Lucky Late Bloomers: The Impact of Early Marriage on Adult Outcomes in Western Kenya

Daniel L. Hicks
University of Oklahoma
hicksd@ou.edu

Joan Hamory Hicks
University of California, Berkeley CEGA
jrhamory@berkeley.edu

[Preliminary and Incomplete: Do Not Cite Without Permission]

January 2013[1]

Abstract:

Using rich data from a longitudinal survey, we examine the impact of early marriage on adult outcomes for women in rural Western Kenya, a region where men and women face markedly different social and labor market circumstances and there is a substantial gender gap across a wide range of development outcomes. The timing of physical maturation in girls has been shown to influence age at first marriage in a quasi-random manner. We find that each additional year that menarche is delayed in our study region is associated with an increase of 0.25 years in marriage age. Using age at menarche as an instrument for age at marriage, we show that delayed marriage increases educational attainment and academic test scores, while surprisingly a range of other adult outcomes such as self-reported health and wellbeing, migration decisions, asset ownership, political participation and attitudes and beliefs are unaffected. Results are similar if we instead examine the effect of age at family formation, which incorporates information on both age at first marriage and age at first pregnancy.

JEL Classification: I32, I15, J12, J13, J16, O12
Keywords: Marriage, fertility, education, labor market, gender

[1] Acknowledgements: The authors acknowledge helpful discussions with Joaquín Artés, Jennifer Graves and Kevin Grier. Any errors are our own.

I. Introduction

Narrowing the gender gap has been identified as a key component of economic development. In spite of this, health, education, and labor market outcomes for women remain significantly worse than for men in most of the developing world. To what degree can this disparity be attributed to gendered social norms and roles within the household? Adolescent marriage, teenage pregnancy, patrilocal exogamy, and weak land rights are not uncommon for girls in many regions. Women typically marry and start families younger than their male counterparts, and bear a larger physical and economic cost to household fertility decisions. Nevertheless, little research has been undertaken to directly explore how significant a causal role social institutions and norms play in hindering female development outcomes.

This paper examines the impact of early marriage and family formation using a rich longitudinal dataset of young adults in a rural district of western Kenya. To overcome omitted variables concerns, we employ a research strategy pioneered by Field and Ambrus (2008), using timing of menarche as an instrumental variable for timing of marriage.[2] Menarche represents a constraint on marriage, and thus the timing of menarche provides an exogenous source of variation by creating a lower bound on this event.[3] We confirm that later physical maturation in girls is associated with later marriage in our study area. Specifically, each additional year in which menarche is delayed is associated with an increase of 0.25 years in marriage age. In our sample, over 32% of women marry within 5 years and 55% within 10 years of reaching reproductive maturation.

We use this instrumental variables strategy to examine the impact of early marriage on a wide range of adult outcomes among Kenyan women. Surprisingly, while delayed marriage appears to increase educational attainment and improve some academic test scores, a broad range of other adult outcomes including self-reported health and wellbeing, migration decisions, asset ownership, political participation, and attitudes and beliefs are largely unaffected.

---

[2] In robustness checks, we also explore age of menarche as an instrument for age of family formation, which we define as the age at which a woman marries or becomes pregnant, whichever comes first. We find that each year that menarche is delayed is associated with an increase of 0.36 years in age of first pregnancy. Field and Ambrose focus on age at first marriage, which is very closely linked to menarcheal age in Bangladesh.

[3] Accounts of marriage practices among the Luhya and Luo, the two main ethnic groups in our study region, suggest that marriage prior to menarche is uncommon in western Kenya (Clark *et al.*, 2010). In fact, 97.6% of women in our sample report menarche before marriage.

In the context of quasi-experimental research methods, validation research is not especially common, particularly for studies which examine social institutions (see Glazerman *et al.,* 2002 for an overview). In this regard, in addition to exploring a wider range of impacts of early marriage, this paper additionally serves as a demonstration of external validity for the findings of Field and Ambrus (2008). Specifically, our findings suggest that the results for the Bangladeshi case, where adolescent and arranged marriage are more prevalent, extend to Kenya, a region characterized by later (although still comparatively early) marriage and by a different set of religions and local customs. The strength of our observed association between marriage and education suggests that their findings are likely to generalize more broadly.

This paper proceeds as follows. Section II discusses the context and examines the existing literature. Section III describes the data and presents summary statistics. Section IV discusses the empirical strategy. Section V explores the impact of early marriage on a range of socio-economic outcomes among Kenyan youth. Section VI concludes.

II. Background

Gender, Marriage, and Fertility in Rural Western Kenya

Our primary study area is Busia, a densely-settled rural farming district of western Kenya bordering Lake Victoria. Busia District is home to a relatively ethnically homogenous population – individuals in our sample are primarily Luhya (87%; the district is also home to smaller numbers of Luo, 7%, and Teso, 5%). Men and women in this region face markedly different social and labor market circumstances. Patrilocal exogamy is prevalent, and women in the region have weak land rights. As in many parts of the developing world, women tend to marry (either formally or informally) at a young age. Men traditionally pay bride price, so while women often marry shortly after leaving school (or leave school to marry), men typically work for several years between school and marriage, in part because they need to accumulate resources to marry.

We focus on a survey of individuals who at the time of data collection were primarily in their mid-twenties. Women in our study area in this age group are significantly more likely to be married than their male counterparts (65.8% versus 50.4%). Furthermore, there is a large gender gap in age at first marriage. On average 12% (1%) of women (men) are married before age 18 and 46% (17%) are married by age 25. The average male spouse in our sample is 5.1

years older than his female partner, and over 97% of unions in our data entail a positive age difference between male and female spouses. Figures 1a and 1b show the distribution of age of first marriage among married women and men, respectively, in our sample.

As might be expected in a poor agrarian society, total fertility in rural western Kenya is high, and is also characterized by a gender gap (CBS, 2004). Since we focus on a group of individuals who are early in their prime child-bearing years, total fertility in our sample is lower than the regional average, which exceeds 5. However, the gender gap is already evident at this earlier age group. Females in our sample have had more pregnancies than partners of male respondents (2.38 versus 2.05). Figures 2a and 2b plot the distribution of the number of children by gender in our sample.

Related Literature

Early marriage has been associated with earlier age at first birth, higher total fertility, lower utilization of maternal healthcare, and lower female education (Jensen and Thornton, 2003). However, rigorously identifying a relationship between early marriage and other adult outcomes as causal is difficult due to the presence of other (difficult to observe) omitted factors which likely affect both marriage timing and other adult outcomes. Such factors could include parent socio-economic status, or gender views. Following Field and Ambrus (2008), we employ an instrumental variables technique in order to identify a causal effect. While this earlier study focused on the impact of early marriage on educational attainment of women in Bangladesh, we examine a broader range of outcomes in the context of rural western Kenya.

III. Data

We employ data from the Kenya Life Panel Survey (KLPS), an unusual longitudinal survey of rural Kenyan youth which has gathered information on a wide range of outcomes (including education, health, labor market, migration, marriage, fertility and social attitudes) in three rounds of data collection between 1998 and 2012. Respondents of this panel data collection effort were involved in one of two school-based development programs which took place between 1998 and 2002. The Primary School Deworming Project (PSDP) provided deworming treatment to children in 75 primary schools in Busia District during 1998-2002, and the Girls' Scholarship Program (GSP) awarded merit-based scholarships to qualifying grades 5 and 6 girls in a separate set of primary schools in Busia District in 2001-2002.[4] Thus, KLPS respondents compose a representative subset of individuals who attended primary school in Busia District at some point during 1998-2002.[5]

We primarily utilize data collected during the third round of KLPS data collection (KLPS-3). A representative half of all KLPS respondents were interviewed in 2011-2012, and we utilize information on these individuals in the present version of this paper.[6] In our analysis, we focus on female KLPS respondents who experienced menarche between the ages of 11 and 19. These women range age from 18 to 33. In what follows we study two subsets - those who are married (60.5% of the full sample), and those who have ever been pregnant (71.3% of the full sample).

IV. Empirical Strategy

Menarche as an Instrumental Variable

Earlier physical maturation in women has been associated with earlier marriage and earlier fertility timing in previous studies (Field and Ambrus, 2008; Ghorry, 2012). The use of menarche as an IV relies on several assumptions. First, age of menarche needs to be correlated with age of marriage. Figure 3 plots mean age of marriage as a function of menarcheal age for our sample of women in Busia, Kenya. As can be seen, later menarche is strongly associated with later marriage. This can also be seen in Figure 5, which plots the distribution of age of marriage while dividing the sample into girls who undergo menarche between ages 11-14 and girls who undergo menarche between 15-19. Again, later onset of menarche is strongly associated with delays in marriage timing. Table 3 presents the results of first stage regressions, which suggest a strong, positive relationship between age of menarche and age at first marriage

---

[4] Both of these programs used randomized methods to study program impacts. For more information on the PSDP, see Miguel and Kremer (2004). For more information on the GSP, see Kremer *et al.* (2009).

[5] The 1998 Kenya Demographic and Health Survey reports that 85% of 8 to 18 year olds in western Kenya were enrolled in school at that time, indicating that our school-based sample is broadly representative of children in the region. Nevertheless, because we focus on a population of individuals who were enrolled in school, it is likely that we may be missing the poorest of the poor in the region, and perhaps some of the earliest marriages (girls who had already dropped out of school to marry). We discuss how this affects our analysis in our conclusion.

[6] The other half of respondents will be approached during 2013 and 2014.

(columns 4-6). Each additional year that menarche is delayed is associated with 0.27 years later marriage, and this relationship is robust to the inclusion of measures of earlier life health status.

The context of our results differ from those of Field and Ambrus (2008) because in the case of Bangladesh, girls are frequently subject to arranged marriages and are often married as close as possible to the exact timing of menarche in order to preserve the bride's virginity. The authors find that over 70% of first marriages occur within 2 years of menarche. In our sample, 32% of women marry within 5 years and 55% within 10 years of menarche. Although nearly 1 in 8 brides are younger than age 18 at marriage, the principal connection we observe between age of menarche and marriage likely reflects marriage market dynamics, in which individuals have more say in choosing their partners.[7] Timing of menarche then represents a firm lower bound on marriage, and an important marker for entry into the marriage market.

A second assumption of this instrumental variables strategy is that the timing of menarche is exogenous, so that common factors do not influence both the timing of physical maturation and later life outcomes. Researchers have argued that in developed nations, variation in the timing of menarche is primarily attributable to genetic differences and not environmental factors (Parent *et al.*, 2003; Ghorry, 2012). Once we examine developing countries, few environmental factors have been robustly linked to age of menarche, with the notable exception being events of extreme malnutrition. For instance, in utero and early life nutrition shocks which lead to stunting in preschool aged children has been correlated with significant increases in age at menarche.[8]

The consensus among researchers is that the timing of menarche is quasi-random with the exception of these extreme shocks to health and nutrition. A check on the association between adult height (a measure of cumulative health) and age of menarche in the KLPS sample suggests that this is not a concern, as the data reveals a weakly significant positive coefficient. This stands in contrast to the bulk of the literature which finds that improved nutrition yields earlier menarche. These findings suggest that menarche can be used as an instrument for age of marriage in this setting.

---

[7] Recent research suggests individual control over marriage partner selection is rising in Western Kenya (Clark et al. 2010).

[8] Note that the literature suggests that timing of menarche is affected by true shocks at particular times in human development, not chronic malnutrition.

However, we undertake several different strategies to alleviate any remaining concerns of omitted variables. The richness of the KLPS data allows us to do this in several ways. First, we proxy for resource availability *in utero* and during early childhood, and ability to smooth shocks to health and nutrition, by using information on parent education and childhood household assets. Second, we include cohort dummies, as year of birth may capture aggregate shocks such as droughts and floods which may have impacted nutritional availability at important stages of development. Most importantly, we include a direct measure of height for age (stunting) during youth obtained during the initial KLPS-1 survey round, which was fielded in 2003-2005 when respondents were primarily in their mid-teens. As we show below, our results are largely unaffected by these strategies.

Furthermore, we follow the practice of Field and Ambrus (2008) and limit the analysis to individuals who report age of menarche within a specified band, owing to the association between environmental factors and chronic medical conditions with very early or very late menarche outcomes. Specifically, we restrict the sample to individuals who report age of menarche between 11 and 19 which encompasses 93.7% of all respondents who are not missing information on age of menarche.[9]

Another key assumption of our analysis is that age of menarche does not affect our outcomes of interest directly, but only through age of marriage. One concern is that age of menarche also affects fertility outcomes. We undertake several strategies to address the fact that menarche influences both marriage and fertility outcomes. First, we create an additional measure, which we call "age at family formation", and define as being the minimum of (1) first age at which an individual becomes pregnant, (2) age at first marriage, and (3) age at first cohabitation (which essentially captures informal marriages). Examining the impact of family formation on our developmental outcomes is intended to capture the age at which maternal or spousal institutions and norms become relevant. We further check the robustness of our IV regressions of menarche for marriage age, by controlling for age of first birth and other fertility outcomes, and by examining the results once we exclude the subset of women who give birth before marriage (formal or informal).

[9] Field and Ambrus (2008) restrict their analysis to those reporting age of menarche between 11-16, which they report covers 90% of the sample. In the KLPS data, to capture 90% of the sample would require including ages 1A through 1B. Recent research has documented variation in the timing of menarche across countries and over time possibly owning to nutrition, stress, climactic and environmental conditions including weather and altitude, and exposure to endocrine disruptors (Parent *et al.*, 2003).

A final concern is that the information we use on menarcheal age is based on interviewee recollection and thus subject to recall error. At the same time, menarche is a significant life event, and most research suggests that accuracy of recall data on this topic is high. Exploring findings from other studies is reassuring here. A number of studies have examined the timing of menarche among African populations, finding that mean age of menarche typically ranges from 12 and 16, with most studies of rural populations documenting mean ages of 14 or 15.[10] In Mumias and Asembo, two small towns in western Kenya located near the primary study district, Leenstra *et al.* (2005) document median menarcheal age ranging from 14.6 to 15.1. This is consistent with the results from our survey for Busia (approximate mean and median age 15).

Estimation Equations

[to be written – discussion of estimation equations, controls]

V. Results [INCOMPLETE]

The second stage results from our instrumental variables strategy for educational outcomes are presented in Table 4. This table contains two specifications for each educational outcome of interest, the first instrumenting for family formation and the second for age of first marriage. Focusing first on column 1 of Panel A, each additional year of delay in family formation (the timing of first marriage or pregnancy) is associated with an increase in educational attainment of 0.49 years. Similarly, the results presented column 2 suggest that focusing solely on age of marriage, each additional year in age at first marriage for women in this region is associated with an increase in female educational attainment of 0.66 years. The remaining columns of Panel A include binary outcome variables, and these coefficients can be interpreted as percentage point increases in the likelihood of completion or attendance respectively. Moving across the table, we can see that delayed marriage and pregnancy is associated with a higher probability of completing primary school (6.8% points), attending secondary school (6.2% points), and completing secondary school (5.8% points).

