Appendix for

Bride Price and Female Education

Nava Ashraf

Harvard University, NBER, BREAD

Natalie Bau

University of Toronto

Nathan Nunn

Harvard University, NBER, BREAD

Alessandra Voena

University of Chicago, NBER

1 July 2016

Appendix A: Theoretical appendix

A. *Bride price and male education*

We discuss an extension of our model that examines how male schooling responds (or does not respond) to the bride price custom. Compared to the case for females, the impact of bride price on male education is more complex and, as we will see, more assumptions are needed to derive unambiguous predictions.

While it is always the case that the bride's parents receive the bride price, who makes the payment is less standard, and varies across families, often depending on the specifics of the situation. While the husband generally is the one paying the bride price, the family, or even extended family and friends, often also contribute.

It turns out that the problem becomes substantially more complicated if one allows the parents to pay the bride price (in full or in part), particularly if educating a son means he will be more likely to marry an educated woman. If the son is the one choosing a spouse, parents may prefer to underinvest in their son's education if they do not value an educated bride enough to justify paying the bride price education premium. If the parents choose the bride, they may choose a lower quality (and cheaper) bride but still invest in their son's schooling, leaving education rates unchanged. In what follows, we will consider the case where the groom (only) pays the bride price.

Since the parents' problems are linear for boys and girls, we consider their problems separately without loss of generality. The problem for a boy's parents is the following:

$$\max_{P\in\{0,1\},c\geq 0} \quad c_1 + \frac{c_2}{1+r} + \delta\left[a_j P_j + \frac{u_2(P_j, I_e, k)}{1+r}\right] \tag{1}$$

$$s.t. \quad c_1 + k \cdot P_j \leq y_1$$

$$c_2 \leq y_2$$

with δ capturing the parents' altruism towards a son, j indexing a son, P_j taking the value 1 if the son j is educated and 0 otherwise, and u_2 denoting the utility a son obtains in the second period.

Following a similar reasoning that we used for daughters, we obtain the probability that a son is educated as:

$$Pr(P_j = 1|k, I_e) = 1 - G\left(\frac{1}{\delta}k - \frac{\Delta u(I_e, k)}{1+r}\right).$$

Hence, parents in bride price groups may decide differently about a son's education relative to parents in other groups if and only if the returns to education $\Delta u_{I_e}(k)$ differ systematically across ethnic groups. Unlike for daughters, there is no direct effect on the choice coming from the bride price education premium. In fact, data from both Indonesia and Zambia show that the levels of male education across ethnic groups do not vary as systematically as the levels of female education do. In the Indonesia 1995 Intercensal Survey, male primary school completion is 2.3 percentage points higher in bride price ethnicities. However, this difference is not statistically significant after controlling for district fixed effects. Moreover, in the IFLS, we do not find any statistically significant difference in test scores for males in bride price groups compared to other ethnic groups. This is consistent with the fact that there is no systematic difference in the marginal boy that can be induced into schooling by a school construction program. In the pooled Zambia DHS, male primary school enrollment is 1.6 percentage points lower in bride price ethnicities and this difference is not statistically significant whether or not we control for district fixed effects.

In terms of the effects of school construction varying across ethnic groups, any difference in the responses across ethnic groups would then be driven by differences between $\frac{\partial \Delta u_{BP}(k)}{\partial k}$ and $\frac{\partial \Delta u_{NoBP}(k)}{\partial k}$, since education levels do not vary systematically across ethnic groups. As shown in appendix table A14, the impact of school construction on male schooling also does not vary in a systematic way between ethnic groups that do or do not traditionally practice bride price.

B. *An equilibrium model of investment in education with and without bride price*

In the paper, we examine comparative statics that take the daughter's value v^f as an exogenous function of ethnicity and the cost of schooling. In this appendix, we allow v^f, and the corresponding u^m for men, to be equilibrium objects which are determined on the marriage market. The goal is to examine whether assumptions **A3-A5** hold when we take into account the equilibrium effects of school construction in a simple case in which spouses only differ by education and matching is frictionless with transferable utility, as in the seminal paper by Chiappori, Iyigun and Weiss (2009), once we have incorporated the bride price custom.

Define ζ_i^f and ζ_j^m to be the agents' values if they remain single, i.e. their labor market earnings, where i indexes a female and j indexes a male, and let ζ_{ij} be the total value of a marriage between i and j. Define V_i to be the material output that women receive in marriage, and U_j to be the one that men receive. Schooling abilities, as in the main model, are a_i and a_j, distributed as

$G(a)$. In line with Chiappori et al. (2009), we introduce preferences for marriage, θ_i and θ_j, which are distributed as $F(\theta)$ on support $[\underline{\theta},\bar{\theta}]$. F and G are independent and identical across genders. Hence, we have that

$$v^f(S_i, I_e, a_i, \theta_i) = a_i S_i + \frac{\zeta_i^f + \max\{V_i + \theta_i, 0\}}{1+r}$$
$$u^m(P_j, I_e, a_j, \theta_j) = a_j P_j + \frac{\zeta_j^m + \max\{U_j + \theta_j, 0\}}{1+r},$$

where, as before, P_j is the education investment decision for a male, and S_i is the education investment decision for a female. Hence, we have decomposed the utility of daughters and sons in the second period as utility they derive from the labor market (ζ_i and ζ_j), and utility they derive in marriage, if they get married (max$\{V_i + \theta_i, 0\}$ and max$\{U_j + \theta_j, 0\}$).

The marriage market. We assume there is no intermarriage between ethnic groups.[1] Each ethnic group features identical masses of women i and men j. Marriage output is defined as

$$z_{ij} = \zeta_{ij} - \zeta_i - \zeta_j.$$

Total marriage surplus is then defined as

$$s_{ij} = z_{ij} + \theta_i + \theta_j - I_e \cdot [b + \pi S].$$

Agents' value when single and marital output only depend on their education:

$$\zeta_{ij} = \zeta_{S_i P_j}, \qquad \zeta_i = \zeta_{S_i}, \qquad \zeta_j = \zeta_{P_j}, \qquad z_{ij} = z_{S_i P_j}.$$

Education leads to a labor market return to schooling (R), which varies by gender:

$$R^f \equiv \zeta_1^f - \zeta_0^f \quad \text{and} \quad R^m \equiv \zeta_1^m - \zeta_0^m.$$

Because different ethnic groups often live in the same area and are likely to face the same labor markets, we do not allow these returns to vary across ethnic groups. This model focuses on the role of bride price in determining the portion of the returns to education $\Delta v_e = R^f + \Delta V_e$ and $\Delta u_{I_e} = R^m + \Delta U_e$ that accrues to men and women in the marriage, ΔV_e and ΔU_e.

We assume that the surplus gains from a woman's education exceed the bride price cost and that spouses' education levels are complementary:

[1] In the Indonesia 1995 Intercensal Survey, 1.50% of married household heads aged 25-45 are in a marriage in which the bride price practice of the husband and wife differ. That proportion is 16.80% in the pooled Zambia DHS.

$$z_{10} - z_{00} > \pi, \qquad z_{11} - z_{01} > \pi, \qquad z_{00} + z_{11} > z_{10} + z_{01}.$$

Last, we assume the output from the marriage of uneducated people is high enough that even a couple in which each spouse has the lowest value of marriage produces positive surplus:

$$z_{00} - I_e \cdot b + 2\underline{\theta} > 0.$$

A stable equilibrium maximizes aggregate surplus (Shapley and Shubik, 1971, Becker, 1973). Therefore, note that because of the condition on the support of θ, maximizing aggregate surplus requires that everyone marries in equilibrium.

Consistent with the data, we consider the case in which more men than women are educated. Later, we will examine which assumptions generate this outcome in equilibrium. Chiappori et al. (2009) show that the unique stable equilibrium in this marriage market is one in which everyone marries and educated women only marry educated men. Moreover, men of the same education all obtain the same share of marital output, and the same is true for women.

This implies that there exist shares of marital output U_S, V_P such that:

$$V_0 + U_0 = z_{00} - I_e \cdot b, \qquad V_1 + U_1 = z_{11} - I_e \cdot [b + \pi],$$
$$V_0 + U_1 = z_{01} - I_e \cdot b.$$

Subtracting these conditions, we have the following expressions for the returns to education in the marriage market:

$$(V_0 + U_1) - (V_0 + U_0) = \Delta U_e = z_{01} - z_{00} \tag{2}$$
$$(V_0 + U_1) - (V_1 + U_1) = \Delta V_e = z_{11} - z_{01} - I_e \pi.$$

As in Chiappori et al. (2009), educated women, who are the side in *short* supply, receive their marginal contribution in marriage with an educated man. Educated men, the side in *excess* supply, receive their marginal contribution to a marriage with an uneducated woman. More importantly in our context, the bride price erodes the contribution of a woman's education to total output, and hence will reduce her marriage market return to education.

