Targeting men, women or both to reduce child marriage.*

Rachel Cassidy[†], Anaya Dam[‡], Wendy Janssens[§], Umair Kiani[¶], and Karlijn Morsink[‖]

May 22, 2025

Abstract

We study whether targeting men, women, or both with the same intervention in the same context is more effective in improving women's and girls' outcomes in settings governed by gendered social norms. Using a cluster-randomized controlled trial in Pakistan, we evaluate an edutainment intervention aimed at delaying girls' marriage. Targeting men reduces child marriage within target households, while targeting women generates sustained reductions in child marriages at village level. Our results can be rationalized by a Bayesian persuasion model, where women are more hesitant to deviate from norms than men. Extending the model to gender-segregated information transmission accounts for the village-level spillover effects.

Keywords: Social Norms, Targeting, Gender, Child Marriage, Edutainment

*We are grateful to Dan Anderberg, Alison Andrew, Chris Barrett, Michele Belot, Lauren Bergquist-Falco, Pierre-Andre Chiappori, Paul Collier, Brian Dillon, Rachel Glennerster, John Hoddinott, Seema Jayachandran, Travis Lybbert, Imran Rasul, Bob Rijkers, Frank Schilbach, Tavneet Suri, and Nicholas Swanson for comments. We thank audiences at the Workshop in Gender and Family Economics 2022, PacDev 2022, 100 Years of Economic Development 2022, NEUDC 2022, Applied Development Economics Conference at Lahore School of Economics 2022, Yale EGC-Asian Development Bank Institute 2023, and seminars at Cornell, University of Chicago, Passau, Tinbergen Institute, Wageningen, and the Institute for Fiscal Studies. We are grateful to Mayan Ali, Vinh Phan, Gabriella Villanueva and Anouk van Veldhoven for excellent research assistance. We are indebted to Saleem Ahmed, Wasim Durrani, Ronald van Moorten, Katinka Moonen, Brigitte Obertop, Mohammed Qazilbach, Carmen Reinoso, and Raffat Shuja. This research project was funded by ESRC-DFID Grant No: ES/N014650/1, Oxfam Novib and Center for Economic Analysis of Risk, Georgia State University. The research in Punjab was conducted in partnership with the Punjab Social Protection Authority. IRB approval was obtained from University of Oxford, Approval No. R56430/RE001. No Objection Certificates were obtained from local authorities in Sindh (No.SO(LE-II)HD/Misc-64/2018) and Punjab (No.21(11)/MISC/PSPAP&D/2017). AEA RCT Registry number AEARCTR-0006098.

[†]Institute for Fiscal Studies & World Bank. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, nor those of the Executive Directors of the World Bank or the governments they represent.

[‡]Utrecht University & University of Oxford

[§]Vrije Universiteit Amsterdam, Tinbergen Institute & Institute of Global Health & Development

[¶]Centre for Economic Research in Pakistan

[‖]Utrecht University, Wageningen University & Cornell University. k.morsink@uu.nl

1 Introduction

Social norms prescribing the behaviour and actions of women and girls are pervasive and persistent.[1] A common explanation for this persistence is that men, as dominant decision-makers, face fewer private incentives to improve outcomes for women (Bernhardt et al., 2018). Interventions targeting men's beliefs about gender norms have been shown to be effective in some contexts, for example to improve female labour force participation (Bursztyn et al., 2020). Similarly, interventions targeting women or adolescent girls directly can also yield positive outcomes (Andrew et al., 2024; Ashraf et al., 2020; Buchmann et al., 2023; Edmonds et al., 2021). However, women and girls frequently face backlash from families and communities when they deviate from traditional gender norms. In certain settings, the private gains to women and girls from changing their behaviour may be outweighed by the social sanctions they may incur for violating these norms (Aizer, 2010; Beaman et al., 2009; Chakraborty and Serra, 2023; Leibbrandt et al., 2018; Macmillan and Gartner, 1999).[2] This trade-off raises the question of whether interventions should target men or women to change behaviour that is governed by gendered social norms.

We experimentally test the effects of targeting men, women, or both in the *same context* with the *same intervention* focused on a high-stakes household decision. Specifically, we investigate parental decisions to delay marriage of their adolescent girls, in the presence of age-of-marriage norms that favour early marriage. Early marriage is a widespread and long-standing custom in rural Pakistan, the setting of our study. On average, in 35% of marriages we observe in our study villages the bride is below 18 years of age at the time of marriage. Furthermore, approximately 60% of parents believe that: they should marry their daughter before 18; others in the community expect them to marry their daughter before 18; there will be sanctions if they deviate from the early age-of-marriage norm. The negative welfare consequences of early marriage, including its impacts on health, education,

[1] For example, norms relating to marriage (e.g., Bicchieri et al. (2014); Buchmann et al. (2023); Bursztyn et al. (2017)), taking up work outside the home (e.g., Bertrand et al. (2015); Jayachandran (2021)), domestic responsibilities and caregiving (e.g., Breen and Cooke (2005)), female genital cutting (e.g., Bellemare et al. (2015); Fors et al. (2021); Gulesci et al. (2021)). See also Jayachandran (2015) and Bursztyn and Jensen (2017) for an overview.