[10] For a general overview see (Padez, 2003). Recent country specific studies include: Sudan (Aziem *et al.*, 2011), Nigeria (Adebara and Ijaiya, 2012), Ethiopia (Zegeye *et al.*, 2009), Gambia (Prentice *et al.*, 2010), Senegal (Simondon, 1998) and Kenya (Leenstra *et al.,* 2005).

Panel B explores indicators of academic performance. [Discuss Math and Reading Tests]. Columns 1 - 4 suggest that delayed marriage and pregnancy are associated with weakly significant gains in math scores, and no improvement in reading scores. Columns 5 and 6 examine girls propensity to sit for the Kenyan Certificate of Primary Education (KCPE). The KCPE examination is supervised by the Ministry of Education and results are used for admission into secondary school.[11] Column 6 suggests that each additional year in age at first marriage is associated with a 13.1% point increase in the likelihood of taking the KCPE test. The final two columns repeat the exercise for the KCSE. The results are weaker for these exams, but consistent with those for the primary school exams - delayed marriage and fertility net better educational outcomes.

Table 5 continues the IV analysis, exploring the impact of delayed marriage and family formation on reported beliefs and attitudes. KLPS respondents were asked to agree or disagree with a range of statements such as" Women should equal rights to men," "Family decisions should be made by men", "Ethnic identity is not very important." Panel A reports a small subset of these questions examining attitudes towards traditional gender roles and social norms. None of these beliefs appear to be impacted by age at first marriage or family formation. Panel B replicates this exercise for a subset of questions on a range of social institutions such as religion and ethnic identification. In all cases except one, we find no significant impact of delayed marriage. The exception is for the statement "Most people can be trusted" where we find a weakly significant decrease in reported trust levels among girls who marry at a later age.

VI. Conclusion [To be Added]

[11] See Ozier (2011) for a comprehensive discussion of the KCPE and KCSE.

References

Baird, Sarah, Joan Hamory Hicks, Michael Kremer, and Edward Miguel (2011). "Worms at Work: Long-run Impacts of Child Health Gains." Mimeo.

Baird, Sarah, Joan Hamory Hicks, and Edward Miguel (2008). "Tracking, Attrition and Data Quality in the Kenyan Life Panel Survey Round 1 (KLPS-1)," *Working Paper C08-151, Center for International and Development Economics Research*, University of California at Berkeley.

Central Bureau of Statistics (CBS) (2004). "Kenya, Ministry of Health (MOH), and ORC Macro." *Kenya Demographic and Health Survey*, 2003. Calverton, MD: CBS, MOH, and ORC Macro.

Clark, Shelley, Caroline Kabiru, and Rohini Mathur (2010). "Relationship Transitions among Youth in Urban Kenya" *Journal of Marriage and Family*, Vol. 72, pp.73-88.

Duflo, Esther, Pascaline Dupas, and Michael Kremer (2011). "Education, HIV, and Early Fertility: Experimental Evidence from Kenya" *Working Paper,* August 25.

Ferré, Céline (2009): "Age at First Child: Does Education Delay Fertility Timing? The Case of Kenya," *Policy Research Working Paper 4833, The World Bank.*

Field, Erica, and Attila Ambrus (2008). "Early marriage, age of menarche, and female schooling attainment in Bangladesh." *Journal of Political Economy* Vol. 116(5), pp. 881-930.

Friedman, Willa, Edward Miguel, Michael Kremer and Rebecca Thornton (2011). "Education as Liberation?" *NBER Working Paper* #16939.

Gorry, Devon Haskell (2012). "The Effect of Menarche on Education: Explaining Black-White Differences." *Working Paper.*

Glazerman, S., D. Levy, and D. Myers (2003) "Nonexperimental Replications of Social Experiments: A Systematic Review." Princeton, NJ: Mathematica Policy Research, Inc.

Ikamari, Lawrence (2005). "The effect of education on the timing of marriage in Kenya." *Demographic Research*, Vol. 12(1), pp 1-28.

Kremer, Michael, Edward Miguel, and Rebecca Thornton (2009). "Incentives to Learn" *Review of Economics and Statistics*, Vol. 91(3), pp. 437-456.

Jensen, Robert, and Rebecca Thornton (2003). "Early Female Marriage in the Developing World." *In Gender, Development and Marriage*, edited by Caroline Sweetman. Oxford: Oxfam GB.

Leenstra T, Petersen LT, Kariuki SK, Oloo1 AJ, Kager PA, Ter Kuile FO (2005) "Prevalence and severity of malnutrition and age at menarche; crosssectional studies in adolescent schoolgirls in western Kenya." *European Journal of Clinical Nutrition*, Vol. 59, pp41-48.

Miguel, Edward and Michael Kremer (2004). "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica*, 72(1), 159-217.

Ozier, Owen (2011). “The impact of secondary schooling in Kenya: a regression discontinuity analysis” *Working Paper.*

Padez, C. (2003). “Age at Menarche of Schoolgirls in Maputo, Mozambique,” Annals of Human Biology, Vol. 30, 487–495.

Parent, A. S., G. Teilmann, A. Juul, N. E. Skakkebaek, J. Toppari, and J. P. Bourguignon (2003). “The Timing of Normal Puberty and the Age Limits of Sexual Precocity Variations Around the World, Secular Trends and Changes After Migration,” *Endocrine Reviews*, Vol. 24, pp. 668–693.

Prentice S, Fulford AJ, Jarjou LM, Goldberg GR, Prentice A (2010) "Evidence for a downward secular trend in age of menarche in a rural Gambian population." *Annals of Human Biology*, Vol. 37(5), pp. 717-721.

Simondon, K. B., Simondon, F., Simon, I., Diallo, A., Benefice, E., Traissac, P. & Maire, B. (1998) "Preschool stunting, age at menarche and adolescent height: a longitudinal study in rural Senegal." *European Journal of Clinical Nutrition* Vol. 52, 412–418.

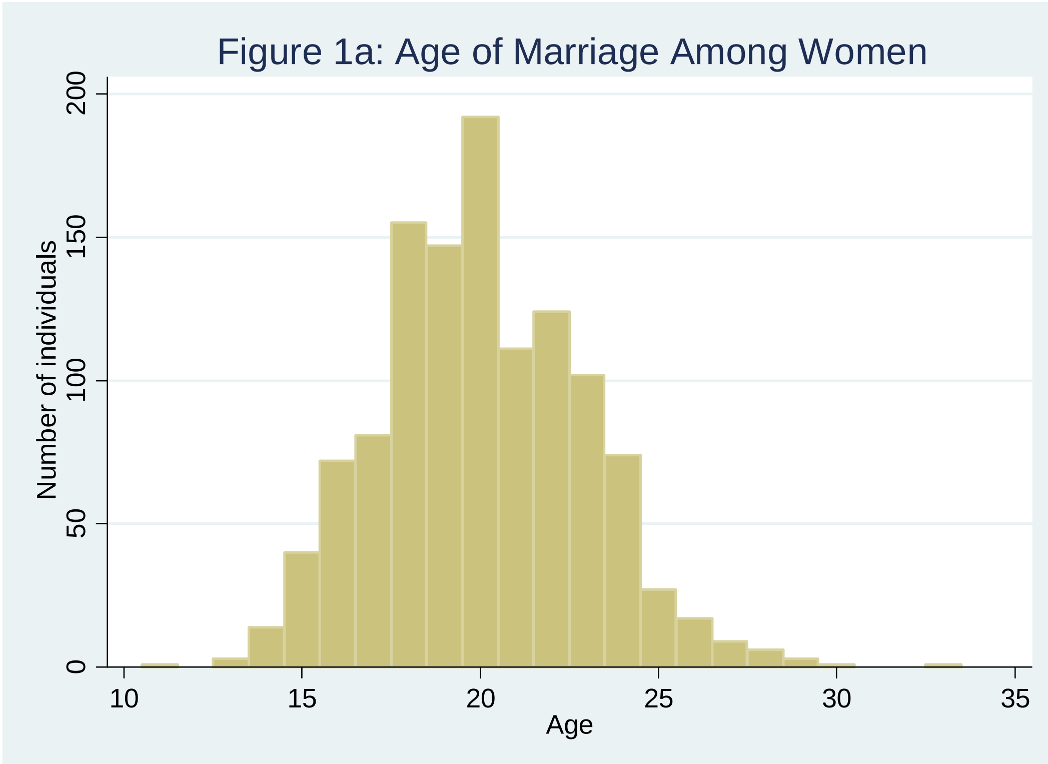


Figure 1a: Age of Marriage Among Women
Number of individuals
0
50
100
150
200
10
15
20
25
30
35
Age

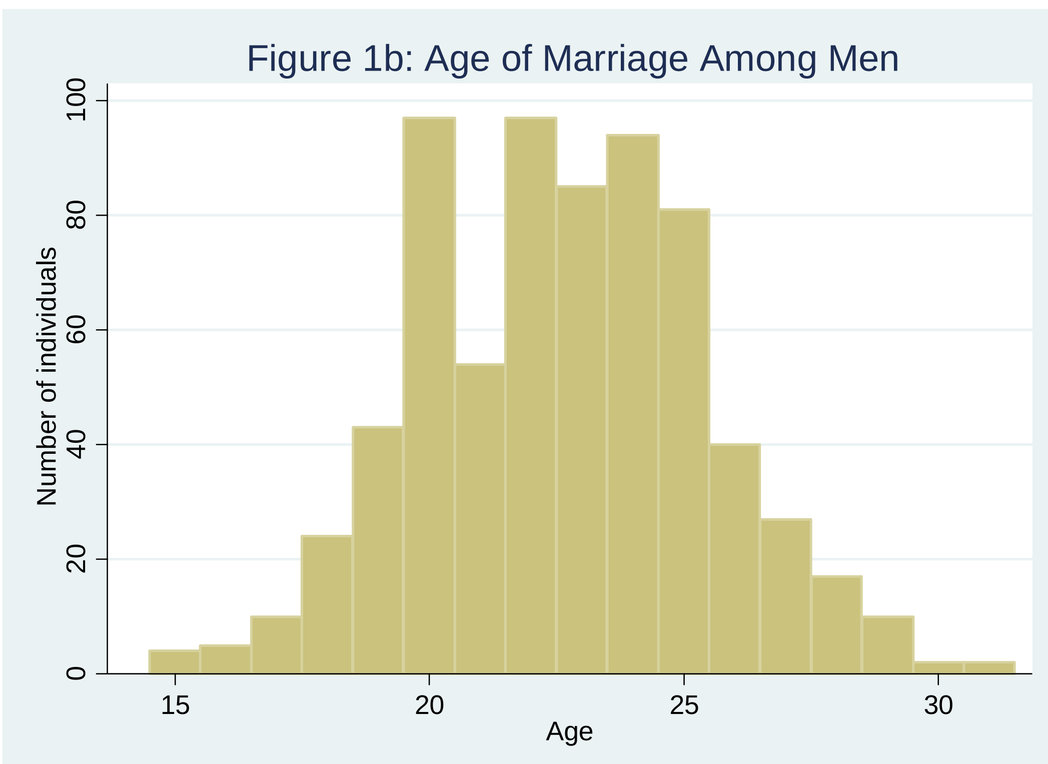


Figure 1b: Age of Marriage Among Men
Number of individuals
0
20
40
60
80
100
15
20
25
30
Age