The investment stage. The education choice problems for females and males are stated above. We have imposed that more men than women are educated. Exploiting the fact that the distribution of ability $G(a)$ is the same for men and women, we have that more men than women are educated in equilibrium if and only if

$$\left(\frac{1}{\gamma} - \frac{1}{\delta}\right) k + \frac{R^m - R^f}{1+r} + \frac{\Delta U_e - \Delta V_e}{1+r} - I_e \frac{\pi}{\gamma(1+r)} > 0. \quad (3)$$

As long as condition 3 is satisfied before and after the school construction, the equilibrium described above holds in both cases.[2] Ignoring the changes in labor market returns caused by the school construction and assuming that the school construction does not affect the technology of home production (z_{ij}), we can also easily verify that assumptions **A3**-**A5** are all met. In particular,

$$\Delta v_{BP}(k) - \Delta v_{NoBP}(k) = \Delta V(I_e = 1) - \Delta V(I_e = 0) = \pi < \frac{\pi}{\gamma},$$

and so **A3** is satisfied. Moreover,

$$\frac{\partial \Delta v_{BP}(k)}{\partial k} = \frac{\partial \Delta v_{NoBP}(k)}{\partial k} = 0 < \frac{1+r}{\gamma},$$

and so **A4** and **A5** are also satisfied.

In addition, this model implies that the bride price is incident on the wife, and hence that the groom's education is not affected by this custom. This is consistent with our empirical findings. Of course, the fact that this is a model with frictionless matching with transferable utility is crucial for these results to hold.

C. *Proofs of the Paper's Predictions*

Proof of prediction 1

Proof (*i*) Compare $1 - G(a^*_{BP}(k))$ and $1 - G(a^*_{NoBP}(k))$. We have that the threshold equals

$$a^*_{BP}(k) = \left(\frac{1}{\gamma}\left[k - \frac{\pi}{1+r}\right] - \frac{\Delta v_{BP}(k)}{1+r}\right) \quad (4)$$

for bride price girls and

$$a^*_{NoBP}(k) = \left(\frac{1}{\gamma}k - \frac{\Delta v_{NoBP}(k)}{1+r}\right) \quad (5)$$

[2] The first term $\left(\frac{1}{\gamma} - \frac{1}{\delta}\right) k$ is driven by the gender preferences of parents: if they care more about sons than about daughters, they are more willing to educate their sons. The second terms captures differential labor market returns, and it is likely to be positive in this context in which women have lower employment rates than men. The third term, which is equal to $\frac{2z_{01}+I_e\pi - z_{11} - z_{00}}{1+r}$, captures differential returns in the marriage markets, and its sign depends on the relative contribution of an educated woman compared to an educated man to the marital surplus. Note that we have imposed a standard supermodularity assumption, which implies that $z_{01} + z_{10} - z_{11} - z_{00} < 0$. Hence, the sign of $\frac{2z_{01} - z_{11} - z_{00} + I_e\pi}{1+r}$ depends on how z_{01} and $z_{10} - I_e\pi$ compare. Finally, the last term $I_e \frac{\pi}{\gamma(1+r)}$ is the impact of the bride price education premium on the parents' budget constraint, which alone should increase female schooling relative to male schooling.

for non-bride price girls. Under **A2**, the threshold in (4) is lower than the threshold in (5).

(*ii*) By the chain rule

$$\frac{\partial Pr(S_i = 1|I_e, a_i, k, \theta_i)}{\partial k} = -g\left(a^*(I_e, k)\right) \cdot \left[\frac{1}{\gamma} - \frac{1}{1+r}\frac{\partial \Delta v_e(k)}{\partial k}\right]. \tag{6}$$

The second term is positive under assumption **A3**. □

Proof of prediction 2

Proof Given the probability density function of ability $g(a_i)$, average ability of educated girls is equal to:

$$E[a_i|S=1] = E[a_i|a_i > a^*_{I_e}(k)] = \int_{a^*_{I_e}(k)}^{\infty} a_i g(a_i|a_i > a^*_{I_e}(k)) da_i$$

By the Leibniz integral rule, $\frac{\partial E[a_i|a_i > a^*]}{\partial a^*} = \frac{g(a^*)}{1-G(a^*)} \{E[a_i|a_i > a^*] - a^*\} > 0$.

Now, $a^*_{NoBP}(k) > a^*_{BP}(k)$ under assumption **A2**. This implies that

$$E[a_i|S=1; I_e = 0, k] > E[a_i|S=1; I_e = 1, k].$$

□

Proof of prediction 3

Proof Compare the two partial derivatives:

$$\frac{\partial Pr(S_i = 1|k, I_e = 1, \theta_i)}{\partial k} = -g\left(a^*_{BP}(k)\right) \cdot \left[\frac{1}{\gamma} - \frac{1}{1+r}\frac{\partial \Delta v_{BP}(k)}{\partial k}\right]$$

v.s.

$$\frac{\partial Pr(S_i = 1|k, I_e = 0, \theta_i)}{\partial k} = -g\left(a^*_{NoBP}(k)\right) \cdot \left[\frac{1}{\gamma} - \frac{1}{1+r}\frac{\partial \Delta v_{NoBP}(k)}{\partial k}\right]$$

For the derivative to be more negative for bride price females, it has to be the case that

$$g\left(a^*_{BP}\right) - g\left(a^*_{NoBP}\right) > \frac{\gamma}{1+r}\left[g\left(a^*_{BP}\right)\frac{\partial \Delta v_{BP}(k)}{\partial k} - g\left(a^*_{NoBP}\right)\frac{\partial \Delta v_{NoBP}(k)}{\partial k}\right].$$

Under **A5**, we can define $\frac{\partial \Delta v_{BP}(k)}{\partial k} = \frac{\partial \Delta v_{NoBP}(k)}{\partial k} = \frac{\partial \Delta v(k)}{\partial k}$. The above condition is then

$$g\left(a^*_{BP}\right) - g\left(a^*_{NoBP}\right) > \frac{\gamma}{1+r}\left[g\left(a^*_{BP}\right)\frac{\partial \Delta v(k)}{\partial k} - g\left(a^*_{NoBP}\right)\frac{\partial \Delta v(k)}{\partial k}\right],$$

which is implied under **A3**.

Under unimodality of $g()$, low education rates, and assumption **A2**, we have that $g\left(a^*_{BP}(k)\right) - g\left(a^*_{NoBP}(k)\right) > 0$. Thanks to **A4**, we have that:

$$-g\left(a^*_{BP}(k)\right) \cdot \left[\frac{1}{\gamma} - \frac{1}{1+r}\frac{\partial \Delta v_{BP}(k)}{\partial k}\right] < -g\left(a^*_{NoBP}(k)\right) \cdot \left[\frac{1}{\gamma} - \frac{1}{1+r}\frac{\partial \Delta v_{NoBP}(k)}{\partial k}\right].$$

□

Proof of prediction 4

Proof The response of girls' schooling to the school construction is therefore

$$\frac{\partial Pr(S_i = 1|k, I_e\theta_i)}{\partial k} = -\frac{g\left(a^*_e(k)\right)}{\gamma}\left[\frac{1}{\gamma} - \frac{1}{1+r}\frac{\partial \Delta v(k)}{\partial k}\right].$$

This quantity only varies between ethnic groups because of $(a^*_e(k))$, which maps one-to-one onto the baseline level of schooling $1 - G(a^*_e(k))$, since $G()$ is strictly monotonic. □

Appendix B: Data description

Cross-Cultural Data

Information on bride price practices is taken from the *Ethnographic Atlas* (Murdock, 1967) and LeBar (1972) for Indonesia and the *Ethnographic Atlas* (Murdock, 1967) and the *Ethnographic Survey of Africa* (Willis, 1966, Whiteley and Slaski, 1950, Schapera, 1953) for Zambia.

Our primary analysis for Indonesia uses the 1995 Indonesia Intercensal Survey, which records 174 different spoken languages. These are matched to 44 ethnic groups from the *Ethnographic Atlas* and LeBar (1972). To undertake the matching, we exploited the previous matching of ethnic groups to languages undertaken by Alesina, Giuliano and Nunn (2013), where the 1,265 ethnic groups of the *Ethnographic Atlas* were matched to one of 7,612 language groups in the *Ethnologue: Languages of the World* (Gordon, 2005). All but 11 of the 172 language groups in the Indonesia Intercensal Survey could be matched to an ethnicity from our sources. These comprise 0.29 percent of the observations with non-missing language data.

Our baseline analysis for Zambia uses the four rounds of the Zambia Demographic and Health Surveys (1996, 2001, 2007, and 2013). The Zambia DHS reports 65 distinct ethnic groups. Of these, we are able to match 53 of them to 30 more-coarsely defined ethnic groups from the *Ethnographic Atlas* and the *Ethnographic Survey of Africa* (Willis, 1966, Whiteley and Slaski, 1950, Schapera, 1953). The remaining unmatched groups are small and comprise less than 2.5 percent of the DHS sample. The matching was done by hand, relying heavily on Murdock (1959).

The ethnicity-level control variables used in the analysis (female participation in agriculture and lineage type) are taken from the *Ethnographic Atlas*. We create a measure of female participation in agriculture using variable *v*54 ("sex differences: agriculture") from the *Ethnographic Atlas*. The original variable records ethnic groups as belonging to one of the following categories: (1) males only, (2) males appreciably more, (3) differentiated but equal participation, (4) equal participation, (5) female appreciably more, (6) females only, and (7) absent or unimportant activity. Using this information, we create a measure of female participation in agriculture that takes on the value of one for categories 5 and 6 and zero otherwise, for non-missing values. We also create a second indicator variable that equals one if either agriculture was not present or was unimportant – i.e., category (6) – or if information was missing for the ethnicity.

