesh and the Sampling Strategy

Andhra Pradesh (including the state of Telangana) is the fifth-largest state in India. In 2002, when the Young Lives study started in Andhra Pradesh, only 27% of the population was living in urban areas. While agriculture was important and contributed to 30% of the state domestic product, economic activity was increasingly shifting away from agriculture. In 2004-05, the rural poverty rate was 11%, while the urban poverty rate was 28%. Andhra Pradesh was close to all India averages on various measures of human development. In 2004-05, the Andhra Pradesh's infant mortality rate was 53 as opposed to 57 for the whole country. The average life expectancy in Andhra Pradesh was 63.9, while it was 65.4 in India. Adult literacy and enrollment rates were 50.9 and 87.6 in Andhra Pradesh, respectively, and 61.8 and 82.1 in India (Centre for Economic and Social Studies, 2007).

Figure A.1. Andhra Pradesh



Andhra Pradesh has three historically distinct socio-cultural regions, Coastal Andhra, Rayalseema, and Telangana (Figure A.2). These divisions are divided into twenty-three districts (Figure A.3). Each district composed of three to five divisions, and each division has ten-fifteen *mandals*. A *mandal* is the lowest administrative tier of the government of Andhra Pradesh.

The sample for the Young Lives study was drawn such that the three regions the region were well represented. In addition to the metropolitan city of Hyderabad, two districts from each region, one poor and one non-poor were drawn. Next, from each district *mandals* were chosen to be representative of key development indicators. These chosen mandals, along with three urban slums chosen from Hyderabad, were the primary sampling units (sampling sites). From a list of villages that would have had at least fifty eight-year-olds and at least a hundred one-year-olds, sample villages were randomly chosen from the selected sampling sites.

Households covered by the Young Lives study are wealthier than the average household in Andhra Pradesh. A comparison with the households surveyed by the Demographic and Health

Survey, 1998-99 reveals that the surveyed households also had better access to services. Despite these biases, the sample covered by the Young Lives study captures the diversity of children in poor households in Andhra Pradesh.

Figure A.2. Socio-cultural Regions in Andhra Pradesh

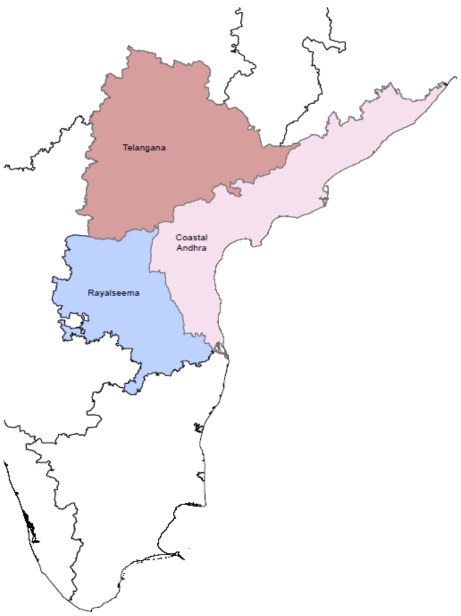
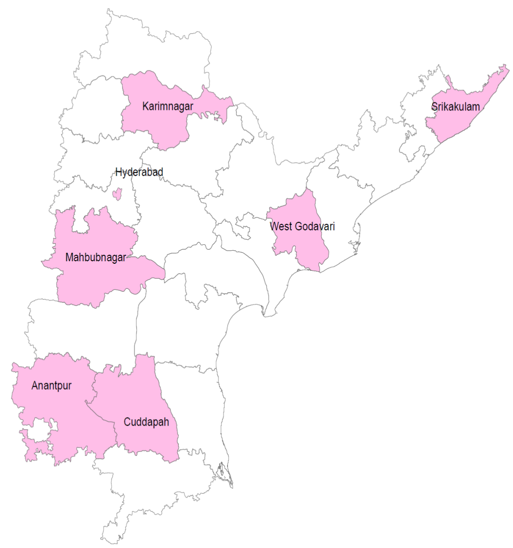