Figure 2a: Number of Children Among Women

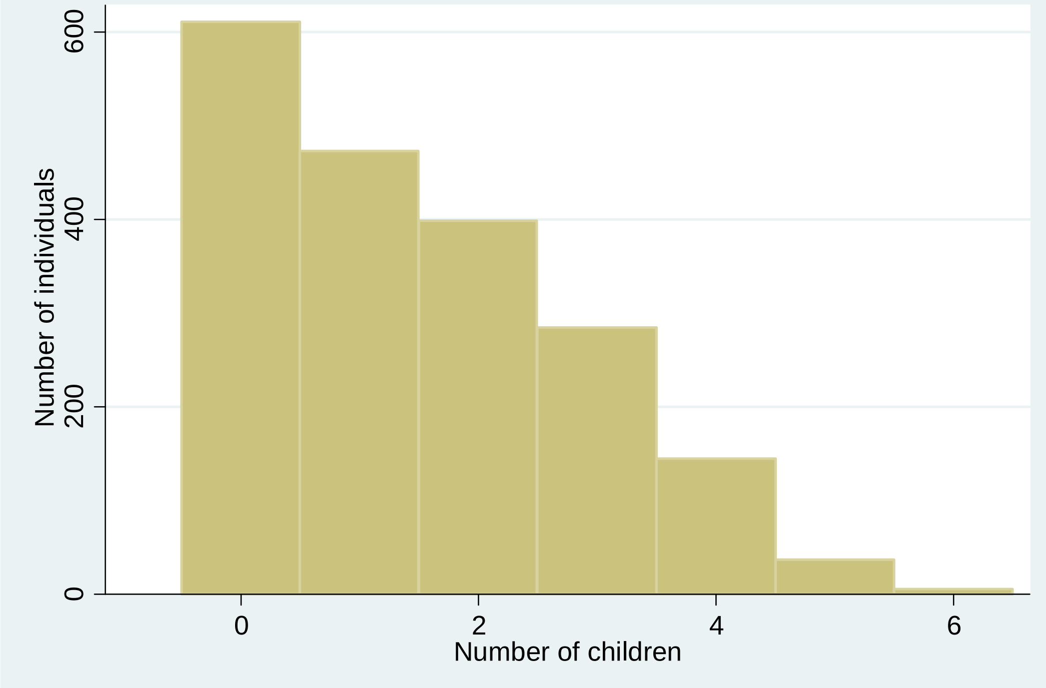


Figure 2b: Number of Children Among Men

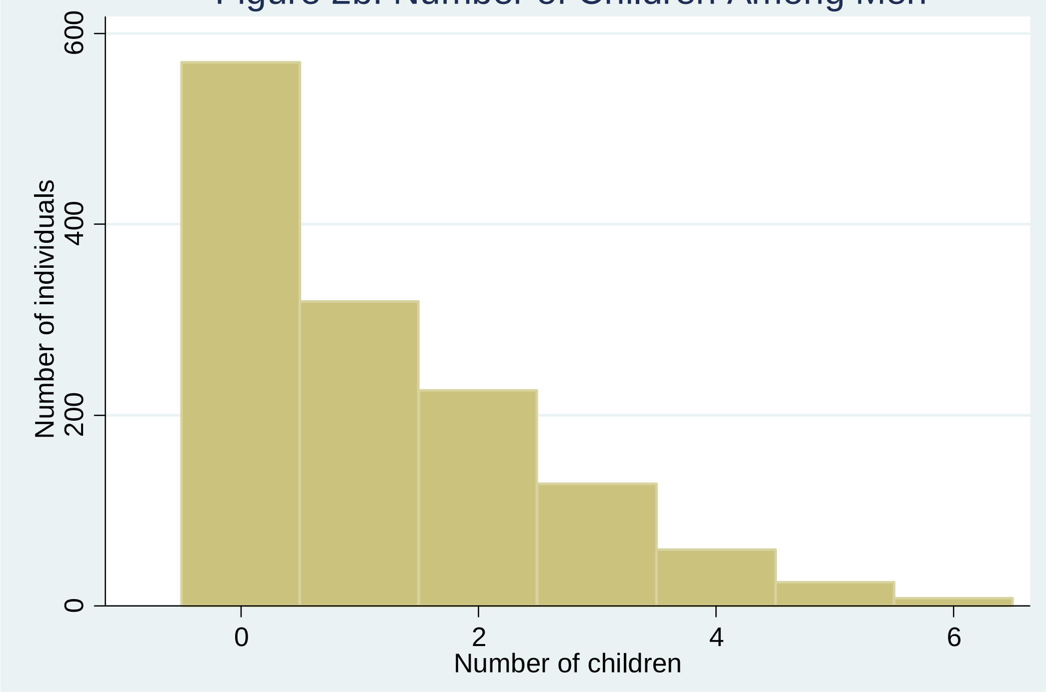


Figure 3: Age of First Marriage and Age of Menarche

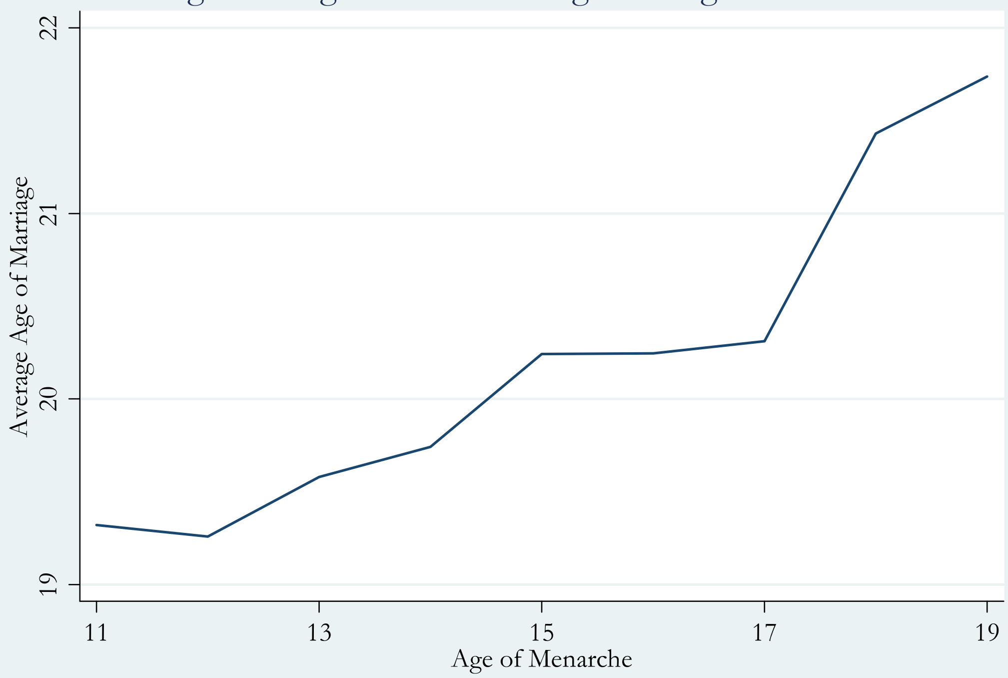


Figure 4: Age of First Pregnancy and Age of Menarche

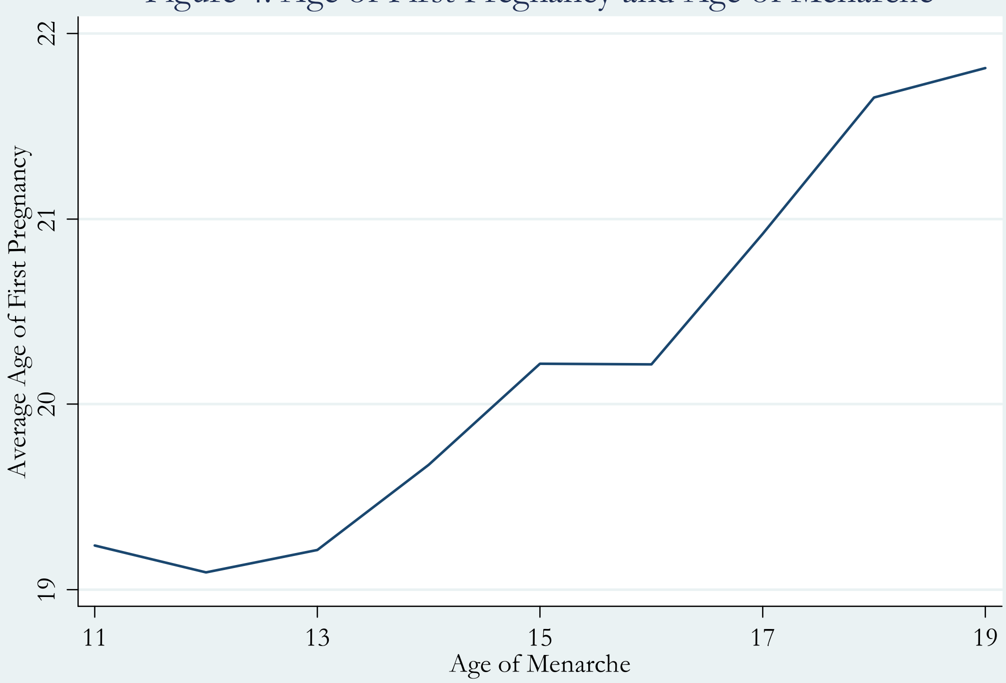


Figure 5: Distribution of Age of Marriage by Age of Menarche

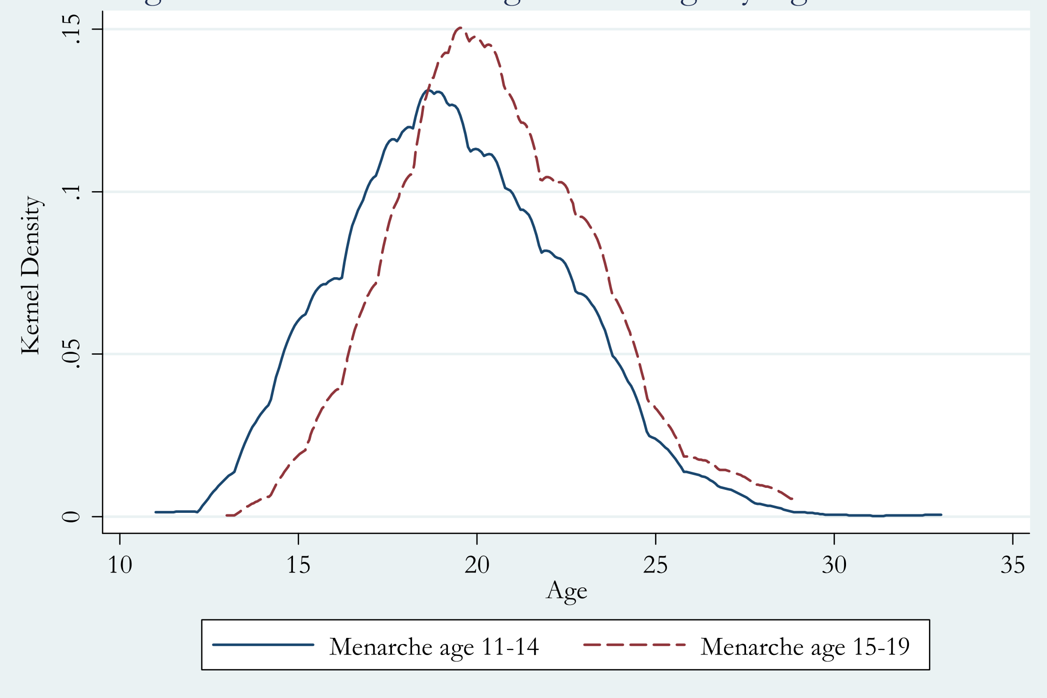
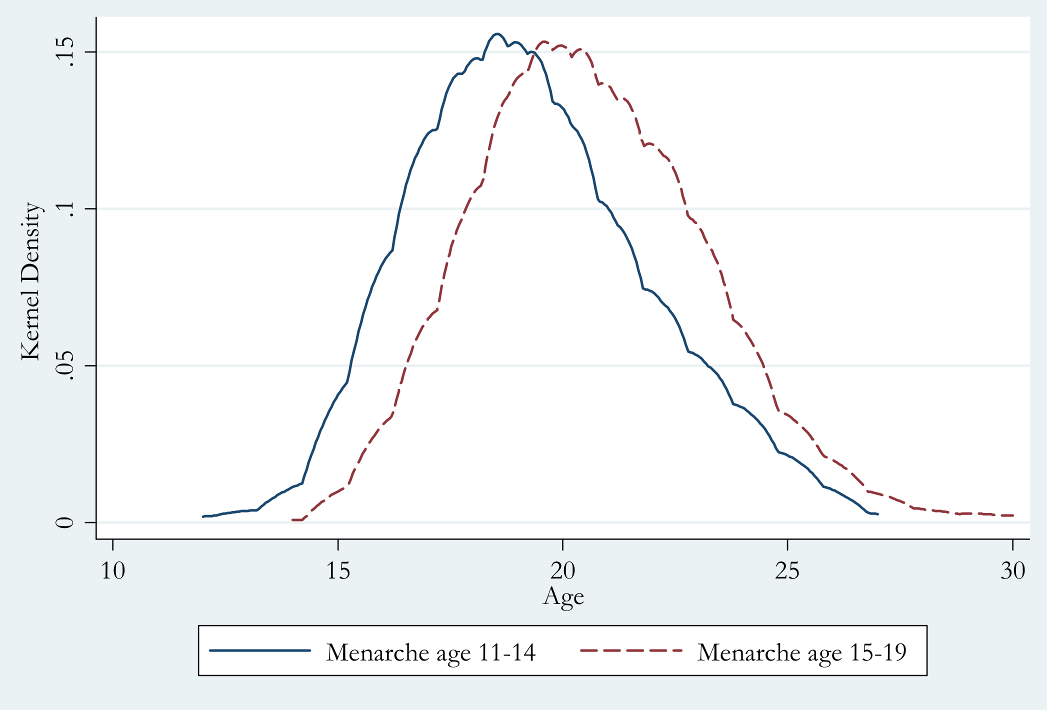


Figure 6: Distribution of Age of First Pregnancy by Age of Menarche



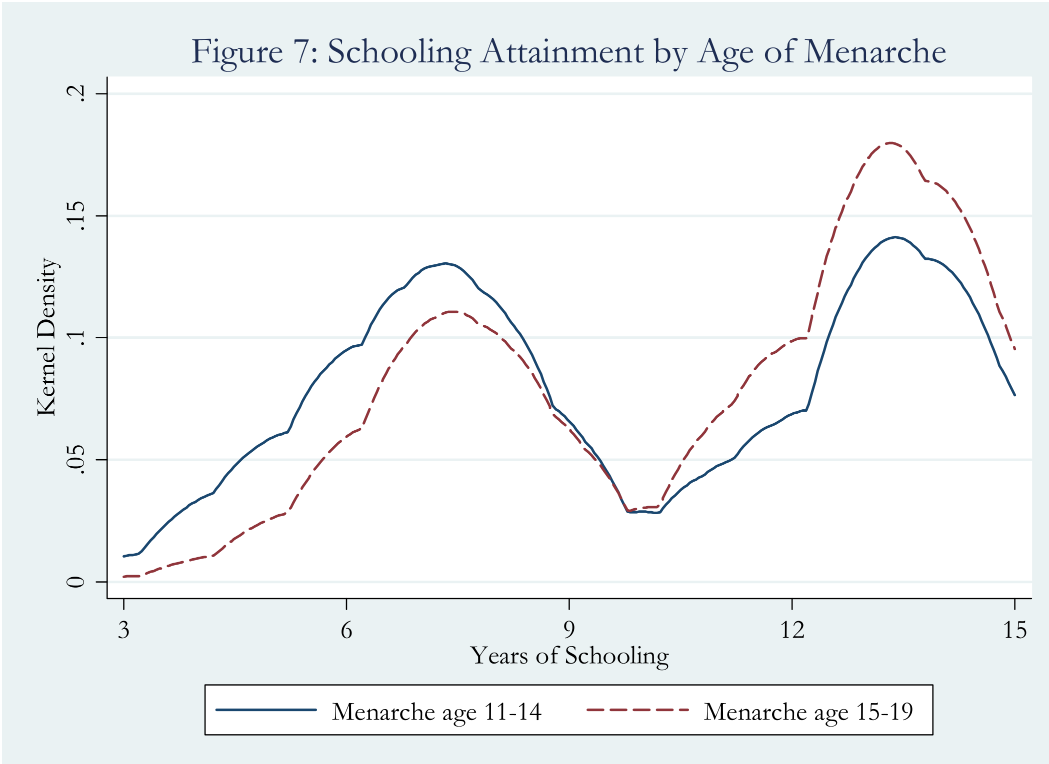


Figure 7: Schooling Attainment by Age of Menarche
Kernel Density
0
.05
.1
.15
.2
3
6
9
12
15
Years of Schooling
Menarche age 11-14
Menarche age 15-19

Table 2: Summary Statistics

Variable	Obs	Mean	Standard Deviation
Individual Characteristics			
Age of Menarche	1956	14.88	1.87
Height (cm), at KLPS-3 Survey	1879	162.30	6.37
Height Z-Score at KLPS-1 Survey	1315	-0.02	1.11
PSDP Sample	1956	0.72	0.45
PSDP Treatment	1956	0.48	0.50
GSP Treatment	1956	0.14	0.35
Age at KLPS-3 Survey	1956	24.91	2.46
Marriage and Fertility Outcomes			
Ever Married	1956	0.65	0.48
Age of First Marriage	1180	20.00	2.93
Ever Pregnant	1956	0.75	0.43
Age of First Pregnancy	1360	20.00	2.61
Age of Family Formation (First Marriage or Pregnancy)	1408	19.38	2.64
Household Characteristics			
Father's Education (years)	1956	8.03	3.57
Mother's Education (years)	1956	5.94	3.55
Number of Siblings	1955	5.43	2.82
Educational Outcomes			
Years of Education	1939	10.45	3.24
Attended Primary (1/0)	1956	1.00	0.00
Completed Primary (1/0)	1956	0.75	0.43
Attended Secondary (1/0)	1956	0.60	0.49
Completed Secondary (1/0)	1956	0.51	0.50
Attended College or Vocational (1/0)	1939	0.08	0.27
Completed College of Vocation or Attended Univ (1/0)	1939	0.29	0.46
Completed University	1939	0.01	0.08
Reading Score	1956	0.47	0.88
Math Score	1956	0.43	0.77
Mean effect for Reading and Test Score	1956	0.45	0.75
Took Kenyan Certificate of Secondary Examination (KCSE)	1956	0.29	0.45
Passed Kenyan Certificate of Secondary Examination (KCSE)	644	0.39	0.49
Took Kenyan Certificate of Primary Examination (KCSE)	1956	0.65	0.48
Took Kenyan Certificate of Primary Examination (KCSE)	1399	0.74	0.44