A second control variable is a measure of the presence of matrilineal inheritance. This was based on variable $v43$ ("descent: major type"). The original variable groups ethnicities into one of the following categories: (1) patrilineal, (2) duolateral, (3) matrilineal, (4) quasi-lineages, (5) ambilineal, (6) bilateral, and (7) mixed. We construct a matrilineal indicator variable that takes on the value of one if variable $v43$ codes an ethnic group as belonging to group 3, and zero otherwise.

Indonesia

Indonesian Family Life Survey

The Indonesian Family Life Survey (IFLS) is an ongoing longitudinal study of households in Indonesia covering over 30,000 individuals. Data is gathered from 13 of Indonesia's 27 provinces, and the study is considered representative of 83 percent of the Indonesian population. This paper uses data from rounds 3 and 4 of the IFLS (Strauss, Beegle, Sikoki, Dwiyanto, Herawati and Witoelar, 2004, Strauss, Witoelar, Sikoki and Wattie, 2009), which, unlike previous rounds of the IFLS, includes questions about individuals' ethnicities. The first panel of appendix table A11 presents summary statistics on educational attainment for males and females, as well as household wealth, for all respondents aged 25–45. The second panel reports summary statistics on female and male marriage age and ln$(brideprice)$ for approximately 2,400 marriages documented in the round 3 IFLS and 3,200 marriages in the round 4 IFLS where bride price was paid. While marriage ages do not differ by whether the groom or bride belonged to a bride price ethnicity, average bride prices are significantly higher in bride price ethnic groups.

Indonesia Intercensal Survey

The Indonesia Intercensal Survey is a large-scale, nationally representative population survey of Indonesia carried out between the 1990 and 2000 censuses. It is housed by the Minnesota Population Center (1995). Importantly, it includes data on primary language spoken, which can be linked to ethnicity and matched to an ethnic group's bride price custom in the *Ethnographic Atlas*. It also contains information on educational attainment, birth year, and birth district which, following Duflo (2001), can be combined with data on the number of schools built in 1974 as part of a large-scale school construction program. Appendix table A1 presents summary statistics for the two sub-samples of this data set that we analyze in this paper. The first sample is used to estimate the impact of school construction for bride price and non-bride price females. This sample is composed of a treated group of individuals who were 2–6 at the time of school construction (1974) and an un-treated group of individuals who were 12–24 at the time of school construction. The second sample, which is used to compare the enrollment patterns of school-aged females in bride price and non-bride price ethnicities, consists of all individuals between the ages of 5 and 22.

Zambia

Data from the Zambia Fertility Preferences Study

Data on bride price amounts and beliefs about bride price and education are drawn from unique survey data collected in Lusaka in Fall 2014 as part of an experimental study on family planning. The study involves 715 couples living in the catchment area of Chipata clinic, a poor peri-urban segment of Lusaka. Each spouse of these couples was interviewed in private and was asked a series of questions on the practice of *lobola*, leading to a total of 1,430 observations.

Appendix table A13 reports summary statistics for the key variables.

Demographic and Health Survey

To study the effect of school construction on the enrollment of bride price and non-bride price children in Zambia, we pool the 1996, 2001, 2007, and 2013 rounds of the Zambia *Demographic and Health Survey*. When we analyze how school construction impacts school enrollment, we limit the sample to primary-school aged children (5–12), since most new schools are primary schools. The first panel of table A9 presents summary statistics for enrollment, wealth, and local female employment rates for these groups. Once we control for district (column 7), the only significant difference between the bride price and non-bride price groups is the female employment rate. When we analyze whether daughters in bride price ethnicities are more likely to be enrolled relative to daughters in non-bride price ethnicities, we use a sample of all school-aged children in the pooled DHS (ages 5–22). Summary statistics for this group are presented in the second panel of appendix table A7. The summary statistics show that bride price females are more likely on average to be enrolled in primary school.

Additionally, data from the pooled 2001, 2007, and 2013 DHS surveys allow us to test whether bride price is correlated with gender bias.[3] To do so, we form three indices for male and female respondents separately. The first index is the portion of decisions that the respondent says are undertaken either jointly by the husband and wife, or by the wife alone. The second index is the portion of times a respondent replied that a husband was justified in beating his wife, and the last index is the portion of times a respondent said that a wife was justified in refusing a husband sex. Since different questions were asked in different years and of different genders, we list the questions for each index below and note in parentheses the survey years a question was asked and whether it was asked of males or females or both. The questions in the first index take the form of, "Who has the final say on...?" The options were "health care" (2001, females; 2007, females; 2013, females; 2013, males), "making large household purchases" (2001, females; 2001, males; 2007, females; 2007, males; 2013, females; 2013, males), "visits to family or relatives" (2001, females; 2001, males; 2007, females; 2007, males; 2013, females; 2013, males), "deciding what to do with money wife earns" (2001, males; 2007, males), "deciding how many children to have" (2001, males; 2007, males), "deciding what to do with money husband earns" (2007, females; 2013, females), and "on making household purchases for daily needs" (2007, females; 2007, males; 2013, females; 2013, males).

The questions for the second index take the form, "Wife beating justified if..." The options were "if she goes out without telling him" (2001, females; 2001, males; 2007, females; 2007, males; 2013, females; 2013, males), "if she neglects the children" (2001, females; 2001, males; 2007, females; 2007, males; 2013, females; 2013, males), "if she argues with him" (2001, females; 2001, males; 2007, females; 2007, males; 2013, females; 2013, males), "if she refuses to have sex with him" (2001, females; 2001, males; 2007, females; 2007, males; 2013, females; 2013, males), and "if she burns the food" (2001, females; 2001, males; 2007, females; 2007, males; 2013, females; 2013, males).

The questions in the third index take the form, "reason for not having sex:", and the possible answers are "husband has STD" (2001, females; 2001, males; 2007, females; 2007, males; 2013, females; 2013, males), "husband has other women" (2001, females; 2001, males; 2007, females; 2007, males; 2013, females; 2013, males), "recent birth" (2001, females; 2001, males) and "tired, mood" (2001, females; 2001, males; 2007, females; 2007, males; 2013, females; 2013, males).

[3]The gender bias questions were not asked in the 1996 DHS.

Appendix C: 2SLS Education Estimates

This sections reports details of the specifications behind the estimates of the impact of education on bride price amounts, using the same procedure as in Duflo (2001), where school construction is used as an instrument for the educational attainment of women from bride price ethnic groups. As we have shown, for non-bride price ethnic groups, there was no impact on female education and therefore no first-stage predictive power. In line with Duflo (2001), we allow the effect of school construction to vary by a child's age in 1974, restricting the effect to be 0 if a child was older than 12 in 1974. Following Duflo (2001), we also restrict the sample to those born between 1950 and 1972. Unfortunately, the resulting sample of couples from bride price ethnicities who were asked questions about bride price is only 311.

The first-stage estimating equation is:

$$I(Completed\,Primary)_{idkt} = \alpha_d + \alpha_k + \alpha_t + \sum_{a=2}^{12} \beta_a\, Intensity_d \times I(age_{1974} = a)_k + \sum_j \mathbf{X'}_d \mathbf{I}_k^j \mathbf{\Gamma}_j + \epsilon_{idkt}, \quad (7)$$

where d denotes the district, i denotes the individual, t denotes the survey year, and k denotes the cohort. α_d denotes district fixed effects, α_k cohort fixed effects, and α_t survey-round fixed effects. $\sum_j \mathbf{X'}_d \mathbf{I}_k^j \mathbf{\Gamma}_j$ are the cohort-specific controls for the INPRES sanitation program, the enrollment rate in 1971, and the total number of school-aged children in 1971.

The second-stage equation is:

$$y_{idkt} = \alpha_d + \alpha_k + \alpha_t + \gamma I(CompletedPrimary)_i + \sum_j \mathbf{X'}_d \mathbf{I}_k^j \mathbf{\Gamma}_j + \mu_{idkt}, \quad (8)$$

where the outcome variable y_{idkt} is either the value of the bride price paid or the natural log of the bride price.

Appendix table A12 reports the 2SLS estimates. Column 1 reports the first-stage estimates, and shows that the instruments are jointly significantly (F-statistic of 2.13). Columns 2 and 3 report 2SLS estimates of the effect of primary schooling on bride price and log bride price amounts. Although the point estimates are imprecise, they corroborate the results from the OLS regressions. Completing elementary school increases bride price payments by 590 percent ($p < .01$).

Because the self-reported bride price amounts in the IFLS are most likely reported in nominal terms, in columns 4–6, we re-estimate the specifications in columns 1–3, but include marriage year fixed effects. The new first stage is quite weak, with a F-statistic of 1.62. The 2SLS estimates of the effect of primary school completion on bride price values are large and significant at the 1 percent level (column 5), while the effect on log bride price (column 6) is large (200 percent) but statistically insignificant.

Overall, the 2SLS estimates confirm the OLS estimates. However, because the point estimates are imprecise, the results should be interpreted with the appropriate caution.