[2] See also Buchmann et al. (2023), Edmonds et al. (2021) and Andrew et al. (2024) who show that girls who change behaviour and deviate from gendered norms experience backlash from families or communities – for example through a negative penalty in the marriage market, less favorable assessments by parents of their daughter's behaviour, and reduced mental health.

domestic violence, and labour market participation for both women and their children, are well documented (e.g., Chari et al. (2017); Field and Ambrus (2008); Hicks and Hicks (2019); Jensen and Thornton (2003)). While both fathers and mothers in our study are aware of these consequences and share similar beliefs about the health benefits of delaying marriage, mothers are systematically less likely than fathers to want to deviate from the early age-of-marriage norm.[3] Even if delaying marriage offers substantial private benefits to women and their daughters, concerns about social repercussions may lead women to be more reluctant to act on these benefits. This suggests that targeting men could potentially be more effective in altering marriage decisions in this context.

We conduct a cluster-randomized controlled trial (RCT) of an educational entertainment ("edutainment") intervention that discusses the costs and benefits of delaying marriage. The intervention is implemented in 177 villages across two provinces in Pakistan — where the village constitutes the relevant marriage market.[4] We study whether targeting men, women, or both is more effective in delaying girls' marriages. The intervention consists of a mobile cinema screening of a locally produced street-theatre performance, followed by facilitated group discussions. The performance was developed in collaboration with local NGOs and enacted by local actors, with the aim of reducing early marriage. Based on extensive piloting and focus group discussions, we identified key factors that are relevant in our context and were therefore incorporated in our intervention: the potential benefits of delaying marriage, such as **improved health outcomes** and **increased spousal quality;** and the potential costs associated with **deviating from prevailing early age-of-marriage norms**.[5]

To implement the RCT, we randomly selected ten households with unmarried adolescents in each village (henceforth, "target households"). On average, these households represent approximately 17% of those with children on the marriage market in each village.[6] Villages

[3]Our baseline data show that mothers have a stronger preference for complying with community age-of-marriage norms than fathers (Table 2). Similar patterns have been confirmed in the literature on female genital cutting (e.g. Bellemare et al. (2015); Gage and Van Rossem (2006); Sagna (2014)).

[4]Our village-level marriage data show that 90% of marriages observed in our study happen between spouses from the same village.

[5]In our context, unlike in Buchmann et al. (2023) and Andrew et al. (2024) there are no immediate gains in terms of the adolescent girls' education from delaying marriage, as the large majority of girls do not complete any education beyond the age of 13.

[6]In our 44 control villages we observe 310 marriages over 28 months, and thus approximately 3 marriages per village per year. Over that same period of time, we observe 44 marriages in our target households, so that is roughly 0.51 marriages in a year. So 17% of the "village marriage market" is treated.

were then randomly assigned to one of four treatment arms: the edutainment intervention was offered in target households to women only, men only, both, or no one (control group). Other households in our study villages outside the target households did not participate in the edutainment intervention.

We conducted a panel survey at baseline, midline, and endline with approximately 1,700 target households, covering 5,100 individuals: an adolescent boy or girl aged 13–17 and their primary male and female caregivers (henceforth, "fathers" and "mothers"). The midline survey was conducted six months after the intervention, while the endline took place approximately eighteen months post-intervention. The surveys collected data on marriage outcomes, beliefs about health risks and spousal quality, such as expected spousal education conditional on the girl's age-of-marriage, and incentivized measures of beliefs about age-of-marriage norms. In addition, we collected monthly observational data on all marriages in the 177 study villages throughout the study period, recording the age of marriage and the origins of both brides and grooms.

By endline, 71% of the sample of adolescent girls are either married, or past the age of 18 and so no longer at risk of child marriage. Recall that girls in target households were aged 13-17 and unmarried at baseline. In the control group, by midline 6% are married and by endline 12%. We find no significant effect on child marriage in target households, at midline or at endline, when the intervention is targeted at women. In contrast, targeting men leads to a significant reduction in the likelihood of child marriage: a 4.1 percentage point decline by midline – a 66% reduction relative to the control mean, and a 5.3 percentage point decline by endline – a 43% reduction relative to the control mean. When both men and women are treated jointly, we observe no significant effect at midline. However, by endline, child marriage declines by 5.2 percentage points – a 42% reduction relative to the control mean. These findings are also robust to hazard model specifications. **Comparing effect sizes across treatment arms**, we find that at midline, the reduction in the probability of child marriage in the Male treatment arm is marginally significantly larger than in the Female+Male arm and significantly larger than in the Female arm. By endline, we cannot reject the null hypothesis that the effects across all three treatment arms are equal.