Notes:

Table 3: First Stage Regression Results

DepVar:	Age of Family Formation			Age of First Marriage		
	(1)	(2)	(3)	(4)	(5)	(6)
Age of menarche	0.354***	0.335***	0.344***	0.272***	0.251***	0.259***
	(0.051)	(0.062)	(0.051)	(0.053)	(0.067)	(0.054)
Years of schooling completed by father	0.041	0.010	0.052*	0.070*	0.055	0.086**
	(0.028)	(0.032)	(0.029)	(0.036)	(0.048)	(0.036)
Years of schooling completed by mother	0.080**	0.057	0.067*	0.063	0.053	0.046
	(0.034)	(0.039)	(0.035)	(0.040)	(0.050)	(0.043)
Number of siblings	0.019	0.011	0.008	0.071*	0.056	0.062
	(0.031)	(0.041)	(0.031)	(0.043)	(0.048)	(0.043)
Height for age z-score, KLPS-1		0.232**			0.185	
		(0.092)			(0.121)	
Height, cm, KLPS-3			0.010			0.002
			(0.012)			(0.018)
PSDP and GSP Controls	Y	Y	Y	Y	Y	Y
Number of observations	1,407	907	1,352	1,179	749	1,134
R2	0.109	0.106	0.104	0.097	0.091	0.096

Note: Main analysis sample includes KLPS and PSDP surveys. Sources for information on covariates are described in the text. Robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: IV Regressions of Education Attainment on Family Formation and Age of Marriage

Panel A:	Years of Educ		Completed Primary		Attended Secondary		Completed Secondary	
DepVar:	FF	AM	FF	AM	FF	AM	FF	AM
Age at family formation	0.487***	0.658**	0.068***	0.096**	0.062**	0.081**	0.058**	0.070*
	(0.187)	(0.265)	(0.026)	(0.043)	(0.026)	(0.035)	(0.025)	(0.036)
Years of schooling completed by father	0.096**	0.059	0.005	-0.000	0.016***	0.013	0.014***	0.009
	(0.041)	(0.058)	(0.005)	(0.007)	(0.006)	(0.008)	(0.005)	(0.008)
Years of schooling completed by mother	0.023	0.022	0.003	0.004	-0.001	-0.001	0.005	0.006
	(0.031)	(0.042)	(0.005)	(0.007)	(0.004)	(0.005)	(0.004)	(0.005)
Height, cm	0.021	0.024	0.004	0.004	0.005*	0.005*	0.001	0.001
	(0.018)	(0.021)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Number of observations	1,336	1,118	1,352	1,134	1,352	1,134	1,352	1,179
R2	0.193	0.106	0.189	0.035	0.185	0.144	0.153	0.045

Panel B:	Reading Score		Math Score		Took KCPE Exam		Took KCSE Exam	
DepVar:	FF	AM	FF	AM	FF	AM	FF	AM
Age at family formation	-0.009	-0.058	0.063*	0.077	0.099***	0.131***	0.028*	0.011
	(0.048)	(0.072)	(0.036)	(0.055)	(0.025)	(0.046)	(0.014)	(0.019)
Years of schooling completed by father	0.034***	0.042***	0.017**	0.016*	0.005	-0.001	0.014***	0.013**
	(0.012)	(0.015)	(0.009)	(0.010)	(0.005)	(0.007)	(0.005)	(0.006)
Years of schooling completed by mother	0.011	0.010	0.008	0.006	0.002	0.001	0.002	0.005
	(0.011)	(0.013)	(0.010)	(0.013)	(0.006)	(0.009)	(0.005)	(0.005)
Height, cm	0.012**	0.012**	0.007*	0.009**	0.003	0.004	0.001	0.001
	(0.005)	(0.006)	(0.004)	(0.004)	(0.003)	(0.003)	(0.002)	(0.002)
Number of observations	1,352	1,134	1,352	1,134	1,352	1,134	1,352	1,134
R2	0.064	-0.040	0.127	0.080	0.229	-0.058	0.165	0.101

Note: Main analysis sample includes KLPS and PSDP surveys. Sources for information on covariates are described in the text. Robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: IV Regressions of Attitudes and Beliefs on Family Formation and Age of Marriage

Panel A: Gender Attitudes	Women should have equal rights		No one has a right to use physical violence		A husband should help with chores if a wife works		Family decisions should be made by men		Women can make good politicians	
DepVar:	FF	AM	FF	AM	FF	AM	FF	AM	FF	AM
Age at family formation	-0.026	-0.037	-0.028	-0.037	0.016	0.018	0.030	0.031	0.000	0.003
	(0.020)	(0.031)	(0.020)	(0.031)	(0.021)	(0.031)	(0.025)	(0.035)	(0.026)	(0.040)
Years of schooling (father)	0.007	0.010	-0.003	0.010	0.004	0.004	-0.002	-0.006	-0.003	-0.004
	(0.005)	(0.007)	(0.004)	(0.007)	(0.004)	(0.005)	(0.005)	(0.007)	(0.005)	(0.007)
Years of schooling (mother)	-0.001	-0.001	0.001	-0.001	-0.001	-0.002	-0.007	-0.006	0.004	0.005
	(0.003)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)	(0.004)	(0.005)
Height, cm	0.001	0.001	-0.001	0.001	0.002	0.003	0.001	0.001	0.006**	0.007***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)
Number of observations	1,352	1,134	1,352	1,134	1,352	1,134	1,351	1,133	1,352	1,134
R2	-0.040	-0.120	-0.041	-0.120	0.024	0.011	0.049	0.013	0.023	0.030

Panel B: General Attitudes	Ethnic identity is not "very important"		Religious identity is not "very important"		Agree with regular/open/honest elections		Democracy is preferable kind of government		Most people can be trusted	
DepVar:	FF	AM	FF	AM	FF	AM	FF	AM	FF	AM
Age at family formation	-0.001	-0.002	0.001	0.002	-0.002	-0.003	0.025	0.027	-0.024*	-0.041**
	(0.006)	(0.007)	(0.001)	(0.002)	(0.016)	(0.025)	(0.030)	(0.047)	(0.013)	(0.020)
Years of schooling (father)	0.000	-0.000	-0.000	-0.001	-0.002	-0.002	0.006	0.005	-0.001	0.001
	(0.001)	(0.001)	(0.000)	(0.000)	(0.002)	(0.003)	(0.005)	(0.007)	(0.003)	(0.004)
Years of schooling (mother)	-0.002*	-0.001	0.000	0.000	0.001	0.001	-0.002	-0.004	0.004	0.003
	(0.001)	(0.001)	(0.000)	(0.000)	(0.002)	(0.002)	(0.005)	(0.006)	(0.004)	(0.004)
Height, cm	-0.000	-0.000	-0.000	-0.000	-0.001	-0.001	0.000	-0.000	0.003**	0.003**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.003)	(0.003)	(0.001)	(0.002)
Number of observations	1,350	1,133	1,347	1,129	1,351	1,133	1,351	1,133	1,352	1,134
R2	0.012	0.002	-0.000	-0.006	0.015	0.024	0.033	0.027	-0.023	-0.171

Note: Main analysis sample includes KLPS and PSDP surveys. Sources for information on covariates are described in the text. Robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: IV Regressions of Attitudes and Beliefs on Family Formation and Age of Marriage

Panel A: Gender Attitudes	**Living with Inlaws (Patrilocal Exogamy)**		**Age of spouse at marriage**		**Own age - spouse's age, at marriage**		**Educational level of spouse (years)**		**Own education - spouse's education**	
DepVar:	**FF**	**AM**	**FF**	**AM**	**FF**	**AM**	**FF**	**AM**	**FF**	**AM**
Age at family formation	-0.009	-0.008	0.598	0.763	0.185	0.237	-0.005	0.008	0.517*	0.658*
	(0.015)	(0.023)	(0.437)	(0.518)	(0.397)	(0.518)	(0.224)	(0.292)	(0.269)	(0.340)
Years of schooling (father)	-0.011***	-0.012***	0.174**	0.137	-0.125*	-0.137	0.029	0.029	0.061	0.032
	(0.004)	(0.004)	(0.075)	(0.088)	(0.070)	(0.088)	(0.040)	(0.043)	(0.050)	(0.063)
Years of schooling (mother)	0.010*	0.011*	-0.111	-0.101	0.098	0.101	0.120**	0.117**	-0.102*	-0.094
	(0.005)	(0.006)	(0.074)	(0.070)	(0.069)	(0.070)	(0.047)	(0.046)	(0.059)	(0.069)
Height, cm	-0.002	-0.002	-0.009	-0.007	0.006	0.007	0.007	0.009	0.015	0.017
	(0.002)	(0.002)	(0.037)	(0.036)	(0.036)	(0.036)	(0.019)	(0.019)	(0.021)	(0.024)
Number of observations	1,352	1,134	1,105	1,104	1,104	1,104	1,126	1,124	1,110	1,108
R2	0.022	0.026	0.091	0.159	0.050	0.054	0.035	0.041	-0.060	-0.190

Panel B: General Attitudes	**Felt ready to marry at time of marriage**		**Brideprice paid**		**Marriage is informal**		**Still married to first spouse**		**Has a cowife**	
DepVar:	**FF**	**AM**	**FF**	**AM**	**FF**	**AM**	**FF**	**AM**	**FF**	**AM**
Age at family formation	-0.012	-0.013	-0.043*	-0.056*	0.062**	0.081**	0.043*	0.056*	0.019	0.024
	(0.030)	(0.040)	(0.025)	(0.032)	(0.029)	(0.040)	(0.022)	(0.031)	(0.025)	(0.032)
Years of schooling (father)	0.001	0.002	-0.004	-0.002	0.005	0.002	0.007	0.005	-0.009	-0.010*
	(0.006)	(0.007)	(0.006)	(0.007)	(0.007)	(0.008)	(0.005)	(0.006)	(0.006)	(0.006)
Years of schooling (mother)	0.006	0.005	0.005	0.005	-0.008	-0.008	-0.010*	-0.010*	0.000	0.001
	(0.006)	(0.006)	(0.005)	(0.005)	(0.009)	(0.010)	(0.005)	(0.006)	(0.005)	(0.005)
Height, cm	0.004	0.004	0.001	0.001	0.000	0.001	-0.000	-0.000	-0.004*	-0.004*
	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Number of observations	1,131	1,129	1,134	1,132	1,137	1,134	1,136	1,134	1,118	1,116
R2	-0.029	-0.047	-0.021	-0.029	-0.116	-0.187	0.049	-0.002	-0.016	-0.027

Note: Main analysis sample includes KLPS and PSDP surveys. Sources for information on covariates are described in the text. Robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

Male Earnings Inequality and the Age of Marriage of Women in India

A. V. Chari[*] Annemie Maertens[†] Sinduja Srinivasan[‡]

PRELIMINARY DRAFT

Abstract

We provide the first evidence in a developing country context of the causal effect of widening income inequality on the age of marriage of women. The logic is similar to that of a job search model. In this case, widening income inequality makes high-earning men relatively more attractive candidates than before, and increases the willingness of the woman to wait for offers from such candidates, and thereby delays her marriage. We utilize data from India to show that this mechanism does indeed have appeared to have increased the average age at marriage. We are able to exploit the richness of the data to rule out a number of alternative explanations for our findings.

1 Introduction

The phenomenon of early marriage of women is widespread in developing countries. Data from various Demographic and Health Surveys (DHS) conducted between 2003 and 2005 reveal that

[*]Rand Corporation. Email: achari@rand.org

[†]Graduate School of Public and International Affairs, University of Pittsburgh. Email: maertens@pitt.edu

[‡]Pardee Rand Graduate School. Email: sinduja@rand.org

as many as 27.1% of women in India between the ages of 15 and 19 are married (the legal minimum age is 18 years). The corresponding figures are 46.1% in Bangladesh, 42% in Chad, 32.9% in Malawi, 50.4% in Mali, 38.2% in Mozambique and 31.7% in Nigeria.

The available data also indicate, however, that the age of marriage of women has been gradually increasing, although there is some heterogeneity in these trends between countries in this regard (Singh and Samara 1996, Harwood-Lejeune 2000, Choe et al 2001, Heaton et al 2002, Rashad and Osman 2003, Westoff 2003, Jensen and Thornton 2003). The trend towards later marriage has been hypothesized to reflect factors such as better education (as well as better access to education), improved labor market opportunities, demographic change, urbanization and exposure to mass media (see for example Jensen and Thornton 2003 and Mensch, Singh and Casterline 2006).

A less obvious mechanism, and one that has not been evaluated in the developing country context, is that increasing income inequality may delay marriage decisions. This mechanism may be understood in the context of a search model of marriage-market matching as in Loughran (2000 & 2002). In this model, a woman who is on the marriage market sequentially samples marriage offers from prospective spouses, evaluating them in terms of their income. In deciding whether to accept a particular offer, she must decide whether the benefits from waiting for a better offer outweigh the costs of staying unmarried during this period. Widening income inequality makes high-earning men relatively more attractive candidates than before, and increases the willingness of the woman to wait for offers from such candidates, and thereby delays her marriage. Loughran (2002) and Gould and Paserman (2003) provide evidence that this mechanism has significantly reduced marriage rates in the United States over the last few decades.

Utilizing nationally representative household survey data from India, we provide the first evidence for such a phenomenon in the context of a developing country. The postponement of marriage is of acute significance in developing countries. In addition to its role in the fertility transition, studies have found that the woman's age at marriage correlates strongly with a number of later health outcomes (Bruce 2003, Clark 2004, Nour 2006, Raj et al 2009, Santhya et al 2010). Early marriage has been associated with low contraceptive use, miscarriages, multiple unwanted pregnancies, domestic violence, depression and even increased HIV risk.