Appendix D: Appendix tables

Table A1: Summary Statistics by Bride Price Practice for the 1995 Indonesia Intercensal Data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Bride Price		Non-Bride Price		Pooled Sample			
					Without Controls		With Controls	
	Mean	SD	Mean	SD	Difference	SE	Difference	SE
				Panel A. School Construction Sample				
Age	34.451	7.041	34.414	7.131	0.036	0.055	0.036	0.052
Female Primary Completion	0.639	0.480	0.605	0.489	0.034***	0.005	0.048*	0.027
Male Primary Completion	0.745	0.436	0.722	0.448	0.023***	0.005	0.032	0.020
Schools per 1000 School-Aged Children	2.220	1.089	1.991	0.790	0.229***	0.006	–	–
Matrilineal	0.093	0.290	0.109	0.312	-0.017***	0.002	-0.185**	0.048
Female Agriculture	0.037	0.189	0.034	0.182	0.003*	0.001	-0.074**	0.034
				Panel B. School Enrollment Sample (Ages 5-22)				
Age	12.750	4.930	13.234	4.990	-0.484***	0.029	0.186	0.145
Female Enrollment	0.610	0.488	0.577	0.494	0.033***	0.004	0.006	0.015
Male Enrollment	0.635	0.481	0.619	0.486	0.016***	0.004	-0.009	0.011
Matrilineal	0.084	0.277	0.126	0.333	-0.042***	0.002	-0.176**	0.043
Female Agriculture	0.037	0.188	0.045	0.208	-0.009***	0.001	-0.054*	0.029

Notes: This table reports summary statistics for the 1995 Indonesia Intercensal data. Columns 1 and 2 present means and standard deviations for ethnicities that traditionally practice bride price. Columns 3 and 4 present summary statistics for non-bride price ethnicities. Column 5 presents the difference in means and column 6 presents the standard error of the difference. Column 7 presents the coefficient on bride price in a regression of the row-name variables on bride price practice, district of birth fixed effects, and in the case of the school construction sample, treated or non-treated cohort fixed effects. Column 8 presents the standard error of the bride price coefficient, clustered at the district level. *, **, and *** indicate significance at the 10, 5, and 1% levels.

Table A2: Indonesia School Construction Regressions, Accounting for Other Cultural Traits

	(1)	(2)	(3)	(4)
	Dep var: Indicator variable for the completion of primary school			
	Baseline Regression	Matrilineal	Female Agriculture	Both
$I_k^{Post} \times Intensity_d \times I_e^{BridePrice}$	0.026**	0.026**	0.026**	0.026**
	(0.012)	(0.012)	(0.012)	(0.012)
$I_k^{Post} \times Intensity_d \times I_e^{NoBridePrice}$	-0.001	0.000	-0.003	-0.002
	(0.010)	(0.010)	(0.0105)	(0.011)
Ethnicity FE $\times I_k^{Post}$	Y	Y	Y	Y
Ethnicity FE $\times Intensity_d$	Y	Y	Y	Y
District FE $\times I_e^{BridePrice}$	Y	Y	Y	Y
Duflo Controls $\times I_e^{BridePrice}$	Y	Y	Y	Y
Duflo Controls	Y	Y	Y	Y
District FE	Y	Y	Y	Y
Cohort FE	Y	Y	Y	Y
F-test	2.84	2.50	3.13	2.78
Number of observations	65,403	65,403	65,403	65,403
Clusters	240	240	240	240
Adjusted R^2	0.184	0.184	0.184	0.184

Notes: The table reports re-estimates of the pooled Indonesia school construction regressions for females including controls for triple interactions of ethnicity-level characteristics, $Intensity_d$ and I_k^{Post}. The Duflo controls consist of cohort fixed effects interacted with the number of school-aged children in the district in 1971, cohort fixed effects interacted with the enrollment rate in 1971, and cohort fixed effects interacted with the regency-level implementation of a water and sanitation program under INPRES. These are the same controls as used in Duflo (2001). Standard errors are clustered at the district-of- birth level. *, **, and *** indicate significance at the 10, 5, and 1% levels.

Table A3: Bride Price Practice and the INPRES School Expansion in the 2010 Census Data

	(1)	(2)	(3)	(4)	(5)
	Dep var: Indicator variable for completion of primary school				
	Males	Females	Females	B.P. Females	Non B.P. Females
$I_k^{Post} \times Intensity_d$	0.016*	0.011		0.017**	-0.015
	(0.009)	(0.010)		(0.008)	(0.011)
$I_k^{Post} \times Intensity_d \times I_e^{BridePrice}$			0.017**		
			(0.008)		
$I_k^{Post} \times Intensity_d \times I_e^{NoBridePrice}$			-0.015		
			(0.011)		
Ethnicity FE $\times I_k^{Post}$	N	N	Y	Y	Y
Ethnicity FE $\times Indensity_d$	N	N	Y	Y	Y
District FE $\times I_e^{BridePrice}$	N	N	Y	Y	Y
Duflo Controls $\times I_e^{BridePrice}$	N	N	Y	N	N
F-test			3.70		
Duflo Controls	Y	Y	Y	Y	Y
District FE	Y	Y	Y	Y	Y
Cohort FE	Y	Y	Y	Y	Y
Number of observations	1,747,727	1,700,856	1,700,436	476,176	1,224,260
Adjusted R^2	0.116	0.176	0.194	0.196	0.183
Clusters	263	263	263	263	259

Notes: This table reports estimates of the Indonesia school construction regressions using the Indonesia 2010 Census data. The education attainment data come from a ten percent sample of the 2010 Indonesia Census. Bride price data are from the *Ethnographic Atlas* (Murdock, 1967) and LeBar (1972). I_k^{Post} refers to the treated cohort, born between 1968 and 1972. The untreated cohort is born between 1950 and 1962. The treatment level is the number of schools built in a district per 1,000 people in the school-aged population. The Duflo controls consist of cohort fixed effects interacted with the number of school-aged children in the district in 1971, cohort fixed effects interacted with the enrollment rate in 1971, and cohort fixed effects interacted with the regency-level implementation of a water and sanitation program under INPRES. These are the same controls as used in Duflo (2001). The subscript d indexes districts, i individuals, k cohorts, and e ethnic groups. Standard errors are clustered at the birth-district level. *, **, and *** indicate significance at the 10, 5, and 1% levels.

Table A4: Placebo Test of Bride Price Practice and the INPRES School Expansion Results in the 1995 Indonesia Intercensal Data

	(1)	(2)	(3)	(4)	(5)
	Dep var: Indicator variable for completion of primary school				
	Males	Females	Females	B.P. Females	Non B.P. Females
$I_k^{PlaceboPost} \times Intensity_d$	-0.004	-0.006		0.015	-0.004
	(0.006)	(0.005)		(0.014)	(0.007)
$I_k^{PlaceboPost} \times Intensity_d \times I_e^{BridePrice}$			0.015		
			(0.014)		
$I_k^{PlaceboPost} \times Intensity_d \times I_e^{NoBridePrice}$			-0.004		
			(0.007)		
Ethnicity FE $\times I_k^{Post}$	N	N	Y	Y	Y
Ethnicity FE $\times Indensity_d$	N	N	Y	Y	Y
District FE $\times I_e^{BridePrice}$	N	N	Y	Y	Y
Duflo Controls $\times I_e^{BridePrice}$	N	N	Y	N	N
Duflo Controls	Y	Y	Y	Y	Y
District FE	Y	Y	Y	Y	Y
Cohort FE	Y	Y	Y	Y	Y
F-test			1.80		
Number of observations	54,812	53,640	45,799	6,833	38,966
Number of clusters	254	247	232	140	210
Adjusted R^2	0.100	0.137	0.134	0.158	0.128

Notes: This table reports estimates of the Indonesia school construction regressions using a placebo treatment status instead of the true treatment status. The education attainment data are taken from the 1995 Indonesia Intercensal Survey. Bride price data are from the *Ethnographic Atlas* (Murdock, 1967) and LeBar (1972). $I_k^{PlaceboPost}$ refers to the placebo treated cohort, who are aged 12–17 in 1974. The placebo untreated cohort is aged 17–24 in 1974. $Intensity_d$ is the number of schools built in a district per 1,000 people in the school-aged population. The Duflo controls consist of cohort fixed effects interacted with the number of school-aged children in the district in 1971, cohort fixed effects interacted with the enrollment rate in 1971, and cohort fixed effects interacted with the regency-level implementation of a water and sanitation program under INPRES. These are the same controls as used in Duflo (2001). The subscript d indexes districts, i individuals, k cohorts, and e ethnic groups. Standard errors are clustered at the birth-district level. *, **, and *** indicate significance at the 10, 5, and 1% levels.