To address potential concerns about **reporting bias**, we rely on marriage reports from adolescents — a practice common in recent studies on child marriage, as adolescents may have fewer incentives to misreport their marriage age. We find a 95% consistency rate in

marriage reports across adolescents, mothers, and fathers. All our findings are robust to using marriage reports from fathers, mothers, or the adolescent child. Additionally, we verify marriage status by tracking the residence of girls, leveraging the fact that married girls move out of their parental household, and 90% marry within the same village, allowing for reliable tracing. Finally, kernel density plots of reported marriage ages show no evidence of bunching around the legal marriage age, further supporting the validity of our data.

At the village level, where the data include all adolescent girls in the village, we observe that on average 35% of marriages in the control group have a bride who was below 18. Unlike the household-level findings, we find that targeting women, either alone or jointly with men, leads to significant reductions in village-level child marriage rates for girls at both midline and endline. By endline, targeting women alone reduces the probability of observing a child marriage in the village-level data by 21 percentage points – a 60% decline relative to the control mean. Targeting both women and men jointly results in a 25 percentage point reduction – equivalent to 72% of the control mean. In contrast, targeting men only generates significant reductions at midline, but these effects do not persist at endline. The comparisons of effects sizes at midline and endline suggests that the impact of targeting men alone fades over time.

To rationalize these findings, we build on a model of Bayesian persuasion between spouses, as developed in the companion paper, Anderberg et al. (2024).[7] That model posits that parents weigh the uncertain benefits of delaying marriage against potential social sanctions for deviating from community age-of-marriage norms. Parents differ systematically in the weight they place on these sanctions, with mothers facing higher stigma costs from norm deviation than fathers, which we support with data.[8] The edutainment intervention introduces *new* information to one or both parents regarding the health and spousal quality benefits of delaying marriage, potentially shifting their preferred marriage timing. If parents disagree on the preferred marriage timing and only one parent is treated, they may attempt to persuade their spouse to align the marriage decision with their own prefer-

[7]The theoretical companion paper develops a model of intimate partner violence arising from disagreement between spouses about important household decisions, which is applied to child marriage as an illustration. They reference the household-level results on child marriage, as presented in full in this paper, and not the village-level results.

[8]In psychology, law and criminology it is well-established that women are less willing to deviate from norms and perceive a higher risk of social sanctions (Blackwell, 2000; Eagly and Carli, 1981; Grasmick et al., 1993; Kuang et al., 2020; Nolen-Hoeksema, 2004; Steinberg and Monahan, 2007; Tibbetts and Herz, 1996).

ence. Given mothers' stronger aversion to norm deviation, mothers that are treated alone may persuade untreated fathers to marry early, whereas fathers that are treated alone may persuade untreated mothers to delay.

To account for the divergence in child marriage outcomes between target households and the broader village level, we extend the model of Anderberg et al. (2024) to incorporate equilibrium child marriage outcomes at the village level. Specifically, we assume that parents who receive new information from the intervention share it within their gender-specific networks, for which we provide empirical support. Given mothers' greater reluctance to deviate from community age-of-marriage norms, we also assume that they have stronger incentives than fathers to shift these norms – which we empirically demonstrate is what happens[9] – and to disseminate information about the benefits of delaying marriage, as that would lower or remove costs of stigma from delaying marriage for them.[10] Given the gender-segregated nature of information transmission, our model shows that, under plausible diffusion rates, information provided to men fades out over time before reaching the broader community. In contrast, when women are targeted, alone or jointly with men, the information reaches all households, leading to a new equilibrium with delayed marriages. These predictions are consistent with our village-level findings.

Our findings have strong policy implications. In contexts such as Pakistan where child marriage is prevalent, even if parents wish to deviate from the prevailing norm and delay their daughter's marriage, they may get "stuck" marrying off their daughters early due to coordination failure across households. In theory, a signal to coordinate on a new norm could shift the norm in the community to later marriage (Buchmann et al., 2023). We provide evidence that a community-based intervention can act as a signal and facilitate coordination on a new norm when delivered to the right decision-makers (in our case, audiences involving women). Such mechanisms pertain not only to child marriage but potentially also to other practices that are harmful for women and girls, such as female genital cutting (Bellemare

[9]In Section 2 we refer to the components of a social norm, as defined by Bicchieri (2006). Two of the five components of a norm are 1) village-level changes in the likelihood of child marriage (empirical expectation); 2) beliefs of mothers' about community members' beliefs about the acceptable age of marriage (contingency). This supports our assumption that indeed norms shift. We show in Online Appendix Table O20 that our results are not driven by our intervention correcting misperceptions of the norm.