The role of income inequality in delaying marriage is particularly relevant in the Indian setting because the turn towards economic liberalization taken during the 1990s appears to have increased income inequality, both over time as well as in spatial terms (Banerjee and Piketty 2001, Deaton and Dreze 2002, Sundaram and Tendulkar 2003a and 2003b, Sen and Himanshu 2005 and Pal and Ghosh 2007). The available evidence indicates that this growing inequality has been largely driven by an increase in incomes at the upper tails of the income distribution, while incomes of other groups have stagnated. Banerjee and Piketty (2005) use income tax data to show that incomes of the top 1% in India increased by 50% in the 1990s, and that their share of total income nearly doubled.

To evaluate the link between earnings inequality and marriage in India, we relate (age-specific) marriage rates to male earnings inequality in the relevant marriage market, which we define as including all out-of-school unmarried adults of marriageable age belonging to the same caste and residing in the same state. Our definition of marriage markets takes into account the fact that inter-caste marriages in India are rare, and that ethno-linguistic differences between states limit the geographical boundaries of the marriage market. This natural segmentation of marriage markets in India provides substantial variation in the marriage market experiences of women, and we are able to utilize this variation to establish the effect of earnings inequality on marriage while controlling for spatial and inter-caste differences in socio-economic characteristics that may influence marriage decisions.

Consistent with the predictions of the marital search model, we find that women in marriage markets characterized by higher male earnings inequality have significantly lower (age-specific) rates of marriage. While Loughran (2002) and Gould and Paserman (2003) examine the effect of income inequality on marriage-rates, we are able to go a step further and establish the extent of time to which marriage is postponed. We are able to do so because our data contain a subsample of ever-married women for whom we know the age of marriage. The results indicate that a doubling of the earnings of men at the 90th percentile relative to median male earnings reduces the (age-specific) propensity of women to get married, resulting in an average delay of marriage by nearly 2 years.

We proceed to consider and rule out a number of alternative hypotheses that may explain the results. Specifically, we show that the results are not due to (i) Earnings dispersion proxying for educational premia that may increase the incentives to obtain higher education (and thereby delay marriage), (ii) Increased duration of labor market search on the part of men as a response to greater dispersion in the potential male earnings dispersion, (iii) Increased duration of marriage market search on the part of men as a response to greater dispersion in female earnings, or (iv) Regional or caste-based differences in marriage norms. We are also able to show that women in marriage markets with high inequality are on average not different from women in low inequality markets, in terms of characteristics such as height and the age at which they reached menarche. Because neither of these variables represents a short-term decision that should be affected by earnings inequality, they are suitable for testing for differences in observable characteristics between women in high- and low-inequality markets.

An important benefit of postponing marriage in the developing country setting is that it can increase female educational attainment (Brien and Lillard 1994, Lindstrom and Brambila Paz 2001, Fields and Ambrus 2008). Consistent with this prediction, we find that women in marriage markets with greater male earnings inequality also obtain more education. This additional education is concentrated at the point in the educational trajectory at which most women find themselves when they are on the marriage market, i.e. greater male earnings inequality results in increased rates of matriculation and high-school graduation among women, but does not correlate either with the probability of receiving zero education or the probability of completing primary school. We also show that greater earnings inequality does not affect the educational attainment or school enrolment of children who are too young to be on the marriage market. These results confirm that male earnings inequality impacts female educational attainment only at the time of marital search and not before.

Our conclusion is that there is a strong link between male earnings inequality and female age at marriage, and that this relationship indeed appears to reflect the effect of inequality on marital search duration. In turn, the increase in search duration appears to improve educational attainments, resulting in significantly higher rates of matriculation and high-school completion. While our data do not allow for an analysis of trends over time, the results in this paper suggest the strong possibility that widening income inequality in India may have had some salutary effects working through the marriage market.

The paper proceeds as follows. Section 2 outlines a simple model of marital search to motivate the empirical analysis. Section 3 describes the data and the variables to be used in the analysis. Sections 3 and 4 present the empirical strategy and the results and Section 5 concludes.

2 Theoretical Model

This section presents a simple model that is used to derive predictions about how the duration of a woman's marriage market search is related to the mean and the variance of the male earnings distribution. In substance, the model is no different from a standard labor market search model. For simplicity of exposition, therefore, we present a simple version of the model that embodies some functional form assumptions, while referring readers to Mortensen (1986) for the derivation of the results in the more general case.

We model the marriage market search as an extension of a simple labor market search model (see among others, McCall 1970), in which an infinitely-lived woman sequentially samples marriage offers from a known distribution of offers, before finally accepting an offer. As in Loughran (2002), we assume that women prefer high-earning males over low-earnings males and that this is the only dimension of groom heterogeneity, so that the male earnings distribution characterizes the distribution of offers. However, the empirical analysis allows the marriage market to be structured horizontally along another important dimension: caste. This is a partial equilibrium model, that focuses attention on how changes in the characteristics of the male earnings distribution affect the marital search duration of women.

Denote by F the cumulative distribution function corresponding to the male earnings distribution, and let f denote the associated density function. For simplicity, we assume that the woman receives exactly one offer in each period and that this offer is an independent and random draw from F. We denote the earnings associated with this offer by x. The woman may either accept or reject this proposal. If she chooses to reject it, she will have to wait one more period, during which she will receive a utility of c - this is the per-period utility from remaining single, and may be negative for some women. Recognizing the iterative nature of the problem we follow a traditional dynamic programming framework.

Denote by $V_{acc}(x)$ the lifetime value obtained by accepting offer x. Assuming that the woman is risk-neutral, we can write:

$$V_{acc}(x) = \frac{x}{1-\beta} \tag{1}$$

where β is the discount rate.

Denote by V_{rej} the value of rejecting offer x. Because of the recursive nature of the problem, we can write:

$$V_{rej} = c + \beta E[\max\{\frac{x}{1-\beta}, V_{rej}\}] \tag{2}$$

The form of the solution to this optimization problem can be shown to be the following: There is a reservation earnings level, denoted by R, above which the woman will accept any offer, and below which she will reject any offer and instead choose to wait another period. We can therefore re-write (2) as:

$$\begin{aligned} V_{rej} &= c + \beta[\int_R^\infty \frac{x}{1-\beta} dF(x) + \int_{-\infty}^R V_{rej} dF(x)] \\ &= c + \frac{\beta}{1-\beta}\int_R^\infty x dF(x) + \beta F(R) V_{rej} \end{aligned} \tag{3}$$

We also note that if the offer exactly equals the reservation level, R, the woman will be indifferent between accepting and rejecting, i.e. we must have:

$$V_{rej} = V_{acc}(R) = \frac{R}{1-\beta} \tag{4}$$

Combining (3) and (4), we can write, after some algebra:

$$[r + 1 - F(R)]R = rc + \int_R^\infty x dF(x) \tag{5}$$

where $r = \frac{1-\beta}{\beta}$.

This equation may be solved to obtain the reservation earnings level, R. Associated with R, one can define the per-period probability of "escape" from the marriage market, q:

$$q = P(x > R) = 1 - F(R) \tag{6}$$

Lower values of q imply greater search duration (in expectation), or equivalently, higher age at marriage. Formally, the expected number of periods of search before an offer is accepted can be shown to be $1/q$.

2.1 Comparative Statics

We now present the key comparative statics of interest. We note that the outcome of interest is the expected search duration, which is inversely related to q.

The comparative static results presented here are derived by assuming that F is $N(0,1)$. This allows us to invoke some standard results regarding the normal distribution that simplify the algebra. It can be verified however (see Mortensen 1986) that the results obtained here are fully general, and are not dependent on the normality assumption.

First, we consider the effect on q of changes in c, while holding fixed the earnings distribution F. Note that in this case it is sufficient to know the effect on R in order to infer the effect on q (this follows from Equation (6) above). For different levels of c, the reservation earnings will be different. It can be shown (please see Appendix for the proof) that higher levels of c correspond to higher values of R and hence lower values of q:

$$\frac{dR}{dc} = \frac{r}{r+q} > 0 \tag{7}$$

This result is intuitive, and says that women who obtain little utility (or great disutility) from remaining single are less likely to turn down offers. While this is a testable prediction, Section X explains why the nature of the data does not allow us to test it.

Next, we examine how changes in the earnings distribution, F, affect search duration. A complication in this case is that when F changes, the effect on search duration cannot automatically be inferred from the change in R.

We begin by considering the effect of a marginal rightward shift in the earnings distribution, while holding constant the variance, i.e. the mean of the distribution increases. The effects of this change are summarized below:

$$\frac{dR}{d\mu} = \frac{q}{r+q} < 1 \tag{8}$$

$$\frac{dq}{d\mu} = (1 - \frac{dR}{d\mu})\phi(R) > 0 \tag{9}$$

where ϕ denotes the standard normal density function. A right-ward shift in the offer distribution tends to increase the reservation level, R, which is intuitive. However, R increases less than one-for-one with μ, and in turn this implies that the per-period escape probability increases, and so the expected age at marriage must fall. This is a testable prediction that we are able to take to the data.

Next, we consider an increase in the standard deviation of the offer distribution, while holding constant the mean. The effects of this change are as follows:

$$\frac{dR}{d\sigma} = \frac{\phi(R)}{r+q} > 0 \tag{10}$$

$$\frac{dq}{d\sigma} = [R - \frac{dR}{d\sigma}]\phi(R) \tag{11}$$

The increase in earnings dispersion unambiguously increases the reservation level, R. However, this time, the direction of the effect on q is not unambiguous, and will depend on the term in square brackets in (11).

To understand these results, note first that for any initial value of R, the expected earnings of individuals above R increases with a mean-preserving spread of F. This implies that the value of rejecting an offer equal to R has increased, which in turn will cause the woman to revise R upwards (because initially, the value of accepting an offer equal to R was equal to the value of rejecting it). Thus the reservation level unambiguously rises for all women, regardless of their initial value of R.

Next, note that the mean-preserving spread increased the probability mass to the right of R (i.e. the escape probability, $P(x > R)$) in the case of women for whom R was above the median, while decreasing $P(x > R)$ in the case of women for whom R was below the median. For the latter group, as R rises in response to the mean-preserving spread, $P(x > R)$ falls further, whereas for initial reservation values above the median, the net effect of the increase in R on $P(x > R)$ is ambiguous, and depends on how large the increase in R is. For this reason, the mean-preserving spread has a theoretically ambiguous effect on search duration for women whose initial reservation levels were above the median.

Nevertheless, it is likely that for most values of R, the net effect on search duration is positive, i.e. q falls (see also Burdett and Ondrich 1985 for discussion on this point). Figure 1 below illustrates this in the context of our specific functional form assumptions, by graphing $R - \frac{dR}{d\sigma}$ against different values of R (where we have assumed a discount rate of $\beta = 0.95$). The figure also graphs the standard normal cumulative distribution function.

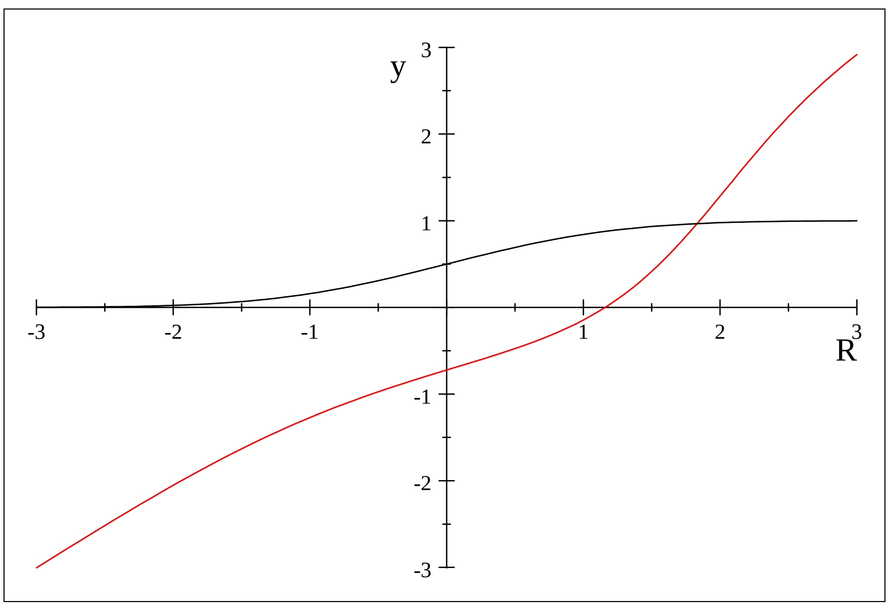


Figure 1

The figure shows that $R - \frac{dR}{d\sigma}$ is indeed negative over a substantial range of values of R, accounting for an extremely largely proportion of the probability mass. It is only in the far right tail of the distribution of R values that $R - \frac{dR}{d\sigma}$ (and hence $\frac{dq}{d\sigma}$) turn positive. Nonetheless, the question of whether increasing earnings dispersion increases search duration and the age at marriage is ultimately an empirical one, and will be tested in the data.

We have so far considered a mean-preserving spread as an example of rising inequality. However, the available evidence indicates that for a number of English-speaking countries as well as India and China, recent increases in income inequality are largely being driven by increases in income for people in the upper tails of the income distribution (Atkinson, Piketty and Saez 2011). In the empirical analysis, we will accordingly separate out the impact of increases in upper-tail incomes (relative to median incomes) from the impact of of reductions in lower-tail incomes.

While analytical proofs are difficult, it is possible to intuitively predict the direction of impact in these cases. Consider an increase in income at the 90th percentile (say), relative to the median income. For women with reservation levels below the 90th percentile (this includes the majority of women), this would increase the value of rejecting an offer equal to R. This would result in an increase in R and (most likely) an increase in search duration. However, a reduction in income at the 10th percentile (say) relative to median incomes, would have no effect on reservation levels and search durations for women whose initial reservation levels were above the 10th percentile (this likely includes the majority of women). These asymmetric predictions can be tested in the data.

3 Data

Overview of the data: We use the India Health Development Survey 2005 (IHDS) to analyze the relationship between earnings inequality and marriage in India. The IHDS is a nationally representative survey of 41,554 rural and urban Indian households in all twenty-eight states and five union territories.[1] IHDS administered household, school, village, and medical facility surveys, collecting poverty, health, employment, economic, and social data. The breadth and depth of IHDS permits extensive analysis of human and economic development in India.

[1] Lakshwadeep and the Andaman and Nicobar islands were excluded.

THe IHDS sample consists of several sub-samples; the main division is between re-interviewed households, originally part of the 1994-1995 Human Development Profile of India (HDPI) survey sample, and newly sampled households. The re-interview sampling strategy of IHDS was fairly intricate, to account for the original HDPI sampling frame and household attrition, while meeting the goal of re-interviewing 50 percent of HDPI households.[2]

To ensure a nationally representative sample, IHDS supplemented the re-interviewed households with an extension sample, surveying an additional eleven states/union territories and all urban areas, which were not included in the original HDPI population. Based on the state population, villages and towns were sampled with a probability proportional to size.[3] In rural areas, twenty households were interviewed in each village. In urban areas, to achieve the minimum target of 45 survey households per town, at least three Census-defined neighborhoods were selected at random in each town, and fifteen households were interviewed in each neighborhood.