Figure A.3. Selected Districts in Andhra Pradesh



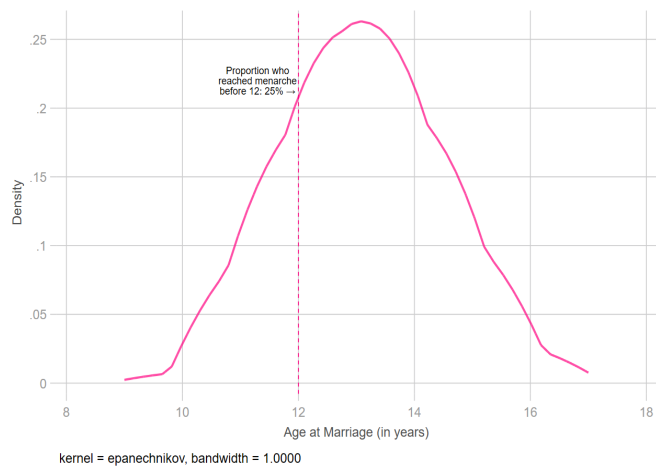
References

Centre for Economic and Social Studies. (2007). *Human Development Report 2007.* Retrieved from https://www.undp.org/content/dam/india/docs/human_revelop_report_andhra_pradesh_2007_full_report.pdf

B Additional Figures and Tables

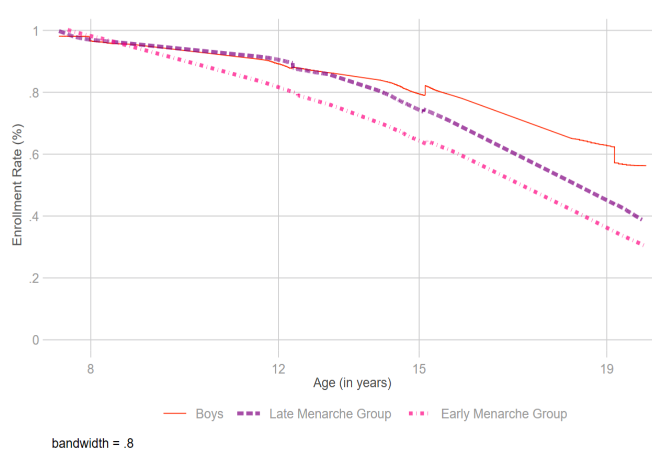
Figures

Figure B.1. Age at Menarche in Andhra Pradesh



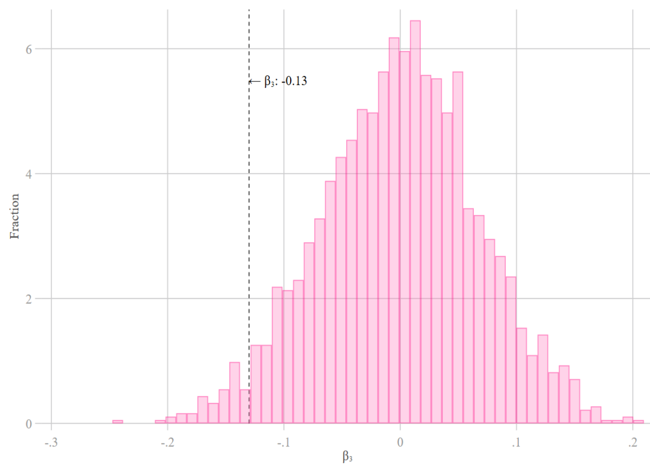
Note: Data are taken from the IHDS. The solid (pink) line represents the distribution of the age at menarche for the state of Andhra Pradesh.

Figure B.2. Enrollment Rate by Gender and Menarche Status - locally weighted regression



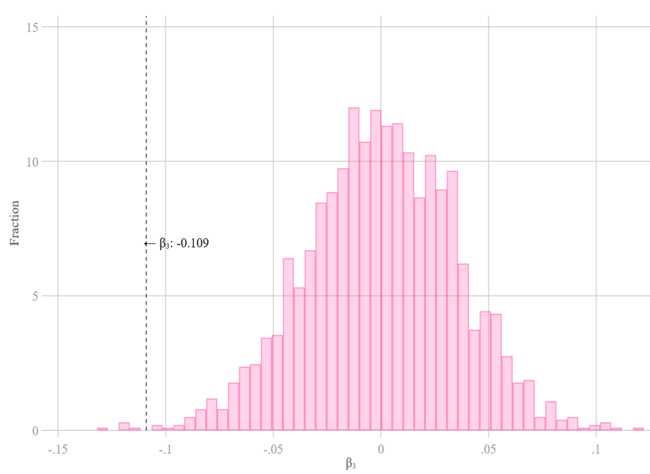
Note: Data are taken from the Young Lives study. The solid (red) line represents the trends in enrollment rate for boys, the dashed (purple) line represents the trends in enrollment rate for girls in the late menarche group and the dash-dotted (pink) line represents the trends in enrollment rate for girls in early menarche group.