[10]It is important to note that we are assuming that the behaviour that is regulated by the age-of-marriage norm is the *choice to delay a daughter's marriage* and not the *choice to talk about the intervention.*

et al., 2015; Fors et al., 2021; Gulesci et al., 2021).

We expand a literature that aims to **change outcomes for women and girls in a context where gender norms prescribe their behaviour** (Bellemare et al., 2015; Bertrand et al., 2015; Bicchieri et al., 2014; Buchmann et al., 2023; Bursztyn et al., 2017; Efferson et al., 2020; Fors et al., 2021; Gulesci et al., 2021; Jayachandran, 2021). While women and girls are commonly targeted, a nascent body of work has started targeting men (Bursztyn et al., 2017; Dhar et al., 2022). We contribute to this literature by providing a first clean test of whether to target men or women with the *same* intervention in the *same* context.

We also connect to a literature on **norm change in communities** over time. Studies show that community engagement (Andrew et al., 2024), social signalling (Buchmann et al., 2023; Karing, 2024) as well as the introduction of an intermediate, less harmful norm (Corno and La Ferrara, 2022; Gulesci et al., 2021) can be effective to facilitate coordination on a new norm. Our study provides an alternative explanation — that the transmission of new information from individuals targeted by the intervention can endogenously result in non-targeted households also changing behaviour.

We also contribute to the literature on **interventions that seek to reduce child marriage;** see Malhotra and Elnakib (2021) for a review. Laws prohibiting child marriage — where they exist — often suffer from implementation issues in low-resource contexts, and have limited to no impact on delaying child marriage (Collin and Talbot, 2017; McGavock, 2021; Wilson et al., 2022). Economists have typically focused on poverty and consumption-smoothing as drivers of early marriage (Corno et al., 2020; Corno and Voena, 2023; Tapsoba, 2023). Many interventions address these drivers by easing households' financial constraints (Baird et al., 2010, 2011, 2019; Duflo et al., 2015). While such interventions have been shown to delay marriage, they require relatively large upfront costs, and are typically conditional on schooling. Such interventions may have less impact in contexts such as Pakistan where girls leave school very early. Instead, we evaluate an intervention that addresses information and beliefs. Our intervention is low-cost and it can be straightforwardly replicated by local NGOs elsewhere. It produces sizeable impacts on delaying marriage, with spillovers at the village level.

Finally, we contribute to a growing literature on the effectiveness of **edutainment interventions** in enhancing outcomes for women, e.g., health, education, women's empower-

ment, as well as broader social cohesion (Armand et al., 2020; Banerjee et al., 2019; Donati et al., 2022; Glennerster et al., 2023; Green et al., 2020; Jensen and Oster, 2009; La Ferrara et al., 2012; Roy et al., 2019; Siddique et al., 2024). We compare the impact of broadcasting the same content to different groups of individuals.

child-married

2 Context and child marriage intervention

Rates of early marriage remain high in Pakistan: Data from the most recent round of the Demographic and Health Surveys (DHS) in Pakistan shows that 26% of women aged 20-29 were child-married, and 46% of women aged 40-49 were child-married. Our study takes place in the provinces of Sindh and Punjab, where the legal age of marriage for girls and boys is 18 years. In our study villages, in particular, early marriage is widespread and persistent – in the parents' generation 47% of women were married before 18, and approximately 60% of parents at baseline state that marrying a girl before 18 is acceptable. During our study period, 35% of marriages in our control villages had a bride that was below 18. In our study areas, marriage is a decision made jointly by parents. In 90% of households, fathers are expected to make final decisions about marriage, but mothers are highly involved in 70% of households. These marriage decisions are made in the presence of community age-of-marriage norms that favour early marriage, especially for girls. Following the various components of the definition of a social norm[11] by Bicchieri (2006, pp. 11-28) we show that 35% of married girls were married before 18. In addition, approximately 60% of parents: believe that community members marry their adolescent close to this common age (empirical expectation, Table 1 row 3); believe that there exists an early age-of-marriage norm (contingency, Table 1 row 2); prefer to conform to the age-of-marriage norm (conditional preference, Table 1 row 4 and 6); believe that they should marry their daughter before 18; believe that others believe that they should marry their daughter before 18 (normative expectations, Table 1 row 5), and believe that others will prefer to confirm to the norm and may sanction behaviour (sanctions; Table 2 all rows). Our village-level marriage data shows that 90% of marriages observed in our study happen between spouses from the same village – making the village the relevant marriage market – and nearly all parents believe other community members find it acceptable to marry off a girl below 18.

[11]Note that Bicchieri (2006: p.12) states that a social norm can exist in a population even if only a subset of the population adheres to empirical regularities that define a social norm.