Description of the earnings variables: The 2005 round of the IHDS elicited detailed information on the employment, occupation and earnings of individuals in the household.[4] The survey elicited the annual earnings of individuals from employment outside the home. It also elicited the net income that the household received from farm and business activity, in addition to the the contribution of each household individual to the farm or business, in terms of days worked in the year.

From these data, we construct two measures of (annual) individual earnings. The first measure includes only earnings from paid employment outside the home. The second measure also includes net income from household farm and business work. To obtain this measure, we assign

[2] Additional details of the re-interview sampling strategy are available from Desai and Vanneman (2008).

[3] Villages were selected from the 1991 Census village listing to match the HDPI sampling design; towns were selected using the 2001 Census list (Desai and Vanneman, 2008).

[4] Because the 1994 HDPI did not elicit comparable information, we could not make use of the 1994 data in this study.

household farm and business income to each household individual in proportion to the number of days that individual spent during the year on that activity (relative to the total number of days that household members devoted to the activity).

In our empirical analysis we present results separately using each of these earnings measures. On the one hand, the first measure provides an incomplete picture of the earnings distribution relative to the second measure, because a substantial proportion of individuals tend to work on their own farm or business. On the other hand, earnings from wages and salaries tend to be relatively stable and may be thought of as measures of permanent income. In contrast, because farm and business income are likely to be much more stochastic, the cross-sectional distribution of these earnings may be a somewhat noisy approximation to the cross-sectional distribution of "permanent" earnings from these activities.

Definition of marriage market: Caste and state (keeping only Hindus). We use the sample weights in the IHDS to reconstruct the earnings distribution corresponding to each marriage market. We only consider the earnings of marriageable males between ages 15 and 30, who are not currently enrolled in school. From the earnings distribution we compute (i) Mean earnings, (ii) Standard deviation of earnings, (iii) Various percentiles of the distribution, including the 90th, 50th and the 10th. We also repeat this exercise to obtain measures of mean and dispersion of earnings distributions for women.

Sample of women: Our main analysis uses two overlapping samples of women. The first sample is derived from the household roster, which provides the age and marital status of each household member. We will refer to this as the "full sample". The second sample is based on the sample of ever-married women (one was randomly chosen from each surveyed household) who were administered a separate questionnaire - for these women, we know marriage histories, in addition to their educational attainment as well as anthropometric measurements. We will refer to this as the "ever-married sample".

4 Empirical Strategy

4.1 Baseline regressions

Our baseline regressions investigate the effects of male income dispersion on female marriage using a simple specification. We begin by looking at how male earnings inequality affects the probability of being married conditional on age, using the full sample of women obtained from the household roster. We employ the following linear probability model:

$$y_{ics} = \alpha + \beta_1 \sigma_{ics} + \beta_2 \mu_{ics} + \gamma Age_{ics} + \eta_c + \eta_s + e_{ics} \tag{12}$$

where y_{ics} is an indicator that takes the value 1 if individual i belonging to caste c and residing in state s is married; σ_{ics} is a measure of the dispersion in male (log) earnings in marriage market defined by caste c and state s; μ_{ics} denotes the mean of the male (log) earnings distribution in the marriage market; Age_{ics} denotes the age of the individual; η_c denotes a caste fixed effect; η_s denotes state fixed effect and e_{ics} is an unobserved error term. The inclusion of caste fixed effects allows us to control for marital norms and socio-economic characteristics that differ across castes. The inclusion of state fixed effects controls for regional differences in socio-economic conditions that may impact on the marriage decisions of women, as well as for regional differences in marriage norms (in practice, as we explain below, we are able to control for fixed effects at a much finer geographic level).

The coefficient of interest is β_1, which represents the effect of a one unit increase in male earnings dispersion on the individual's probability of being married. The marriage market search hypothesis suggests that β_1 should be negative, i.e. greater dispersion in male earnings should result in a lower per-period probability of getting married. However, while earnings inequality may reduce the proportion of married women at any given age, it is unlikely do so at a constant rate - because most women in India are married by the age of 30, the proportions of married women in different marriage markets should eventually converge as women grow older. To estimate the rate of catch-up, we estimate the following variant of (12):

$$y_{ics} = \alpha + \beta_1 \sigma_{ics} + \beta_2 \mu_{ics} + \gamma Age_{ics} + \delta \sigma_{ics} Age_{ics} + \eta_c + \eta_s + e_{ics} \tag{13}$$

where the coefficient on the interaction between earnings dispersion and age, δ, is expected to be positive.

One concern with the analysis above is that earnings dispersion may be changing rapidly over time. As a result, the proportion of married women in a given marriage market may not really reflect the effects of the current earnings dispersion, especially if many of the married women were married several years prior to the survey. To mitigate this concern, we limit the regression sample to individuals between the ages of 15 and 30.

We implement the regressions for each of the three different measures of male earnings that we defined earlier (Section X). The regressions are weighted using the survey sampling weights, and the standard errors are corrected for clustering at the marriage market level. The actual regression specification departs from (12) in two ways. Following Loughran (2002), we measure earnings inequality using differences between earnings at different points in the earnings distribution. Specifically, we consider the difference between earnings at the 90th and 50th percentiles, and the difference between earnings at the 50th and the 10th percentile. The marital search hypothesis suggests that the marital decision should be more sensitive to inequality in the upper tail of the distribution than to inequality in the lower tail. In particular, for a woman whose reservation is at the median of the male earnings distribution, widening inequality in the lower half of the distribution should not affect her marital decision. Secondly, while the marriage market is defined at the state-caste level, the large sample sizes allow us to control for fixed effects at the level of the primary sampling unit (i.e. village level).

While the analysis based on specifications (12) and (13) can tell us how the male earnings dispersion affects the probability of being married conditional on age, it does not directly indicate the extent to which marriage is delayed as a consequence of increased earnings dispersion. This question may however be answered by turning to the sample of ever-married women for whom we know the age at marriage. For this sample, we employ the following regression specification:

$$AgeMarriage_{ics} = \alpha + \beta_1 \sigma_{ics} + \beta_2 \mu_{ics} + \eta_c + \eta_s + e_{ics} \tag{14}$$

where $AgeMarriage_{ics}$ denotes the age at which woman i belonging to caste c and residing in state s got married.

We want to use specification (14) to examine whether women in higher-inequality markets get married later than women in lower-inequality markets. Because some of the women in the married sample were married many years ago (when earnings inequality may have been very different), it would be preferable to restrict attention to those women who were married recently. We would also prefer to exclude older women, because even though some of them may be recently married, they may have been on the marriage market for a long time and current earnings inequality may not really explain their late marriages. Our solution to these issues is to include in the regression sample all married women between the ages of 15 and 25, while restricting our attention within this group to those women who got married within the last two years (restricting the sample further to women who were married in the last year reduces the sample drastically). This sample is much smaller than the full sample (which was based on the household roster and included married as well as unmarried women), and so we control for regional differences in socio-economic characteristics by including district-level fixed effects in these regressions (in place of village fixed effects).[5]

[5]It should be noted that specification X does not include as a control the current age of the married woman. This is

4.2 Testing alternative hypotheses

We now consider some alternative hypotheses that may explain the observed correlations between male earnings inequality and female marriage. First, the characteristics of the male earnings distribution may be proxying for expected female earnings. In particular, it may be the case that the dispersion in the male earnings distribution captures skill/education premia, and these returns to higher education may motivate women to pursue higher education (and thus delay their marriage). To address this possibility, we introduce as controls in the regressions the mean male earnings corresponding to each of four categories of educational attainment: (i) Fewer than 10 years of education, (ii) 10-11 years of education (i.e. secondary school completion), (iii) Higher secondary completion (i.e. at least 12 but less than 15 years of education) and (iv) College degree (15 years of education). We also examine the effect of controlling for *female* earnings conditional on these levels of educational attainment. We note, however, that the female earnings variables are potentially endogenous in these regressions. Our intention is not to interpret the coefficients on these variables, but to examine whether their inclusion affects the coefficients on the male earnings inequality variables.

Second, increasing dispersion in the male earnings distribution may extend the labor market search of men and thereby delay their own entry into the marriage market (or make them ineligible for marriage). This may reduce the number of available men in the marriage market, possibly delaying female marriage. To address this concern, we introduce as controls in the regressions (i) The overall sex-ratio in the marriage market, defined as the number of unmarried (eligible) men to the number of unmarried (eligible) women, and (ii) The ratio of working unmarried men to the number of unmarried women. Once again, these sex ratios are likely to be endogenous and their

because the restriction that the regression sample only includes women who were married in the last two years implies that current age and age at marriage are almost identical within this sample.

coefficients should therefore not be interpreted causally.

Third, it is possible that marital search duration of men is extended by greater female earnings dispersion (and that this also extends the marriage search of women in equilibrium) - if male earnings dispersion is correlated with female earnings dispersion, then this would show up as a negative correlation between female age at marriage and male earnings dispersion. Our prior is that this hypothesis is unlikely to be true, given the traditional division of responsibilities in Indian households, according to which men are the bread-winners while women are assigned to home production. However, the increasing trend towards labor market participation of women in recent years may have increased the importance of potential labor market earnings of women as a measure of their quality in the marriage market.

We examine this possibility in two ways: (1) We directly control in the regressions for measures of female earnings dispersion (with the caveat that these variables are potentially endogenous), and (2) We use regression specification (12) to look at the effect of female earnings dispersion on the probability of eligible men being married, i.e. our regression sample includes only eligible men between the ages of 15 and 30, while the measures of mean and dispersion (the explanatory variables) correspond to the female earnings distribution.

A fourth alternative hypothesis is that caste-specific norms regarding education and marriage may create a spurious correlation between marriage rates and male earnings inequality. Although our regressions controlled separately for caste and village/district fixed effects, one may still be concerned that there still remains sufficient variation in marriage norms across the states of India, even within the same caste. To address this possibility, we re-estimate our earnings dispersion measures, but this time allowing the marriage market to be defined by the intersection of caste and district (as opposed to caste and state). We then estimate specifications (12), but now controlling for state-caste interaction dummies. This specification restricts the comparisons to individuals within the same caste and state, but facing different local marriage markets.

4.3 Effect on female educational attainment

Marriage in India (and in many other developing countries) is strongly linked with education. In the majority of cases, schooling is discontinued immediately upon cohabitation (if not much earlier). We hypothesize therefore that to the extent that increased male inequality delays marriage, it may also have an impact on female educational attainment by keeping women in school longer. To test this hypothesis, we first employ specification (14), using completed educational attainment of ever-married women as the dependent variable in order to determine how many extra years of education are added by greater earnings inequality.

Next, we attempt to determine where in the educational trajectory the additional years (if any) accrue. To do this, we construct a set of indicator variables: The first takes the value 1 if the woman received zero years of education and 0 otherwise; the second takes the value 1 if the woman did not complete primary school (i.e. received fewer than 6 years of education); the third takes the value 1 if the woman did not matriculate (i.e. received fewer than 10 years of education); the fourth takes the value 1 if the woman did not complete high school (i.e. received fewer than 12 years of education); the fifth takes the value 1 if the woman did not receive any college education (i.e. received fewer than 13 years of education); the sixth takes the value 1 if the woman did not complete her college education (i.e. received fewer than 15 years of education). We then estimate linear probability models using these indicators as dependent variables, once again utilizing specification (14).

The hypothesis of marital search suggests that greater search duration due to widening earnings inequality should largely impact educational attainment at the later stages of the educational trajectory because that is when women are likely to be on the marriage market, i.e. it may affect rates of matriculation and high-school completion (and even possibly college education), but is unlikely to move women from zero to positive levels of education or to raise primary school completion rates.[6] In a sense, this provides a falsification test for our hypothesis: If greater earnings inequality is associated with a significantly differential proportion of women completing lower levels of education, then this would suggest either that marital search duration is not the channel by which earnings inequality affects education and/or that our measures of earnings inequality are correlated with unobservable socio-economic characteristics that affect education (and possibly marriage).

A related falsification test that we implement is to look at the schooling of girls who are not yet of marriageable age. If, for this group, greater earnings inequality is associated with more years of education and/or a greater propensity to be enrolled in school, then this would again suggest either that the channel of influence is not marital search duration or that earnings dispersion is correlated with unobserved socio-economic characteristics that affect education. To implement this test, we use regression specification (12), but this time looking at (i) completed years of education and (ii) current school enrolment of unmarried girls under the age of 10 (who are most likely not on the marriage market).

4.4 Are women in high-inequality markets different?

We present two tests to determine whether our earnings inequality measures are simply proxying for unobserved socio-economic characteristics. Specifically, we look at the correlation between the earnings inequality measures and two outcomes that are not decision variables, namely the height of the woman and the age at which she attained menarche (both of these variables are

[6] We do not entirely rule out the possibility, however, because there is a small percentage of women who are enrolled in school in their late teens but have only completed four or five years of education, so for these individuals an increase in search duration may result in primary school completion.

reported for the ever-married women sample). Because neither of these variables represents a short-term decision that should be affected by earnings inequality, they are suitable for testing for differences in observable characteristics between women in high- and low-inequality markets. Both height and age at menarche reflect childhood nutrition, however, so they may be correlated with earnings inequality if the latter is correlated with unobserved socio-economic characteristics.

We employ the regression specification in (14) to test whether our measures of earnings dispersion have any predictive power for these outcomes. The test that looks at age at menarche is of particular interest because women typically get married only after they attain menarche (indeed, in our sample of ever-married women, age at menarche is a very strong predictor of age at marriage). If individuals in marriage markets with greater earnings inequality also happen to be associated with poor childhood nutrition (which would act to delay menarche), this might result in a spurious positive correlation between earnings inequality and age at marriage.

4.5 Robustness checks

Lastly, we examine the robustness of the results to varying the regression samples. In the regressions corresponding to specifications (12) and (13), we had limited the sample to individuals between the ages of 15 and 30, arguing that including older individuals may be problematic because many of the married individuals may have gotten married many years ago. We will relax this restriction to include all individuals up the age of 35, and test whether the results are robust to this relaxation. In the regressions using the ever-married sample, we had argued that it may be preferable to avoid including older married women. We will examine the robustness of the results to expanding the regression sample to include the set of ever-married women under the age of 30.

5 Results

Table 1 reports the results of the baseline regressions that relate the probability of being married to earnings inequality in the marriage market. We report the results corresponding to each of the three measures of male earnings described in Section X. The effect of upper-tail earnings inequality on the probability of being married (captured by the coefficient on the Male 90-50 earnings difference) is stable across the different earnings measures. Because the explanatory variable is the difference between the 90th and 50th percentiles of the log earnings distribution, the coefficient on this variable may be interpreted as saying that if earnings at the 90th percentile were to double relative to median earnings, then this would reduce the proportion of married women (for the average age-group) by 6-7%. However, the effect of greater lower-tail inequality on the probability of being married is small and insignificant, suggesting that the average woman's reservation earnings level is likely to be near the median of the male earnings distribution. Table X also presents the estimates from the regression model in X, which includes an interaction between upper-tail earnings inequality and the age of the individual. As expected, the effect of earnings inequality on the probability of being married is still negative, but this effect declines as women grow older, by around 1 percentage point for each year.