Figure B.3. Permutation Tests for Estimates of β_3



Note: This figure shows the distribution of estimates of β_3 from a permutation test of for equation 5.1 when girls in the early menarche group are compared to girls in the late menarche group. For each child in the comparison sample, menarche status was randomly generated and these assignments were used to compute β_3 2000 times. This figure shows the distribution of the estimates of β_3

Figure B.4. Permutation Tests for Estimates of β_3



Note: This figure shows the distribution of estimates of β_3 from a permutation test of for equation 5.1 when girls in the early menarche group are compared to boys. For each child in the comparison sample, menarche status was randomly generated and these assignments were used to compute β_3 2000 times. This figure shows the distribution of the estimates of β_3

Tables

Table B.1. Tests for Parallel Trends: Logit

	Girls in the Early Menarche Group vs Boys		Girls in the Early vs Late Menarche Groups
	(1)	(2)	(3)
Early menarche X Round 3	0.141		
	(0.219)		
Early menarche X Round 4		-0.111	0.040
		(0.276)	(0.236)
Control average	0.808	0.561	0.448
Sampling site FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Note: This table reports the results from a difference-in-differences specification (see Equation 5.1) for comparisons that span the second and the third survey rounds (column 2), and the third and the fourth survey rounds (columns 3 and 4). Controls include household size, caste, sex of the household head, wealth index, mom's literacy status, age of the head of the household, indicator of whether the child speaks the majority language, number of younger siblings, number of older siblings, number of older sisters, number of older brothers, test scores at age eight, weight at age eight, and BMI at age eight. Sampling-site fixed effects are included. Robust standard errors clustered at the level of the sampling site are reported.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

Table B.2. Girls in the Early and the Late Menarche Group- BMI, Height and Caretaker's Perception of Child's Health

	(1) BMI	(2) Height (in cms.)	(3) Caretaker's Perception of Child's Health Relative to Peers
Round 2	1.657***	22.515***	-0.056
	(0.244)	(2.350)	(0.114)
Early menarche X Round 2	1.641***	2.123*	0.160
	(0.339)	(1.165)	(0.096)
Observations	1003	1007	1014
R-squared	0.826	0.876	0.535
Control average	15.56	139.1	0.425
Individual fixed effect	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Note: This table reports the results from a difference-in-differences specification (see Equation 5.1). Controls include household size, caste, sex of the household head, wealth index,mom's literacy status, age of the head of the household, indicator of whether the child speaks the majority language, number of younger siblings, number of older siblings, number of older sisters, number of older brothers, test scores at age eight, weight at age eight, and BMI at age eight. Sampling-site fixed effects are included. Robust standard errors clustered at the level of the sampling site are reported.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

Table B.3. Other Investments in Girls in the Early Menarche Group in Round 2

	Food Intake (last seven days)						
	Exp. (in rupees)	Cereal	Roots	Legumes	Milk	Eggs	Meat
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Early menarche	969.268	0.029	-0.037	-0.037	0.068*	0.007	0.037
	(1785.812)	(0.035)	(0.046)	(0.054)	(0.038)	(0.033)	(0.031)
Observations	501	507	507	507	507	486	478
R^2	0.128	0.255	0.208	0.224	0.0845	0.117	0.136
Control average	7559.8	0.251	0.380	0.647	0.171	0.0922	0.0627
Sampling site FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports the results from a single-difference specification (see Equation 5.3) using data from the second survey round. Controls include household size, caste, sex of the household head, wealth index,mom's literacy status, age of the head of the household, indicator of whether the child speaks the majority language, number of younger siblings, number of older siblings, number of older sisters, number of older brothers, test scores at age eight, weight at age eight, and BMI at age eight. Sampling-site fixed effects are included. Robust standard errors clustered at the level of the sampling site are reported.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

Table B.4. Time Use Patterns of Older Siblings, Proportion of Time in the Day

	(1) Work	(2) Household Chores	(3) Care	(4) Sleep	(5) Play	(6) School	(7) Study
Early menarche	0.002	-0.004	-0.000	-0.007	-0.013	-0.014	-0.016
	(0.012)	(0.004)	(0.002)	(0.020)	(0.020)	(0.016)	(0.012)
Observations	496	496	496	496	496	496	496
R^2	0.0807	0.141	0.198	0.174	0.0794	0.170	0.0678
Control average	0.0155	0.0211	0.00561	0.304	0.158	0.196	0.0577
Sampling site FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports the results from a single-difference specification (see Equation 5.3) for data from the second survey round when children were twelve. Controls include household size, caste, sex of the household head, wealth index,mom's literacy status, age of the head of the household, indicator of whether the child speaks the majority language, number of younger siblings, number of older siblings, number of older brothers, number of older sisters, test scores at age eight, weight at age eight, and BMI at age eight. Sampling-site fixed effects are included. Robust standard errors clustered at the level of the sampling site are reported.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