Both mothers and fathers are aware of these norms, and do not differ in their beliefs about such norms at baseline. Mothers are, however, significantly more averse to deviating from the community age-of-marriage norm than fathers. Table 2 reports gender differences in parents' preferences for compliance with community age-of-marriage norms and expected community responses to child marriage for girls. Significantly more mothers than fathers report that they: agree that marriage should not happen if the community disapproves; expect that others will not support a daughter who wants to delay a marriage; and expect that the community would do nothing if the girl is married young. Mothers also expect fewer community members to go to the police or pressure the family if a girl is married young. This lower willingness of women to deviate from social norms for fear of social sanctions, is well-established in psychology, law and criminology. Women are recognized to perceive a higher risk of social sanctions, including shame and embarrassment, than men (Blackwell, 2000; Grasmick et al., 1993; Nolen-Hoeksema, 2004; Tibbetts and Herz, 1996). Women are also more susceptible to pressure associated with deviating from common behaviours of their social groups (Eagly and Carli, 1981; Kuang et al., 2020; Steinberg and Monahan, 2007).

Table 1: Gender differences in parents' early age-of-marriage norms

	(1) Female	(2) Male	(3) Difference	(4) Observations
1. Agree that marrying a girl who is less than 18 is acceptable	0.588 (0.493)	0.657 (0.476)	-0.070 (0.302)	380
2. Community members out of 10 that believe that that less than 18 is an acceptable or best age for a girl's marriage	6.131 (2.356)	6.652 (2.274)	-0.521 (0.101)	380
3. Community members out of 10 that marry their adolescent close to the common marriage age in the community	6.523 (1.845)	6.238 (1.916)	0.285 (0.140)	380
4. Likelihood that you marry your adolescent also close to the common marriage age in the community	6.121 (2.382)	6.077 (2.174)	0.043 (0.850)	380
5. Community members out of 10 that expect you to marry your adolescent close to the common marriage age in the community	6.261 (1.960)	5.978 (2.041)	0.283 (0.188)	380
6. Likelihood that you marry your adolescent as per the community's expectation	5.955 (2.425)	6.083 (2.196)	-0.128 (0.606)	380

Notes: The table shows attitudes and beliefs of fathers and mothers about early age-of-marriage norms for girls. Data from the control group at endline is used, as many of these variables were not collected at baseline. Columns 1 and 2 present the means and standard deviations (in parentheses) of the variables for mothers and fathers respectively. Column 3 presents differences in means, with p-values in parentheses from robust standard errors clustered at the village level using bivariate regressions of gender on the variables of interest. Column 4 shows the total number of observations for mothers and fathers combined. The first row presents a binary variable that is equal to one if the respondent agrees or strongly agrees to the questions that marrying a girl who is 12-15 or 16-17 is an acceptable age. The second row represents respondents' beliefs about the number of community members out of 10 that believe that less than 18 is an acceptable or best age for a girl to get married. We take the highest value among the responses for the following categories: under 14 as the best age, 14-15 as the best age, 16-17 as the best age, 12-15 as an acceptable age, and 16-17 as an acceptable age. The third row represents respondents beliefs about the number of community members out of 10 that marry their adolescent close to the community's common age of marriage. The fourth row represents respondents expectation about the likelihood that they will therefore also marry their adolescent close to the community's common age of marriage. The fifth row represents respondents beliefs about the number of community members out of 10 that expect the respondent to marry their adolescent close to the community's common age of marriage. The sixth row represents respondents expectation about the likelihood out of 10 that they will conform to the community's expectations of marrying the adolescent at that common age. For all variables, a higher value is indicative of early age-of-marriage-norms. Stars indicate: * 1 percent ** 5 percent * 10 percent level of significance.

Our intervention, which took place during the first six months of 2019, consisted of a mobile cinema screening of an educational and entertaining (edutainment) street-theatre performance. Street theater is a popular art form in South Asia, and uses emotion, immersion, and perspective-taking to address sensitive social and political themes, whilst raising awareness among the public in an entertaining way. To ensure standardisation of the content across communities, the play was pre-recorded and screened to target households as a mobile cinema. The movie of the street theatre lasted approximately 10 minutes.

The play was written and directed by local NGOs and performed by local actors. Our local partner organizations led the content development — combining evidence about households' decisions about the timing of daughters' marriage[12] with their own experience working

[12]E.g., household income, such as dowry and bride price (Chari et al., 2017; Corno et al., 2020; Corno and Voena, 2023; Jensen and Thornton, 2003), health of children and grandchildren (Chari et al., 2017; Jensen and Thornton, 2003), spousal and match quality (Adams and Andrew, 2024),

in the study areas, and focus group discussions they held separately with men, adolescent boys, women, and adolescent girls in pilot communities. While beliefs about the likelihood of health benefits, spousal quality benefits and costs in terms of deviation from community age-of-marriage norms were recognized in all groups, there were differences between the groups in terms of the perceived importance of each of these aspects. Our edutainment therefore brought out the rights of women and girls, the benefits in terms of spousal quality of delaying marriage, and the perceived costs of deviating from prevailing age-of-marriage norms. The latter were included since the literature on edutainment emphasizes the need to portray struggles that individuals may face in changing behaviour, and how they overcome them, rather than focusing solely on benefits (e.g., Banerjee et al. (2019)). This content was conveyed through characters embodying various positions on early marriage and enacting the identified everyday situations. Our local partners obtained approval for their programming, and we received approval for our study from the local governments of Punjab and Sindh after review of the protocols and questionnaires.[13]