Table A5: Indonesia School Construction Results with Effect of School Construction by Age in 1974

	(1)	(2)	(3)
	Dep var: Indicator variable for the completion of primary school		
	All Females	Bride Price Females	Non-Bride Price Females
$Intensity_d \times I(Age_{1974} = 2) \times I_e^{BridePrice}$	0.011	0.011	
	(0.015)	(0.015)	
$Intensity_d \times I(Age_{1974} = 3) \times I_e^{BridePrice}$	0.039*	0.039*	
	(0.023)	(0.023)	
$Intensity_d \times I(Age_{1974} = 4) \times I_e^{BridePrice}$	0.012	0.012	
	(0.018)	(0.018)	
$Intensity_d \times I(Age_{1974} = 5) \times I_e^{BridePrice}$	0.046***	0.046***	
	(0.012)	(0.012)	
$Intensity_d \times I(Age_{1974} = 6) \times I_e^{BridePrice}$	0.026*	0.026*	
	(0.015)	(0.015)	
$Intensity_d \times I(Age_{1974} = 7) \times I_e^{BridePrice}$	-0.016	-0.016	
	(0.021)	(0.021)	
$Intensity_d \times I(Age_{1974} = 8) \times I_e^{BridePrice}$	-0.023*	-0.023	
	(0.014)	(0.014)	
$Intensity_d \times I(Age_{1974} = 9) \times I_e^{BridePrice}$	0.027*	0.027*	
	(0.015)	(0.015)	
$Intensity_d \times I(Age_{1974} = 10) \times I_e^{BridePrice}$	-0.006	-0.006	
	(0.015)	(0.015)	
$Intensity_d \times I(Age_{1974} = 11) \times I_e^{BridePrice}$	0.069***	0.069***	
	(0.022)	(0.022)	
$Intensity_d \times I(Age_{1974} = 12) \times I_e^{BridePrice}$	0.015	0.015	
	(0.025)	(0.025)	
$Intensity_d \times I(Age_{1974} = 2) \times I_e^{NoBridePrice}$	-0.002		-0.002
	(0.013)		(0.013)
$Intensity_d \times I(Age_{1974} = 3) \times I_e^{NoBridePrice}$	-0.005		-0.005
	(0.013)		(0.013)
$Intensity_d \times I(Age_{1974} = 4) \times I_e^{NoBridePrice}$	0.009		0.009
	(0.013)		(0.013)
$Intensity_d \times I(Age_{1974} = 5) \times I_e^{NoBridePrice}$	-0.010		-0.010
	(0.013)		(0.013)
$Intensity_d \times I(Age_{1974} = 6) \times I_e^{NoBridePrice}$	-0.033***		-0.033***
	(0.011)		(0.011)
$Intensity_d \times I(Age_{1974} = 7) \times I_e^{NoBridePrice}$	-0.014		-0.014
	(0.014)		(0.014)
$Intensity_d \times I(Age_{1974} = 8) \times I_e^{NoBridePrice}$	-0.008		-0.008
	(0.014)		(0.014)
$Intensity_d \times I(Age_{1974} = 9) \times I_e^{NoBridePrice}$	-0.000		-0.000
	(0.012)		(0.012)
$Intensity_d \times I(Age_{1974} = 10) \times I_e^{NoBridePrice}$	-0.006		-0.006
	(0.013)		(0.013)
$Intensity_d \times I(Age_{1974} = 11) \times I_e^{NoBridePrice}$	-0.003		-0.003
	(0.013)		(0.013)
$Intensity_d \times I(Age_{1974} = 12) \times I_e^{NoBridePrice}$	-0.001		-0.001
	(0.014)		(0.014)
F-test of bride price interactions	7.22		
F-test of non-bride price interactions	1.81		
Number of observations	92,325	13,700	78,625
Clusters	249	183	222
Adjusted R^2	0.171	0.168	0.171

Notes: This table estimates the impacts of school construction flexibly on individuals of different age cohorts, using data from the 1995 Indonesia Intercensal Survey. The effect of school construction on children aged 12 or greater in 1974 is restricted to be zero. Standard errors are clustered at the district level. *, **, and *** indicate significance at the 10, 5, and 1% levels.

Table A6: Effect of Bride Price by Religion in Full Indonesia School Construction Sample

	(1) Dep var: Primary school completion indicator
$I_k^{Post} \times Intensity_d \times I(Muslim)_i \times I_e^{BridePrice}$	0.022
	(0.021)
$I_k^{Post} \times Intensity_d \times I(Muslim)_i$	-0.003
	(0.013)
$I_k^{Post} \times Intensity_d \times I(NonMuslim)_i \times I_e^{BridePrice}$	0.045**
	(0.019)
$I_k^{Post} \times Intensity_d \times I(NonMuslim)_i$	-0.011
	(0.012)
Ethnicity FE $\times I_k^{Post}$	Y
Ethnicity FE $\times Intensity_d$	Y
District FE $\times I_e^{BridePrice}$	Y
Duflo Controls $\times I_e^{BridePrice}$	Y
Muslim	Y
Muslim $\times I_k^{BridePrice}$	Y
Muslim $\times Intensity_d \times I_k^{Post}$	Y
Duflo Controls	Y
District FE	Y
Cohort FE	Y
F-test	4.29
Number of observations	65,403
Adjusted R^2	0.187
Clusters	240

Notes: This table reports estimates of the effect of the INPRES school construction program on primary school completion, allowing for differences by religion (Muslim or not) and traditional bride price practice. Data on religion and educational attainment come from the 1995 Indonesia Intercensal Survey. Bride price data are from the *Ethnographic Atlas* (Murdock, 1967) and LeBar (1972). The F-test tests the joint significance of $I_k^{Post} \times Intensity_d \times I(Muslim)_i \times I_e^{BridePrice}$ and $I_k^{Post} \times Intensity_d \times I(NonMuslim)_i \times I_e^{BridePrice}$. The Duflo controls consist of cohort fixed effects interacted with the number of school-aged children in the district in 1971, cohort fixed effects interacted with the enrollment rate in 1971, and cohort fixed effects interacted with the regency-level implementation of a water and sanitation program under INPRES. These are the same controls as used in Duflo (2001). Standard errors are clustered at the birth district level. *, **, and *** indicate significance at the 10, 5, and 1% levels.

Table A7: Summary Statistics from the 1996, 2001, 2007, and 2013 Zambia Demographic and Health Surveys

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Bride Price		Non-Bride Price			Full Sample		
	Mean	SD	Mean	SD	Difference	Se	Coefficient	Se
			School Construction Sample (Ages 5 to 12)					
Female Enrollment	0.628	0.483	0.613	0.487	0.015**	0.007	0.011	0.013
Male Enrollment	0.595	0.491	0.595	0.491	-0.001	0.007	0.003	0.014
Age	8.328	2.251	8.359	2.263	-0.031	0.024	-0.010	0.026
Schools/Area	0.093	0.278	0.094	0.277	-0.001	0.003	-0.006	0.005
Matrilineal	0.587	0.492	0.493	0.500	0.094***	0.005	0.079	0.070
Female Agriculture	0.896	0.305	0.711	0.453	0.185***	0.004	0.198***	0.038
Wealth Index	2.930	1.340	2.890	1.360	0.040**	0.016	0.038	0.069
Female Employment Rate (All Adults 25-45)	0.637	0.481	0.680	0.466	-0.044***	0.007	-0.031**	0.011
			School Enrollment Sample (Ages 5 to 22)					
Female Enrollment	0.559	0.497	0.550	0.498	0.009*	0.005	0.014	0.010
Male Enrollment	0.603	0.489	0.612	0.487	-0.0008	0.005	0.003	0.010
Age	12.448	5.078	12.384	5.027	0.064*	0.038	0.069	0.068
Schools/Area	0.099	0.287	0.101	0.289	-0.003	0.002	-0.004	0.004
Matrilineal	0.582	0.493	0.488	0.500	0.094***	0.004	0.075	0.068
Female Agriculture	0.895	0.306	0.713	0.452	0.182***	0.003	0.189***	0.037
Wealth Index	3.043	1.358	3.023	1.382	0.020*	0.012	0.030	0.070

Notes: This table reports summary statistics by bride price practice in the pooled 1996, 2001, 2007, and 2013 rounds of the Zambia DHS. The first panel presents summary statistics for the sample of children aged 5–12 used in the school construction analysis, and the second panel of the table presents summary statistics for the sample of children aged 5–22 used in the enrollment analysis. Columns 1 and 2 present means and standard errors for the bride price group, and columns 3 and 4 present means and standard errors for the non-bride price group. Column 5 presents the difference in the means and column 6 gives the standard error of the difference. Column 7 is the coefficient on bride price practice in a regression of the row-name variable on bride price practice and district and year fixed effects. Column 8 is the standard error of this coefficient clustered at the district level. *, **, and *** indicate significance at the 10, 5, and 1% levels.