Table X reports the estimates for the regressions corresponding to specification X, which relates age at marriage of ever-married women to male earnings inequality. Once again, the results are quite comparable across the three different earnings measures, and indicate that a doubling of 90th percentile earnings relative to the median increases the average age at marriage by 1.5-1.8 years. As before, lower tail inequality does not have a statistically significant effect on age at marriage.

Because the regression estimates display only minor variations across the three different earnings measures, we hereonwards restrict attention to the estimates derived using the third and most comprehensive earnings measure (the full set of results for each of the other two measures is however available on request).

We now introduce various controls to test alternative hypotheses that may explain the correlation between male earnings inequality and female marriage. Tables X and X report the results separately for the full sample and the ever-married sample. To determine whether male earnings dispersion is proxying for educational premia, we examine the effect of controlling for mean male earnings conditional on each of four levels of educational attainment (Column 1 in Tables X and X). The results are qualitatively the same as before, although the estimated effects of earnings inequality on marriage rates and the age of marriage are now uniformly larger than before in all specifications. Next, to determine whether greater male earnings dispersion could be correlated with the availability of eligible men, we control for the overall sex-ratio in the marriage market, defined as the number of unmarried (eligible) men to the number of unmarried (eligible) women, as well as the ratio of working unmarried men to the number of unmarried women (Column 2 in Tables X and X). The inclusion of these controls does not have much of an impact on the coefficients on the earnings inequality variables. We then examine whether male earnings dispersion could be standing in for female earnings dispersion, by introducing measures of mean and dispersion corresponding to the female earnings distribution (Column 3 in Tables X and X). Once again, the coefficients on the male earnings inequality variables are largely unaffected by the inclusion of these variables. In addition, the coefficients on the female earnings inequality variables are small and insignificant (but we repeat the caveat that these variables are potentially endogenous).

Table X contains two additional columns. Column 4 reports the results of a regression in which we examine the effect of earnings inequality among eligible women on marriage rates of men. Clearly, there is no symmetric effect of female earnings inequality on male marriage propensities, indicating that our results were not being driven by extended marriage-market search on the part of men as a response to greater female earnings inequality. Column 5 reports estimates of the effect of male earnings inequality on female marriage, but in this specification the marriage-market is defined at the intersection of district and caste (and the earnings distribution accordingly calculated), and we control for state-caste fixed effects. This allows us a finer comparison between individuals belonging to the same caste and residing in the same state, but facing different local marriage markets. This is a particularly demanding specification, not only because it requires sufficient variation in earnings distributions between local marriage markets within the same caste and state, but also because the newly-defined marriage-market cells contain many fewer observations and the corresponding earnings distributions are likely to be noisily measured. The results in Column 5, however, once again confirm that upper-tail earnings inequality reduces marriage rates even under this specification, although the estimated effect is smaller than before (at 2%).

We now consider the effect of male earnings inequality on female educational attainment, using the ever-married sample of women. In these regressions, we control for the mean male earnings conditional on each of the four levels of educational attainment. The returns to these various levels of education are an important consideration in educational decisions, and the results in Tables X and X suggest that their exclusion may bias the estimates of the effect of earnings inequality.

Table X reports the regression results for the educational attainment variables. The dependent variables in Columns 1-6 are, respectively, (i) Completed years of education, (ii) Probability of having obtained zero years of education, (iii) Probability of having obtained fewer than six years of education, (iv) Probability of receiving fewer than 10 years of education, (v) Probability of having received fewer than 12 years of education and (vi) Probability of having received fewer than 15 years of education.

A doubling of 90th percentile earnings relative to the median increased completed years of education by 0.6-2 years, but these effects are not precisely estimated. The coefficient sizes are however comparable to those in Table X, suggesting the possibility that increases in age at marriage result one-for-one in increases in years of education. Turning to the specific attainment variables, we find that earnings inequality has small and statistically insignificant effects on lower-level educational attainment, but it has a significant effect on matriculation and high-school completion rates. In particular, matriculation probability increases by 30% and high-school completion probability increases by 43%. There is some weak evidence that it also increases rates of college completion, but the effects are not statistically significant. These results are consistent with the hypothesis that greater earnings inequality affects female marriage by increasing search duration, and therefore impacts educational attainment at the part of the educational trajectory that women find themselves in when they are on the marriage market.

Table X examines the effect of earnings inequality on educational attainment and current school enrolment of unmarried children below the age of 10. Because these children are plausibly not on the marriage market, earnings inequality should not affect these educational outcomes. The results in Table X confirm this hypothesis: Neither current school enrolment nor completed years of education are correlated with male earnings inequality for this sample.

Table X presents the results of the tests that examine whether male earnings inequality predicts age at menarche and height of women in the ever-married sample. The results are negative: The coefficients on the earnings inequality variables are small and statistically insignificant, indicating that these variables are unlikely to be capturing the effects of unobserved socio-economic characteristics. The results confirm that women in high inequality markets are not different from women in low inequality markets in terms of observable characteristics.

Finally, Tables X and X examine the robustness of the main results to alternative regression samples, as described in the previous section. Table X presents the results using specification X, but this time including in the regression sample all individuals up to the age of 35. The estimated coefficients are quite similar to those obtained using the more limited sample. Table X repeats the full set of regressions that used the ever-married sample, but this time including all ever-married women in the regression sample. Because the sample sizes are now significantly larger in these regressions, we are able to control for fixed effects at the district level. Once again, the results are reasonably robust to this expansion of the sample.

References

Banerjee, A., & Piketty, T. (2005). Top Indian Incomes. *World Bank Economic Review* 19 (1), 1-20.

Block, S. (2004). Maternal Nutrition Knowledge and the Demand for Micronutrient-Rich Foods: Evidence from Indonesia. *Journal of Development Studies*, 40(6), 82-105.

Bloom, D. E., & Reddy, P. (1986). Age Patterns of Women at Marriage, Cohabitation, and First Birth in India. *Demography*, 23(4), 509-523.

Brien, M. J., & Lillard, L. A. (1994). Education, Marriage, and First Conception in Malaysia. *Journal of Human Resources*, 29(4), 1167-1204.

Bruce, J. (2003). Married Adolescent Girls: Human Rights, Health, and Developmental Needs of a Neglected Majority. *Economic and Political Weekly*, 38(4), 378-80.

Boulier, B.L. & Rosenzweig, M.R. (1984). Schooling, search, and spouse selection: Testing economic theories of marriage and household behavior. *Journal of Political Economy*, 92(4), 712–732.

Caldwell, J., Reddy, P., & Caldwell, P. (1983). Causes of Marriage Change in South India. *Population Studies*, 37(3), 343-361.

Choe, M.K., Westley, S.B., & Rutherford, R.D. (2002). Tradition and Change in Marriage and Family Life: The Future of Population in Asia. Honolulu, HI:East-West Center.

Christiaensen, L., & Alderman, L. (2004). Child Malnutrition in Ethiopia: Can Maternal Knowledge Augment the Role of Income? *Economic Development and Cultural Change*, 52(2), 287-312.

Clark, S. (2004). Early marriage and HIV risks in sub-Saharan Africa. *Studies in Family Planning*, 35, 149- 60.

Deaton, Angus, and Jean Dreze (2002). Poverty and inequality in India: A re-examination. *Economic and Political Weekly*, 7 September: 3729-3748.

Field, E., & Ambrus, A. (2008). Early Marriage, Age of Menarche, and Female Schooling Attainment in Bangladesh. *Journal of Political Economy*, 116(5), 881-929.

Gakidou, E., Cowling, K., Lozano, R., & Murray, C. (2010). Increased Educational Attainment and Its Effects on Child Mortality in 175 Countries between 1970 and 1990: A Systematic Analysis. *Lancet* (376), 959-74.

Glewwe, P. (1999). Why Does Mother's Schooling Raise Child Health in Developing Countries? Evidence from Morocco. *Journal of Human Resources*, 34(1), 124-59.

Harwood-Lejeune, A. (2000). Rising Age at Marriage and Fertility in Southern and Eastern Africa. *Journal of Population* 17, 261-280.

Heaton, T.B., Forste, R., & Otterstrom, S.M. (2002). Family transitions in Latin America: First intercourse, first union, and first birth. *International Journal of Population Geography*, 8, 1–15.

Jensen, R., & Thornton, R. (2003). Early Female Marriage in the Developing World. *Gender and Development*, 11(2), 9-19.

Lester, A., Banerji, M., & Desai, S. (2008). Negotiating Marriage: Examining the Gap between Marriage and Cohabitation in India. Working Paper IHDS.

Lindstrom, D.P., & Brambila Paz, C. (2001). Alternative theories of the relationship of schooling and work to family formation: Evidence from Mexico. *Social Biology*, 48(3–4), 278–297.

Loughran, D. (2002). The Effect of Rising Male Wage Inequality on Female Age at First Marriage. *Review of Economics and Statistics*, 84(2).

Loughran, D., & Zissimopoulo, J. (2009). Why Wait? The Effect of Marriage and Childbearing on the Wages of Men and Women. *Journal of Human Resources*, 44(2).

Maitra, P. (2004). Parental Bargaining, Health Inputs and Child Mortality in India. Working Paper Department of Economics Monash University.

Mensch, B. S., Singh, S. & Casterline, J. B. (2006). Trends in the Timing of First Marriage Among Men and Women in the Developing World. In *The Changing Transitions to Adulthood in Developing Countries: Selected Studies*. Washington, DC: The National Academies Press, 2006.

Miller, J. E. (1993). Birth Outcomes by Mother's Age At First Birth in the Philippines. International Family Planning Perspectives, 19(3), 98-102.

Nour, NM. (2006). Health Consequences of Child Marriage in Africa. *Emerging Infectious Diseases*, 12, 1644 –9.

Raj, A., Saggurti, N., Balaiah, D., & Silverman, J. (2009). Prevalence of Child Marriage and Its Effect on Fertility and Fertility-control Outcomes of Young Women in India: a cross-sectional observational study. *The Lancet* (373), 1883–89.

Rashad, H. & Osman, M. (2003). Nuptiality in Arab Countries: Changes and Implications. In N. S. Hopkins (ed.), *Cairo Papers in Social Science: The New Arab Family,* Cairo: The American University in Cairo Press.

Samuel, M., & Meyer, R. (2005). Age at first marriage in Malawi: a Bayesian multilevel analysis using a discrete time-to-event model. *Journal of the Royal Statistical Society*, 168(2), 439-455.

Santhya, K., Ram, U., Acharya, R., Jejeebhoy, S., Ram, F., & Singh, A. (2010). Associations Between Early Marriage and Young Women's Marital and Reproductive Health Outcomes: Evidence from India. *International Perspectives on Sexual and Reproductive Health*, 36(3), 132-139.

Sautman, A. (2009). Partner Search and Demographics: The Marriage Squeeze in India. Working Paper, Brown University.

Sen, Abhijit, and Himanshu (2005). Poverty and inequality in India: Getting closer to the truth.

Reprinted in Angus Deaton and Valerie Kozel (eds). *Data and Dogma: The Great Indian Poverty Debate*. Macmillan, New Delhi: 306-370.

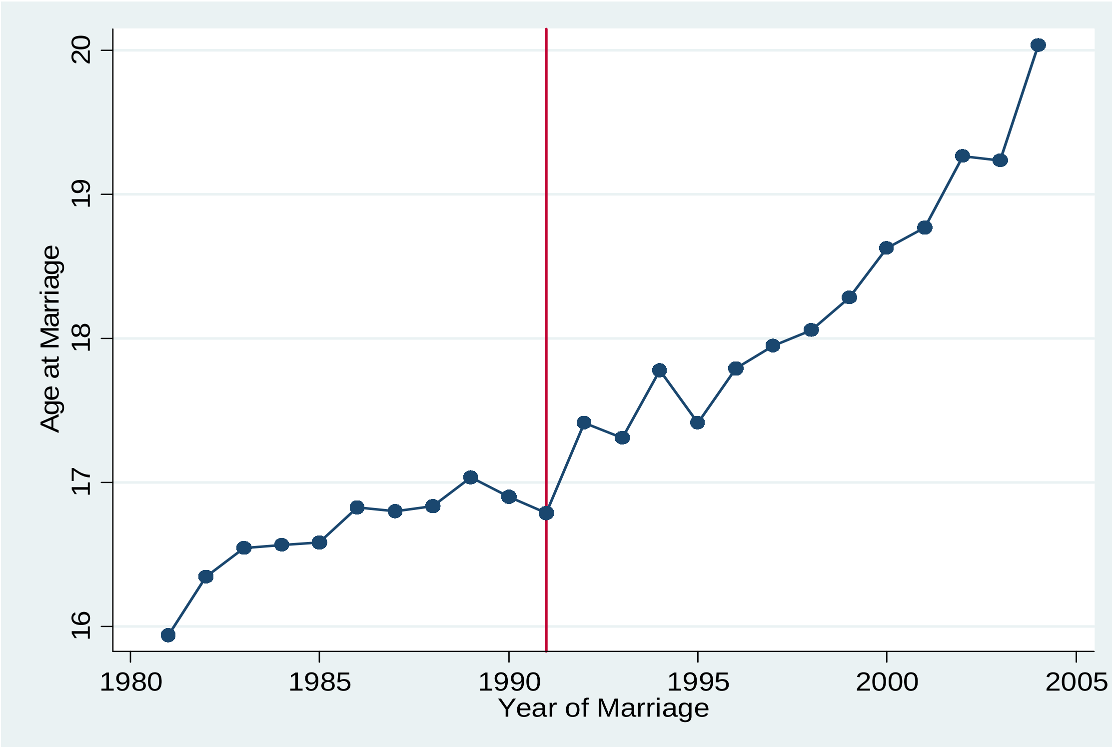
Singh, S., & Samara, R. (1996). Early Marriage Among Women in Developing Countries. *International Family Planning Perspectives*, 22(4), 148-157.