A first example of a scene in the play that highlights the *potential benefit of marriage delay* for the girl and her future household is a scene where a highly educated boy arrives back to the village. He is seen arguing with his parents that he wants to marry a girl that is at least 18, because he has had a realization, during his time away from the village, that marrying a young girl might lead to marriage failure and she might not be mature enough to take care of herself, their children or their household. A second example of a scene about potential benefits of marriage delay shows an old man reminiscing on the tragic memory of his deceased first wife. He tells a community member that she was only 15 when she succumbed to complications during childbirth. He further explained that he remarried when he and his wife were both 19, and he has had a good marriage with many offspring.

and norms governing the age of marriage (Anderson, 2007; Buchmann et al., 2023).

[13]IRB approval was obtained from University of Oxford, Approval No. R56430/RE001. No Objection Certificates from local authorities were obtained in Sindh (No.SO(LE-II)HD/Misc-64/2018) and Punjab (No.21(11)/MISC/PSPAP&D/2017).

Table 2: Gender differences in parents' reluctance to deviate from early age-of-marriage norms

	(1) Female	(2) Male	(3) Difference	(4) Observations
1. Agree that marriage should not happen if the community disapproves	0.576 (0.495)	0.416 (0.494)	0.160 (0.028)***	736
2. Expect that no one in community would support daughter if she wants to delay marriage	0.611 (0.488)	0.582 (0.494)	0.030 (0.048)	736
3. Expect that community would do nothing if they found out a girl is married young	0.804 (0.397)	0.538 (0.499)	0.266 (0.039)***	736
4. Community members out of 10 that believe the community will tell the police if a girl is married young	2.067 (2.102)	4.045 (2.522)	-1.978 (0.202)***	736
5. Community members out of 10 that believe the community will pressure the families if a girl is married young	2.488 (2.322)	4.304 (2.398)	-1.817 (0.216)***	736

Notes: The table shows attitudes and beliefs of fathers and mothers about their reluctance to deviate from early age-of-marriage norms for girls from the baseline. Columns 1 and 2 present the means and standard deviations (in parentheses) of the variables for mothers and fathers respectively. Column 3 presents differences in means with p-values in parentheses from robust standard errors clustered at the village level (unit of randomisation) using bivariate regressions of gender on the variables of interest. Column 4 shows the total number of observations for mothers and fathers combined. The first row presents a binary variable that is equal to one if the respondent agrees or strongly agrees with the statement "If the community disapproves of a marriage, then the marriage should not go ahead even if the parents want it to," and zero otherwise. The second row presents a binary variable that is equal to one if the respondent answered "nobody" to the question "who from the community would support her if your daughter would want to delay a suggested marriage?" The third row presents a binary variable that is equal to one if the respondent answered "nothing" to the question "what would the community do if they found out that a girl was about to be married before the legal marriage age?" A higher value in these rows is associated with a higher preference for norm compliance, where a positive difference between mothers and fathers indicates mothers to be more conservative than fathers. The fourth row presents the respondents' beliefs about how many community members out of 10, averaging over men and women, think that the community would tell the police if they found out that a very young girl was about to be married. The fifth row presents the respondents' beliefs about how many community members out of 10, averaging over men and women, think that the community would speak to or pressure the families if they found out that a very young girl was about to be married. A higher value in these rows is associated with a lower preference for norm compliance, where a negative difference between mothers and fathers indicates mothers to be more conservative than fathers. The number of observations refers to the total number of females and males in all households with adolescent daughters. Baseline data is from Sindh as security situation prevented baseline data collection in Punjab. Stars indicate: * 1 percent ** 5 percent * 10 percent level of significance.

A scene that highlights *potential costs of marriage delay resulting from community norms about the age-of-marriage* shows a marriage ceremony, where the bride is only 12 years old. Community members attending the ceremony are deliberating whether this is an appropriate age of marriage for the girl. In this scene, the adult women and men claim that marrying at this age is what they have always done in this community. A woman then speaks up and says that they want to stick to traditions because they are concerned about the opinions of the community if they delay a girls' marriage. A young nephew then convinces them that delaying marriage for girls reduces health risks, and that traditions need to be reconsidered when they are harmful to young brides.