Table A8: Placebo Test of School Construction and Primary School Enrollment by Bride Price Pratice in the Pooled Zambia DHS (1996, 2001, and 2007)

	(1)	(2)	(3)	(4)	(5)
		Dep var: School enrollment indicator			
	Male	Females	Females	Bride Price Females	Non Bride Price Females
$Schools_{d,t+1}/Area_d$	0.046	-0.069		-0.227*	-0.012
	(0.063)	(0.082)		(0.116)	(0.104)
$Schools_{d,t+1}/Area_d \times I_e^{BridePrice}$			-0.230**		
			(0.116)		
$Schools_{d,t+1}/Area_d \times I_e^{NoBridePrice}$			-0.011		
			(0.109)		
$Schools_{d,t}/Area_d$	0.001	0.070		0.294**	-0.014
	(0.064)	(0.074)		(0.119)	(0.095)
$Schools_{d,t}/Area_d \times I_e^{BridePrice}$			0.295**		
			(0.117)		
$Schools_{d,t}/Area_d \times I_e^{NoBridePrice}$			-0.020		
			(0.101)		
Number of observations	12,073	12,370	12,370	3,554	8,816
Adjusted R^2	0.400	0.393	0.393	0.438	0.375
Clusters	70	70	70	63	69

Notes: This table reports estimates of the differential impact of present and future school building in Zambia on bride price and non-bride price females. The sample consists of children aged 5–12 at the time of the survey in the 1996, 2001, and 2007 rounds of the DHS. We are unable to include the 2013 DHS, since we do not know how many schools will be built in the future. The treatment variable, $Schools_{dt}$, is the number of schools built in a district d by year t (the survey round of the DHS). This is normalized by the area of the district, $Area_d$. $Schools_{d,t+1}$ is the number of schools built by 2001 in 1996, the number of schools built by 2007 in 2001, and the number of schools built by 2013 in 2007. d denotes a district, t denotes a survey year, and e denotes an ethnic group. Standard errors are clustered at the district level. *, **, and *** indicate significance at the 10, 5, and 1% levels.

Table A9: Zambia School Construction Regressions, Controlling for Ethnicity-Level Characteristics

	(1)	(2)	(3)	(4)
	Dep var: School enrollment indicator			
	Baseline	Female Agriculture	Matrilineal	Both
$Schools_{dt}/Area_d \times I_e^{BridePrice}$	0.042***	0.136**	0.070***	0.186***
	(0.014)	(0.066)	(0.019)	(0.069)
$Schools_{dt}/Area_d \times I_e^{NoBridePrice}$	-0.007	0.080	0.021	0.120**
	(0.014)	(0.057)	(0.017)	(0.057)
Age by Round by Bride Price FE	Y	Y	Y	Y
Ethnicity by Round FE	Y	Y	Y	Y
Ethnicity by District FE	Y	Y	Y	Y
F-test	5.22	5.91	5.02	7.36
Number of observations	22,191	22,191	22,191	22,191
Clusters	71	71	71	71
Adjusted R Squared	0.398	0.398	0.398	0.398

Notes: This table reports estimates of the impacts of Zambian school construction, but controlling for traditional female participation in agriculture and traditional matrilineality. d denotes a district, t denotes a survey year, and e denotes an ethnic group. Standard errors are clustered at the district level. *, **, and *** indicate significance at the 10, 5, and 1% levels.

Table A10: Contemporary Bride Price Payments in Indonesia and Zambia

Panel A. Indonesian Sample		
IFLS (2000, 2007)	$I_e^{BridePrice} = 0$	$I_e^{BridePrice} = 1$
Observations	5,428	2,195
$\%(BP > 0)$	91%	92%
Mean payment / per capita GDP	25%	51%
Median payment / per capita GDP	4%	9%

Panel B. Zambian Sample		
ZFPS (2014)	$I_e^{BridePrice} = 0$	$I_e^{BridePrice} = 1$
Observations	525	179
$\%(BP > 0)$	85%	88%
Mean payment / per capita GDP	103%	175%
Median payment / per capita GDP	52%	67%

Notes: This table reports summary statistics on the prevalence of bride price practices and the size of bride prices paid in Indonesia and Zambia. The table draws on data from the 2001 and 2007 rounds of the Indonesia Family Life Survey and from the ZFPS (Fall 2014). The top 1 percent of bride price payments in the IFLS have been trimmed due to the presence of a few very extreme values. Bride price data are from the *Ethnographic Atlas* (Murdock, 1967) and LeBar (1972) for Indonesia and from the *Ethnographic Atlas* (Murdock, 1967) and the *Ethnographic Survey of Africa* (Willis, 1966, Whiteley and Slaski, 1950, Schapera, 1953) for Zambia.

Table A11: Summary Statistics for Adults Aged 25-45 in the Indonesian Family Life Survey

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Bride Price		Non-Bride Price			Full Sample		
	Mean	SD	Mean	SD	Difference	SE	Coefficient	SE
				All Respondents Aged 25 to 45				
Probability a Female Attended Junior Secondary School	0.633	0.482	0.571	0.495	0.062***	0.010	0.041**	0.016
Probability a Female Attended Upper Secondary School	0.485	0.500	0.439	0.496	0.046***	0.010	0.044**	0.018
Probability a Female Attended College	0.100	0.300	0.108	0.311	-0.008	0.006	-0.021	0.013
Probability a Male Attended Junior Secondary School	0.724	0.447	0.640	0.480	0.085***	0.010	0.013	0.016
Probability a Male Attended Upper Secondary School	0.566	0.496	0.495	0.500	0.070***	0.011	0.022	0.017
Probability a Male Attended College	0.146	0.353	0.124	0.330	0.021***	0.007	0.028**	0.014
Household Assets	-0.005	1.373	0.092	1.365	-0.097***	0.019	-0.184***	0.032
				All Married Couples				
Female Marriage Age	22.307	6.333	22.641	6.362	-0.334**	0.168	-0.032	0.271
Male Marriage Age	26.718	7.494	22.076	7.717	-0.358*	0.211	0.426	0.331
Log(Bride Price)	13.240	2.365	12.671	1.886	0.569***	0.058	0.225**	0.112

Notes: This table reports summary statistics for either adult respondents aged 25–45 to rounds 3 and 4 of the Indonesian Family Life Survey or couples who responded to questions regarding a recent marriage. Columns 1 and 2 present the mean and standard deviations of the row-name variables for individuals belonging to ethnic groups with a bride price tradition. Columns 3 and 4 do the same for individuals from non-bride price traditions. Column 5 presents the difference and column 6 presents the standard error of the difference. Column 7 presents the coefficient on bride price practice in a regression of the row-name variable on bride price practice and year and district fixed effects. Column 8 presents the (robust) standard error of the bride price coefficient. The full data set of adults 25–45 contains 37,410 observations. The data set of recent marriages, which includes data on bride price and husband and wives' marriage ages, contains 6,987 observations.

Table A12: Instrumented Effect of Primary School Completion on Bride Price Values in Rounds 3 and 4 of the IFLS

	(1) $I(Completed\ Primary)_i$	(2) $BridePrice_i$	(3) $Ln(BridePrice)_i$	(4) $I(Completed\ Primary)_i$	(5) $BridePrice_i$	(6) $Ln(BridePrice)_i$
$Intensity_d \times I(Age_{1974} = 2)_i$	-0.033			-0.050		
	(0.197)			(0.244)		
$Intensity_d \times I(Age_{1974} = 3)_i$	-0.040			0.074		
	(0.523)			(0.503)		
$Intensity_d \times I(Age_{1974} = 4)_i$	-0.049			0.011		
	(0.132)			(0.158)		
$Intensity_d \times I(Age_{1974} = 5)_i$	-0.233			-0.255		
	(0.294)			(0.221)		
$Intensity_d \times I(Age_{1974} = 6)_i$	-0.122			-0.305		
	(0.360)			(0.505)		
$Intensity_d \times I(Age_{1974} = 7)_i$	-0.117			0.007		
	(0.190)			(0.289)		
$Intensity_d \times I(Age_{1974} = 8)_i$	-0.156			-0.015		
	(0.149)			(0.234)		
$Intensity_d \times I(Age_{1974} = 9)_i$	0.857***			0.941*		
	(0.286)			(0.566)		
$Intensity_d \times I(Age_{1974} = 10)_i$	-0.125			-0.462		
	(0.302)			(0.320)		
$Intensity_d \times I(Age_{1974} = 11)_i$	0.289			-0.419		
	(0.613)			(0.583)		
$Intensity_d \times I(Age_{1974} = 12)_i$	-0.175			0.062		
	(0.212)			(0.366)		
$I(Completed\ Primary)_i$		2,305,491.295**	5.903***		5,336,942.058***	1.949
		(925,994.344)	(1.493)		(1,613,087.934)	(1.211)
Survey-Year FE	Y	Y	Y	Y	Y	Y
District FE	Y	Y	Y	Y	Y	Y
Cohort FE	Y	Y	Y	Y	Y	Y
Duflo Controls	Y	Y	Y	Y	Y	Y
Marriage-Year FE	N	N	N	Y	Y	Y
F-stat	2.13			1.62		
Number of observations	311	311	279	309	309	277
Adjusted R^2	-0.020	-0.261	-0.589	0.065	-2.056	0.381
Clusters	94	94	94	94	94	94

Notes: This table reports two-stage least squares estimates of the effect of primary school completion on the bride price payment received at marriage. Following the specification used in Duflo (2001), we instrument for primary school completion with the interactions of treatment intensity and age in 1974 fixed effects. The effect of school construction on cohorts 12 or greater in 1974 is restricted to be zero. The sample consists of couples from ethnicities with non-token bride price who responded to questions about bride price payment values in rounds 3 and 4 of the Indonesian Family Life Survey. Treatment is assigned based on district of birth. When district of birth is missing, it is assigned based on the respondent's current district of residence. Columns 4–6 include year of marriage fixed effects. d denotes a birth district and i denotes an individual. The Duflo controls consist of cohort fixed effects interacted with the number of school-aged children in the district in 1971, cohort fixed effects interacted with the enrollment rate in 1971, and cohort fixed effects interacted with the regency-level implementation of a water and sanitation program under INPRES. These are the same controls as used in Duflo (2001). Standard errors are clustered at the district level. *, **, and *** indicate significance at the 10, 5, and 1% levels.