Table B.5. Time Use Patterns of Younger Siblings, Proportion of Time in the Day

	(1) Work	(2) Household Chores	(3) Care	(4) Sleep	(5) Play	(6) School	(7) Study
Early menarche	0.034	0.037	0.034	0.048	0.024	0.045	0.032
	(0.034)	(0.034)	(0.034)	(0.028)	(0.042)	(0.030)	(0.033)
Observations	403	403	403	403	403	403	403
R^2	0.0866	0.0817	0.0810	0.128	0.0760	0.180	0.0995
Control average	0.0155	0.0211	0.00561	0.304	0.158	0.196	0.0577
Sampling site FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports the results from a single-difference specification (see Equation 5.3) for data from the second survey round. Controls include household size, caste, sex of the household head, wealth index,mom's literacy status, age of the head of the household, indicator of whether the child speaks the majority language, number of younger siblings, number of older siblings, test scores at age eight, weight at age eight, and BMI at age eight. Sampling-site fixed effects are included. Robust standard errors clustered at the level of the sampling site are reported.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

Table B.6. Robustness Check- Effect on Young Lives Children's Siblings' Enrollment

	(1) vs Late Menarche Group	(2) vs Boys
Round 2	-0.031	-0.066
	(0.046)	(0.053)
Early menarche X Round 2	0.005	0.023
	(0.062)	(0.050)
Observations	1423	1729
R-squared	0.470	0.482
Control average	0.748	0.682
HH fixed effect	Yes	Yes
Controls	Yes	Yes

Note: This table reports the results from a difference-in-differences specification (see Equation 5.1) for comparisons that span the first and the second and the second and the third survey round. Controls include household size, caste, sex of the household head, wealth index,mom's literacy status, age of the head of the household, indicator of whether the child speaks the majority language, number of younger siblings, number of older siblings, number of older brothers, number of older sisters, test scores at age eight, weight at age eight, and BMI at age eight. Household fixed effects are included. Robust standard errors clustered at the level of the sampling site are reported.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

Table B.7. Markers of Gender Performane, by Caste

	(1) Practice purdah	(2) Can go to: Health center	(3) Can go to: Grocery shop	(4) Can go to: Friend's house
Scheduled caste	-0.046*	0.039	0.051**	0.038
	(0.025)	(0.035)	(0.022)	(0.036)
Observations	1774	1774	1774	1774
R^2	0.288	0.216	0.209	0.229
Control average	0.145	0.588	0.802	0.651
Sampling site FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Note: This table reports the results on the association between *markers of gender performance* and caste for the state of Andhra Pradesh using the second round of the nationally representative IHDS Survey. Controls include an indicator for land ownership, logged per capita consumption expenditure, age, age squared, an indicator for whether the woman married before eighteen, number of household members and number of married women between the ages of fifteen and forty-nine. Primary Sampling-site fixed effects are included. Robust standard errors clustered at the level of the primary sampling unit are reported.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

Table B.8. Normalized Test Scores in Round 2

	PPVT		Mathematics	
	(1)	(2)	(3)	(4)
	vs Late Menarche Group	vs Boys	vs Late Menarche Group	vs Boys
Early menarche	0.072	-0.022	0.121*	0.020
	(0.068)	(0.070)	(0.072)	(0.087)
Observations	497	601	498	610
R^2	0.396	0.387	0.449	0.410
Control average	-0.0983	0.0585	-0.0607	0.0470
Sampling site FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Late menarche		-0.148**		-0.117
		(0.051)		(0.069)
Observations		832		842
R^2		0.364		0.378
Control average		0.06		0.04
Girls		-0.130**		-0.079
		(0.046)		(0.068)
Observations		965		975
R^2		0.371		0.390
Control average		0.06		0.05
Sampling site FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Note: This table reports the results from a single-difference specification (see Equation 5.3). Controls include child's enrollment status, household size, caste, sex of the household head, wealth index, mom's literacy status, age of the head of the household, indicator of whether the child speaks the majority language, number of younger siblings, number of older siblings, number of older sisters, number of older brothers, test scores at age eight, weight at age eight, and BMI at age eight. Sampling-site fixed effects are included. Robust standard errors clustered at the level of the sampling site are reported.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.