The movie screening was followed by a group discussion that followed a standardised

format, facilitated by a gender specialist from the local NGOs. The objective of the group discussion was to reiterate the information presented in the movie – the experience of participants with the movie, the positions of the various characters, and the consequences of early marriage. Examples of questions asked are: "*What parts of the story represents real-life happenings in your community?*" or "*From what you have seen today, what are the merits/costs of early child marriage*?" A list of the questions that were used in the group discussion is provided in Appendix A.1.2. Community discussion, sometimes with facilitation, is common after a street-theatre performance. The discussion post-movie lasted for 30 minutes and facilitators encouraged active participation from all attendees. In the second visit to each village approximately three months later, the local NGOs again conducted structured group discussions based around the content of the movie with the same participants. These discussions were led by the same gender specialist and lasted 50 minutes.

3 Experimental design

We conducted a cluster-randomized controlled trial of the edutainment intervention in 177 villages in Sindh and Punjab provinces in Pakistan. The villages were randomly assigned into four treatment groups: 1) Female arm: Targeting women only; 2) Male arm: Targeting men only; Female+Male arm: 3) Targeting the intervention at both genders simultaneously; or 4) Control arm: No intervention. The screening of the theater performance and the group discussions were held in communal areas in the village: typically a compound or a room of a community building. In most villages men and women were not allowed to attend the intervention jointly in the Female+Male arm. Therefore, in this arm, men and women attended separately but simultaneously. Care was taken to inform both groups that the other group would be watching the same screening and would be discussing the same topics at the same time.[14] The objective of the joint arm was to increase the likelihood that spouses discuss the content and prevent the option that they can hide the content of the edutainment, as is possible in the Male only arm or the Female only arm. We deliberately did not let the spouses watch the edutainment jointly with other men and women, as that would have also led to inter-gender dynamics in the sessions as well as in the broader

[14] Any effect of the Female+Male arm should be interpreted as an effect of both members watching and discussing the edutainment with their own gender, and knowing that their spouse did the same. The same-gender discussions within the edutainment and gender session for women in the Female+Male arm, thus, could be comparable to in the Female only arm, and similarly for men in the Female+Male arm versus the Male arm.

community, an additional mechanisms that could have explained any effects.

Our target households in treatment arms were mobilised to participate in the intervention a few days prior to the screening of the theatre performance, by a designated focal person from the village. The gender of the invited individual from target households was according to their village's treatment status — men, women, or both. The intervention was mainly held indoors, and spaces were limited to twenty participants, and only individuals from our target households were invited. Other households in our study villages did not participate, and we estimate that, on average, we treat 17% of the households with children of marriageable age.[15]

3.1 Sampling and randomization

We conducted the sampling and randomization in three stages. First, the local NGOs selected villages for inclusion in the study. To minimize the risk of contamination across villages, we excluded villages that had less than 1.6 kilometers between their outer boundaries, based on a mapping exercise conducted with the local NGOs and local government offices. This left 80 eligible villages from Sindh Province and 97 eligible villages from Punjab Province. Next, we collected baseline village-level data on key village characteristics including presence of and distance to primary and secondary schools, presence of female teachers, distance to nearest town, presence of a health center and tea shop, population size, and mobility of women in the village.

We next conducted a household listing exercise to obtain a census of households in each village that were eligible to participate in our study. The eligibility criteria were that households needed to have at least one unmarried adolescent son or daughter aged 13-17 years and needed to have at least one adult father or male caregiver and one adult mother or female caregiver in the household. Ages of adolescents were verified either through National Identity Card (NIC), or a Birth Registration Certificate (B-Form) where applicable.[16] From the census of eligible households, we randomly selected ten households per village to participate in our study: five households with an adolescent boy ("boy households") and five

[15] See footenote 6.

[16] In Pakistan, citizens who are age 18 years and older are eligible for a National Identity Card. It is possible, given that our villages are remote and rural, that not all households have applied for these cards. Birth Registration Certificates (B-Form) are issued by the local government at the time of birth. They contain the name and date of birth (DOB) of the individual in question and the name and DOB of their parents as well as siblings.

households with an adolescent girl ("girl households"). The focal person and the NGOs explained to target households that a study was being conducted to understand the effects of showing an educational and entertaining movie to men, women, or men and women jointly, and that we would in some cases only invite household members from one gender to attend. In our study areas many community meetings and NGO meetings are gender-segregated, so this is not perceived as unusual. As a result, the planned sample size was 1,770 households (10 households in each of the 177 villages – with three respondents per household: father, mother, and either an unmarried adolescent boy or an unmarried adolescent girl). Some villages did not have a sufficient number of households meeting the selection criteria due to their small size, leading to a final sample size of 1,687 households (5,061 respondents): 756 households (2,268 respondents) in Sindh province, and 931 households (2,793 respondents) in Punjab province.