Table A13: Summary Statistics for ZFPS Data

	Mean	SD	N
$\ln BP_{ie}$	7.510	1.196	442
$I(Primary)_i$	0.943	0.231	442
$I(JuniorSecondary)_i$	0.507	0.501	442
$I(Secondary)_i$	0.267	0.443	442
$MarriageAge_i$	20.446	4.115	442
$I(HusbCompletePrimary)_i$	0.986	0.116	442
$I(HusbCompleteJuniorSecondary)_i$	0.731	0.444	442
$I(HusbCompleteSecondary)_i$	0.468	0.500	442
$HusbandMarriageAge_i$	25.937	6.495	441

Notes: This table contains summary statistics for all couples in the ZFPS data.

Table A14: Determinants of Bride Price Payment Amounts, Including Controls for Female Wealth

	(1)	(2)		(3)	(4)
	Indonesia			**Zambia**	
			Dep Var: Log Bride Price Amount		
$I(CompletePrimary)_i$	0.379***	0.261**	$I(CompletePrimary)_i$	0.127	0.122
	(0.117)	(0.123)		(0.139)	(0.143)
$I(CompleteJuniorSec)_i$	0.492***	0.369***	$I(CompleteJunSec)_i$	0.430***	0.419***
	(0.107)	(0.111)		(0.134)	(0.135)
$I(College)_i$	0.654***	0.329***	$I(CompleteHigherSec)_i$	0.313**	0.303**
	(0.109)	(0.123)		(0.126)	(0.128)
$MarriageAge_i$	0.070*	0.058	$MarriageAge_i$	-0.0745***	-0.0792***
	(0.037)	(0.039)		(0.00968)	(0.0104)
$MarriageAge_i^2$	-0.001**	-0.001*	$MarriageAge_i^2$	-2.78e-08	-3.66e-08
	(0.001)	(0.001)		(9.54e-08)	(9.41e-08)
$I(HusbCompletePrimary)_i$		0.258*	$I(HusbCompletePrimary)_i$		0.0862
		(0.135)			(0.180)
$I(HusbCompletedJunSec)_i$		0.382***	$I(HusbCompleteJunSec)_i$		0.0283
		(0.114)			(0.154)
$I(HusbCollege)_i$		0.370***	$I(HusbCompleteHigherSec)_i$		0.0158
		(0.121)			(0.126)
$HusbMarriageAge_i$		0.024*	$HusbMarriageAge_i$		-0.0452
		(0.014)			(0.0350)
$HusbMarriageAge_i^2$		-0.000	$HusbMarriageAge_i^2$		0.00113*
		(0.000)			(0.0006)
Pre-Marriage Female Wealth Control	Y	Y	Pre-Marriage Female Land/Property Controls	Y	Y
Quadratic in Year of Marriage	Y	Y	Quadratic in Year of Marriage	Y	Y
Ethnicity FE	Y	Y	Ethnicity FE	Y	Y
Survey-Round FE	Y	Y	Survey Round FE	n.a.	n.a.
Number of observations	1,951	1,827	Number of observations	455	454
Adjusted R^2	0.446	0.402	Adjusted R^2	0.269	0.270

Notes: This table reports estimates of the relationship between female education and bride price. The columns regress the natural log of bride price payments at the time of marriage on various covariates, including controls for female wealth (log pre-marriage female-owned assets in Indonesia and land ownership in Zambia). The sample consists of all marriages for which bride price was reported. For Indonesia, bride price and educational attainment data are taken from the 2001 and 2007 rounds of the Indonesian Family Life Survey (rounds 3 and 4). Data on bride price in Zambia are taken from the ZFPS data set. The subscript i indexes a female respondent. Robust standard errors are reported in parentheses. *, **, and *** indicate significance at the 10, 5, and 1% levels.

Table A15: Effects of School Construction on Male Schooling

	(1) **Indonesia** Primary school completion	(2) **Zambia** Primary school enrollment
$I_k^{Post} \times Intensity_d \times I_e^{BridePrice}$	0.009	
	(0.012)	
$I_k^{Post} \times Intensity_d \times I_e^{NoBridePrice}$	0.018**	
	(0.007)	
$Schools_{dt} / Area_d \times I_e^{BridePrice}$		0.034*
		(0.017)
$Schools_{dt} / Area_d \times I_e^{NoBridePrice}$		0.008
		(0.016)
Full set of control variables	Y	Y
F-test	0.40	0.56
Number of observations	63,717	21,772
R^2	0.129	0.397
Number of clusters	247	71

Notes: This table replicates the specifications of table 3 (column 3) and of table 4 (column 3) using a sample of males. The data from Indonesia are from the 1995 Intercensal Survey. The data for Zambia are from the pooled 1996, 2001, 2007, and 2013 Zambia Demographic and Health Surveys. Bride price data are from the *Ethnographic Atlas* (Murdock, 1967) and LeBar (1972) for Indonesia and from the *Ethnographic Atlas* (Murdock, 1967) and the *Ethnographic Survey of Africa* (Willis, 1966, Whiteley and Slaski, 1950, Schapera, 1953) for Zambia. The standard errors are clustered at the district level. *, **, and *** indicate significance at the 10, 5, and 1% levels.

Table A16: Relationship between the Practice of Bride Price and Enrollment for Females in Zambia in the Pooled 1996, 2001, and 2007 DHS

	(1)	(2)
	Dep var: Enrollment indicator	
$I_e^{BridePrice}$	0.033**	0.033**
	(0.013)	(0.013)
Wild bootstrap p-value	[0.024]	[0.016]
Age Controls	Y	Y
District FE	Y	Y
Survey Year FE	Y	Y
Ethnicity Controls	N	Y
Number of observations	23,868	23,868
Adjusted R^2	0.343	0.343
Clusters	29	29

Notes: This table reports estimates of the relationship between bride price customs and female enrollment rates in Zambia in the pooled 1996, 2001, and 2007 DHS. The columns regress an indicator variable for whether a child is enrolled in school on an indicator variable for whether the child is a member of an ethnic group that practices non-token bride price. The sample consists of girls aged 5–22. Bride price data are from the *Ethnographic Atlas* (Murdock, 1967) and the *Ethnographic Survey of Africa* (Willis, 1966, Whiteley and Slaski, 1950, Schapera, 1953) for Zambia. Age controls consist of age and age squared, and cultural controls consist of indicator variables for belonging to a matrilineal ethnicity and belonging to an ethnicity where women traditionally participate more in agriculture than men. Standard errors, clustered at the ethnicity level, are reported in parentheses. p-values obtained using the wild bootstrap procedure with 500 draws are reported in square brackets. *, **, and *** indicate significance at the 10, 5, and 1% levels.

Table A17: Effect of School Construction by Bride Price Practice, Accounting for Pre-treatment Female Education Rates

	(1) Indonesia		(2) Zambia
$I_k^{Post} \times Intensity_d \times I_e^{BridePrice}$	-0.017	$Schools_{dt} / Area_d \times I_e^{BridePrice}$	-0.489
	(0.024)		(0.767)
$I_k^{Post} \times Intensity_d \times I_e^{NoBridePrice}$	-0.032	$Schools_{dt} / Area_d \times I_e^{NoBridePrice}$	-0.518
	(0.024)		(0.769)
$I_k^{Post} \times Intensity_d \times$ Baseline Female Education Controls	Y	$Schools_{dt} / Area_d \times$ Baseline Female Education Controls	Y
Ethnicity FE $\times I_k^{Post}$	Y	Age FE $\times I_e^{BridePrice} \times$ Survey Year FE	Y
Ethnicity FE $\times Intensity_d$	Y	Ethnicity $\times$ Year FE	Y
District FE $\times I_e^{BridePrice}$	Y	Ethnicity $\times$ District FE	Y
Duflo Controls $\times I_e^{BridePrice}$	Y		
Cohort FE $\times I_e^{BridePrice}$	Y		
F-test	1.13	F-test	2.10
Number of observations	65,291	Number of observations	22,180
Adjusted R^2	0.190	Adjusted R^2	0.399
Clusters	232	Clusters	71

Notes: This table reports estimates of the heterogeneous effects of school construction regressions in Indonesia (table 3) and Zambia (table 4) controlling for baseline (pre-treatment) female education rates. The sample for Indonesia consists of women born between 1950 and 1962 and those born between 1968 and 1972. The sample for Zambia consists of girls who were aged 5–12 during the survey year. For Indonesia, educational attainment data are taken from the 1995 Intercensal Data. For Zambia, educational attainment data are taken from the 1996, 2001, 2007, and 2013 Demographic and Health Surveys. Baseline female primary completion rates are calculated using the untreated sample in Indonesia (those aged 12–24 in 1974) and using individuals aged 12–45 in the survey year in Zambia. Bride price data are from the *Ethnographic Atlas* (Murdock, 1967) and LeBar (1972) for Indonesia and from the *Ethnographic Atlas* (Murdock, 1967) and the *Ethnographic Survey of Africa* (Willis, 1966, Whiteley and Slaski, 1950, Schapera, 1953) for Zambia. We calculate the portion of females in these samples who complete primary school at the district by ethnicity level, and include a third-degree polynomial of this variable interacted with $I_k^{Post} \times Intensity_d$ in Indonesia and interacted with $Schools_{dt} / Area_d$ in Zambia as a control in the regression. The subscript d denotes a birth-district in Indonesia and a current district in Zambia, k denotes a cohort in Indonesia, and t denotes a survey year in Zambia. The Duflo controls consist of cohort fixed effects interacted with the number of school-aged children in the district in 1971, cohort fixed effects interacted with the enrollment rate in 1971, and cohort fixed effects interacted with the regency-level implementation of a water and sanitation program under INPRES. These are the same controls as used in Duflo (2001). Standard errors are clustered at the district level. *, **, and *** indicate significance at the 10, 5, and 1% levels.