Sundaram, K., and Suresh D. Tendulkar (2003). Poverty in India in the 1990s: Revised results for all-India and 15 major states for 1993-94. *Economic and Political Weekly*, 15-22 November: 4865-487

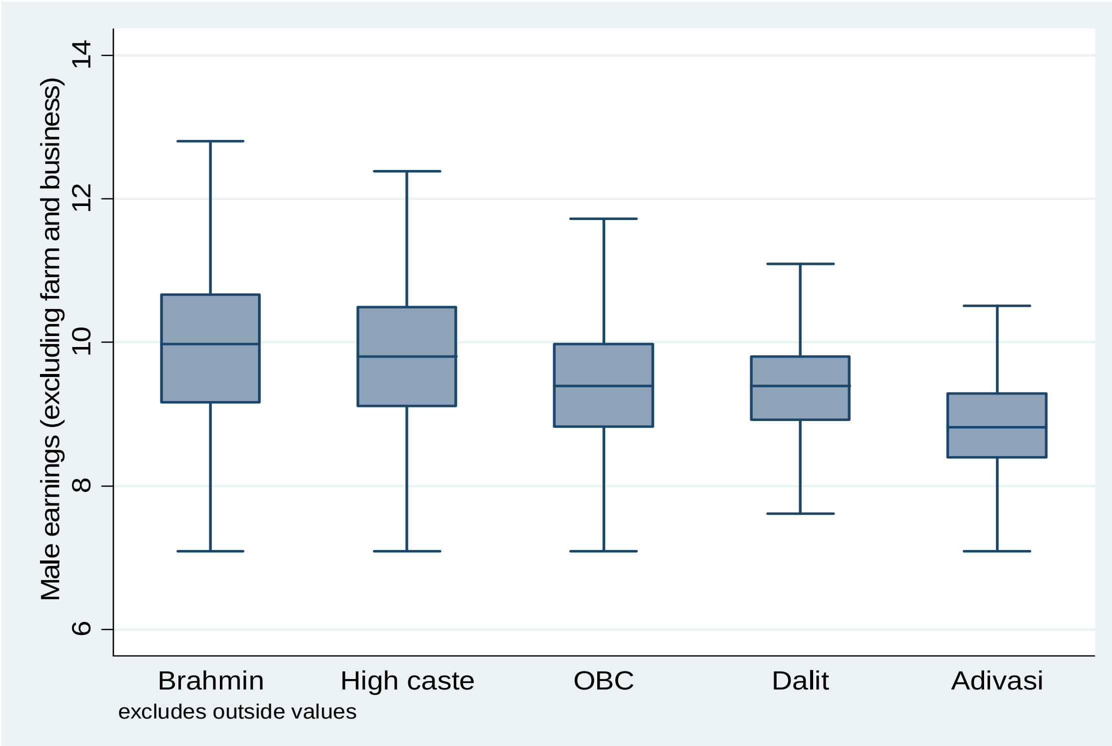
Wang, F., & Yang, Q. (1996). Age at Marriage and First Birth Interval: The Emerging Change in Sexual Behavior among Young Couples in China. *Population and Development Review*, 22(2), 299-320.

Westoff, C. F. (2003). Trends in Marriage and Early Childbearing in Developing Countries. DHS Comparative Reports 5.

Wong, O. M. (2005). The Socioeconomic Determinants of the Age of First Marriage among Women in Hong Kong. *Journal of Family and Economic Issues*, 26(4), 529-550.



Age at Marriage
16
17
18
19
20
1980
1985
1990
1995
2000
2005
Year of Marriage


Male earnings (excluding farm and business)
14
12
10
8
6
Brahmin
High caste
OBC
Dalit
Adivasi
excludes outside values

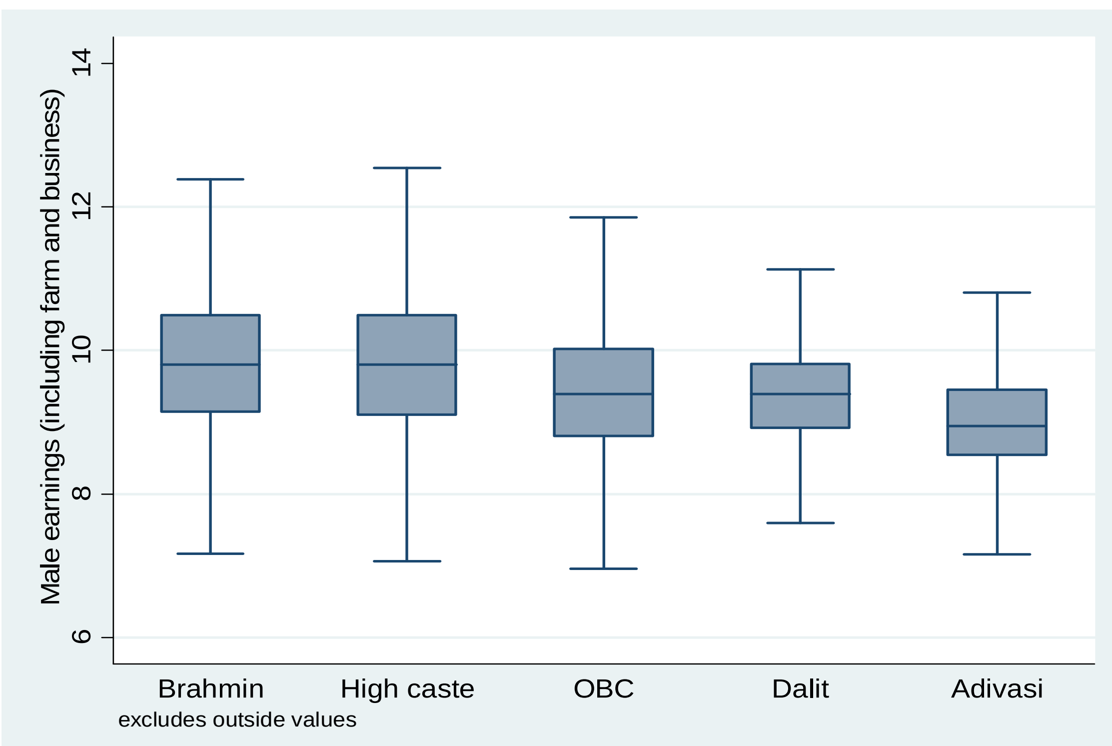

Male earnings (including farm and business)
14
12
10
8
6
Brahmin
High caste
OBC
Dalit
Adivasi
excludes outside values

Table : Effect of male earnings inequality on marriage propensity

	Earnings measure 1		Earnings measure 2		Earnings measure 3	
	(1)	(2)	(3)	(4)	(5)	(6)
Male 90-50	-0.06**	-0.31***	-0.07***	-0.32***	-0.07***	-0.34***
	(0.02)	(0.10)	(0.02)	(0.11)	(0.02)	(0.11)
Male 50-10	-0.02	-0.02	0.01	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Male μ	-0.00	-0.00	0.02	0.02	0.02	0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Age	0.07***	0.06***	0.07***	0.06***	0.07***	0.06***
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)
Age x Male 90-50		0.01**		0.01**		0.01**
		(0.00)		(0.00)		(0.00)
Constant	-0.85***	-0.58**	-1.10***	-0.83***	-1.11***	-0.82***
	(0.22)	(0.24)	(0.25)	(0.26)	(0.25)	(0.25)
Observations	21,314	21,314	21,336	21,336	21,336	21,336
R-squared	0.511	0.512	0.512	0.512	0.512	0.512

*** p<0.01, ** p<0.05, * p<0.1; Standard errors clustered at marriage-market (i.e. state-caste) level;

Table : Effect of male earnings inequality on age at marriage

	Earnings measure 1	Earnings measure 2	Earnings measure 3
	(1)	(2)	(3)
Male 90-50	1.85**	1.75***	1.53**
	(0.73)	(0.66)	(0.61)
Male 50-10	-0.13	-0.57	-0.19
	(0.34)	(0.46)	(0.50)
Male μ	0.48	-0.44	-0.45
	(0.68)	(0.69)	(0.74)
Constant	14.78**	24.83***	24.79***
	(7.20)	(7.02)	(7.57)
Observations	432	433	433
R-squared	0.178	0.175	0.173

*** p<0.01, ** p<0.05, * p<0.1; Standard errors clustered at marriage-market (i.e. state-caste) level;

Table : Testing alternative hypotheses (full sample)

	(1)	(2)	(3)	(4)	(5)
Male 90-50	-0.09***	-0.05**	-0.08***		-0.02**
	(0.02)	(0.02)	(0.03)		(0.01)
Male 50-10	0.02	0.00	-0.00		-0.01
	(0.01)	(0.01)	(0.02)		(0.01)
Male μ	0.03	0.04*	-0.01		-0.07***
	(0.08)	(0.02)	(0.05)		(0.01)
Mean male earnings (<10 yrs of education)	0.05				
	(0.05)				
Mean male earnings (>9 yrs & <12 yrs of education)	-0.02				
	(0.02)				
Mean male earnings (>11 yrs & < 15 yrs of education)	-0.01				
	(0.02)				
Mean male earnings (15 yrs of education)	-0.01				
	(0.01)				
Sex ratio		0.06***			
		(0.01)			
Female μ			0.03	0.00	
			(0.05)	(0.03)	
Female 90-50			0.02	-0.02	
			(0.02)	(0.03)	
Female 50-10			0.01	-0.00	
			(0.02)	(0.02)	
Constant	-1.30***	-1.39***	-1.08***	-1.24***	-0.19*
	(0.29)	(0.24)	(0.27)	(0.34)	(0.11)
Observations	20,760	21,323	21,306	21,510	21,059
R-squared	0.511	0.512	0.511	0.495	0.378

*** p<0.01, ** p<0.05, * p<0.1; Standard errors clustered at marriage-market (i.e. state-caste) level;

Table : Testing alternative hypotheses (ever-married sample)

	(1)	(2)	(3)
Male 90-50	1.75*	1.59***	1.87*
	(0.93)	(0.60)	(1.05)
Male 50-10	-0.16	-0.25	-0.01
	(0.57)	(0.49)	(0.56)
Male μ	2.34	-0.42	0.21
	(1.64)	(0.73)	(1.29)
Mean male earnings (<10 yrs of education)	-1.49		
	(1.16)		
Mean male earnings (>9 yrs & <12 yrs of education)	-0.79		
	(0.63)		
Mean male earnings (>11 yrs & < 15 yrs of education)	-0.31		
	(0.47)		
Mean male earnings (15 yrs of education)	-0.42		
	(0.30)		
Sex ratio		0.44	
		(0.40)	
Female μ			-0.47
			(1.06)
Female 90-50			-0.52
			(0.92)
Female 50-10			-0.10
			(0.46)
Constant	26.03***	23.78***	22.91***
	(9.34)	(7.55)	(7.76)
Observations	417	431	432
R-squared	0.158	0.174	0.174

*** p<0.01, ** p<0.05, * p<0.1; Standard errors clustered at marriage-market (i.e. state-caste) level;

Table: Effect of male earnings inequality on female educational attainment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Yrs of education	No education	Less than primary	Less than 10 yrs	No high school	Did not complete high school	No college education	Did not complete college
Male 90-50	2.00	0.04	-0.09	-0.30**	-0.43***	-0.43***	-0.16	-0.14
	(1.39)	(0.12)	(0.12)	(0.13)	(0.12)	(0.12)	(0.11)	(0.10)
Male 50-10	-0.52	0.09	0.10	-0.04	0.02	0.02	-0.09	-0.05
	(0.80)	(0.07)	(0.07)	(0.08)	(0.09)	(0.09)	(0.07)	(0.06)
Male μ	-2.18	0.12	0.06	0.40	0.31	0.25	-0.02	0.05
	(3.59)	(0.26)	(0.27)	(0.34)	(0.41)	(0.41)	(0.28)	(0.28)
Constant	-5.15	0.96	0.55	0.76	2.39	2.96*	3.95**	2.66**
	(14.83)	(1.36)	(1.44)	(1.39)	(1.61)	(1.66)	(1.52)	(1.28)
Observations	417	417	417	417	417	417	417	417
R-squared	0.230	0.199	0.183	0.239	0.189	0.200	0.246	0.224

Table : Effect of male earnings inequality on education of children not on the marriage market

	(1)	(2)
	Yrs of education	Currently enrolled
Male 90-50	-0.07	-0.08
	(0.11)	(0.08)
Male 50-10	0.05	-0.04
	(0.06)	(0.04)
Male μ	0.30	0.26
	(0.32)	(0.18)
Constant	0.76	2.21**
	(1.30)	(1.10)
Observations	15,663	15,663
R-squared	0.204	0.215

*** p<0.01, ** p<0.05, * p<0.1; Standard errors clustered at marriage-market (i.e. state-caste) level;

Table : Effect of male earnings inequality on height and age at menarche

	(1)	(2)
	Yrs of education	Currently enrolled
Male 90-50	-1.57	0.01
	(4.89)	(0.22)
Male 50-10	-3.04	-0.21
	(2.37)	(0.24)
Male μ	-2.82	0.30
	(7.00)	(0.45)
Constant	183.11**	11.48**
	(73.74)	(4.59)
Observations	433	433
R-squared	0.098	0.221

*** p<0.01, ** p<0.05, * p<0.1; Standard errors clustered at marriage-market (i.e. state-caste) level;

Table : Robustness checks (full sample)

	Earnings measure 1	
	(1)	(2)
Male 90-50	-0.05***	-0.32***
	(0.02)	(0.08)
Male 50-10	0.01	0.01
	(0.01)	(0.01)
Male μ	0.02	0.02
	(0.02)	(0.02)
Age		0.03***
		(0.01)
Age x Male 90-50		0.03***
		(0.00)
Constant	-0.56***	-0.26
	(0.18)	(0.20)
Observations	29,519	29,519
R-squared	0.469	0.470

*** p<0.01, ** p<0.05, * p<0.1; Standard errors clustered at marriage-market (i.e. state-caste) level;

Table : Robustness checks (ever-married sample)

	Age at marriage	Age at menarche	Height (in cm)	Yrs of education	No education	<6 yrs of education	<10 yrs of education	<12 yrs of education	<15 yrs of education
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Male 90-50	1.36**	-0.21	-3.97	2.31	-0.08	-0.11	-0.32**	-0.36***	-0.10
	(0.57)	(0.16)	(4.85)	(1.42)	(0.13)	(0.14)	(0.13)	(0.13)	(0.12)
Male 50-10	0.45	-0.22	1.02	-0.45	0.08	0.15*	-0.04	-0.00	-0.04
	(0.60)	(0.23)	(5.88)	(0.89)	(0.08)	(0.09)	(0.09)	(0.09)	(0.06)
Male μ	-0.41	0.48	4.50	-2.20	0.44	0.23	0.01	-0.09	-0.20
	(0.91)	(0.41)	(11.31)	(3.98)	(0.39)	(0.42)	(0.30)	(0.34)	(0.21)
Constant	23.39**	10.17**	105.16	15.27	-1.18	-2.16	1.07	2.66	1.40
	(9.62)	(3.97)	(116.36)	(14.07)	(1.24)	(1.49)	(1.28)	(1.69)	(1.35)
Observations	718	718	718	695	695	695	695	695	695
R-squared	0.587	0.540	0.482	0.633	0.622	0.584	0.575	0.523	0.535