Table A18: Ethnic groups in our Indonesian sample and whether bride price is practiced

Ethnicity	$I_e^{BridePrice}$	Ethnicity	$I_e^{BridePrice}$
BALINESE	0	ALORESE	1
CHAM	0	AMBONESE	1
DANI	0	BANGGAI	1
ENGGANO	0	BATAK	1
IBAN	0	BELU	1
JAVANESE	0	BUNGKU	1
KENYAH-KAYAN-KAJANG	0	DAWAN	1
KERAKI	0	GORONTALO	1
KUBU	0	ILI-MANDI	1
MARINDANI	0	KEI	1
MENTAWEIA	0	MACASSARE	1
MIMIKA	0	MALAYS	1
MINANGKAB	0	MANOBO	1
REJANG	0	MINAHASANS	1
SASAK	0	MUJU	1
SOROMADJA	0	MUNA	1
SUMBANESE	0	NIASANS	1
SUMBAWANE	0	PANTAR	1
SUNDANESE	0	ROTINESE	1
SUVANESE	0	SUGBUHANO	1
WAROPEN	0	TOBELORES	1
		TOMINI	1
		TORADJA	1

Notes: The table reports the ethnic groups in our Indonesian sample, as well as the value of the bride price indicator variable for that ethnic group.

Table A19: Ethnic groups in our Zambian sample and whether bride price is practiced

Ethnicity	$I_e^{BridePrice}$	Ethnicity	$I_e^{BridePrice}$
BEMBA	0	BWILE	1
BISA	0	ILA	1
CHEWA	0	IWA	1
CHOKWE	0	LUNGU	1
KAONDE	0	MAMBWE	1
KUNDA	0	MBUNDA	1
LALA	0	PL.TONGA	1
LAMBA	0	SALA	1
LOZI	0	SHONA	1
LUANO	0	SOLI	1
LUCHAZI	0	TUMBUKA	1
LUNDA (LUA)	0		
LUNDA (LW)	0		
LUVALE	0		
NYANJA	0		
SHILA	0		
SWAKA	0		
TABWA	0		
USHI	0		

Notes: The table reports the ethnic groups in our Zambian sample, as well as the value of the bride price indicator variable for that ethnic group.

Appendix E: Appendix figures

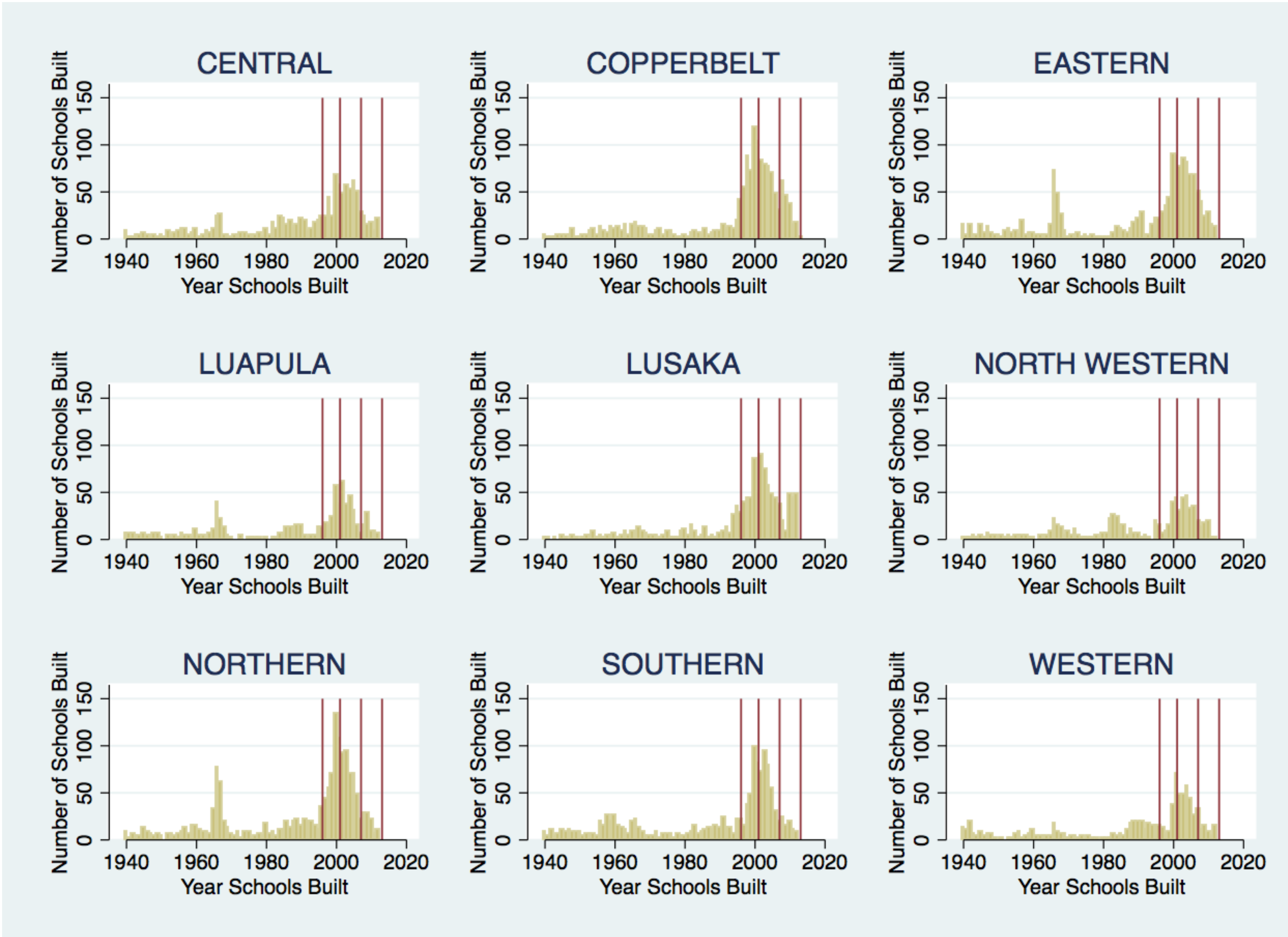

Figure A1: Number of schools constructed each year for each province in Zambia (Ministry of Education, Government of Zambia).

References

Alesina, Alberto, Paola Giuliano, and Nathan Nunn, "On the Origins of Gender Roles: Women and the Plough," *Quarterly Journal of Economics*, 2013, *128* (2), 155–194.

Becker, Gary S., "A Theory of Marriage: Part I," *Journal of Political Economy*, 1973, pp. 813–846.

Chiappori, Pierre-André, Murat Iyigun, and Yoram Weiss, "Investment in Schooling and the Marriage Market," *American Economic Review*, 2009, *99* (5), 1689–1713.

Duflo, Esther, "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review*, 2001, *91* (4), 795–813.

Gordon, Raymond G., *Ethnologue: Languages of the World, 15th Edition*, SIL International, 2005.

LeBar, Frank M., *Ethnic Groups of Insular Southeast Asia. Volume 1: Indonesia, Adaman Islands, and Madagascar*, New Haven: Human Relations Area Files Press, 1972.

Minnesota Population Center, "Integrated Public Use Microdata Series, International: Version 6.1," 1995.

Murdock, George Peter, *Africa: Its Peoples and Their Cultural History*, New York: McGraw-Hill Book Company, 1959.

_ , *Ethnographic Atlas*, Pittsburgh: University of Pittsburgh Press, 1967.

Schapera, I., *The Tswana*, London: International African Institute, 1953.

Shapley, Lloyd S. and Martin Shubik, "The Assignment Game I: The Core," *International Journal of Game Theory*, 1971, *1* (1), 111–130.

Strauss, J., F. Witoelar, B. Sikoki, and A.M. Wattie, "The Fourth Wave of the Indonesian Family Life Survey (IFLS4): Overview and Field Report," 2009.

_ **, K. Beegle, B. Sikoki, A. Dwiyanto, Y. Herawati, and F. Witoelar**, "The Third Wave of the Indonesian Family Life Survey (IFLS): Overview and Field Report," 2004.

Whiteley, Wilfred and J. Slaski, *Bemba and Related Peoples of Northern Rhodesia*, London: International African Institute, 1950.

Willis, Roy, *The Fipa and Related Peoples of South-West Tanzania and North-East Zambia*, London: International African Institute, 1966.