After the baseline survey, we randomly assigned villages to one of the four treatment arms, after stratification first by region/district and second by Mahalanobis distance matching on village-level characteristics.[17] 44 villages were assigned to receive the Male intervention; 45 villages were assigned to receive the Female intervention and 44 villages the Female+Male intervention. The remaining 44 villages were assigned to the Control arm.

4 Data and descriptive statistics

Baseline data were collected in July and August 2018, before the intervention was introduced in the treatment villages. In Punjab Province, we were able to collect village-level data and to conduct the baseline household listing exercise (including adolescent's gender, age and marital status) and select our target households; but we were unable to conduct a full baseline survey due to the security situation at the time of the baseline. Randomization of

[17]Mahalanobis distance calculates the distance between observations using a metric, which standardizes differences across multiple covariates by the covariance matrix. Our study area covered two districts in Sindh, and two in Punjab. Within each district, the Mahalanobis distance score was computed for each of our sample villages based on the following list of village-level baseline variables: boys only primary school; girls only primary school; mixed gender primary school; girls only secondary school; distance to nearest primary girls' school (in minutes); distance to nearest girls' secondary school (in minutes); are girls allowed to leave the compound; distance to the nearest town (in minutes); presence of teashop; whether the village is a main village or sub-village; total number of households; availability of female teachers in girls' school; and presence of a primary health care center. If in a district, the variable had less than or equal to 2 observations or a correlation $>= 0.6$ with other variables it was not included in its score computation. Villages were grouped into groups of 8 villages based on the Mahalanobis distance score, and these 8 villages were subsequently randomly assigned to either one of the treatment arms or the control arm.

villages into treatment arms was conducted after the baseline survey, and before the start of the interventions. The security situation in Punjab subsequently eased, and allowed the intervention to go ahead in Punjab according to the assigned treatment arms. The midline survey was conducted from November 2019 until March 2020, i.e. six months after the intervention had ended in the treatment villages and just before the start of the COVID-19 pandemic. When we visited the target households in Punjab for the midline survey, we included retrospective baseline questions on some outcomes of interest. An endline survey was conducted between September 2020 and March 2021, i.e., 18 months after the intervention and during the later stages of the COVID-19 pandemic. We pre-registered the RCT and submitted pre-analysis plans for the analysis of the midline and endline data.

4.1 Household survey data

Household survey data were collected at baseline in Sindh, and at midline and endline in both provinces with three respondents per household: father, mother and adolescent child. From the father we collected data on household demographic information; education, employment, and marital status of all household members; household financial, wealth indicators, and expenditures as well as decision-making in the household. From all three respondents we collected data on preferences, attitudes, expectations, and beliefs about marriage, with a focus on child marriage (i.e., marriage before age 18). We also asked all three household members about the adolescent child's education, marital status, and (conditional on marriage) the age at the time of marriage and spousal characteristics. For girls who were married, and who had moved out from their parental residence, we also tracked their current residence and confirmed the marriage status with a short module.[18]Enumerators were extensively trained in asking sensitive questions (e.g., surrounding child marriage), in first responder and referral procedures.[19]

4.2 Village-level observational data on marriages

Throughout the study period from September 2018 until March 2021, we collected monthly village-level marriage data in both Sindh and Punjab provinces. We do not use official administrative data on marriage registrations, since pilot investigation showed that few marriages are registered, and we would especially expect child marriage to remain unreg-

[18]Girls move out from their parents house after marriage. 90% marry a spouse from the same village and 95% move within 30km of their parental home

[19]See footenote 13.

istered due to its illegal nature. Instead, our research field coordinators visited a central location in each village at monthly intervals. The field coordinator mapped out all marriages that happened in the village since their last visit by interviewing a series of individuals independently, and continuing to question different individuals until they had cross-checked that the information was complete and correct. Village-months after and including July 2019 are considered post-treatment months, i.e., after completion of the intervention; while village-months before (not including) July 2019 are considered pre-treatment months, i.e. before or during intervention. The data provide a listing of each marriage that took place in that month, the age of the bride and groom, and the origin of the bride and groom as either from the village or another village.

4.3 Balance

Table 3 shows descriptive statistics and balance checks for household-level variables,[20] while Online Appendix Table O1 shows the same for a pre-specified list of village-level variables. We report the mean and standard deviation in each experimental arm, the p-value for the test that the difference in means between each combination of experimental arms is zero, and the normalized differences between each combination of experimental arms. Our household- and village-level variables are well balanced across treatments. The p-value on the difference in means for each household-level variable is never statistically significant. Only one out of 72 tests for village-level variables has a p-value of less than 0.10. The p-value of the F-statistic of joint significance is never significant. Normalized differences in means are never above 0.13 for the household-level variables; and mostly below the rule of thumb of 0.25, as suggested by Imbens and Rubin (2015), for the village-level variables.

Table 3 confirms that, by construction, 50% of target households had an adolescent girl surveyed ("girl households") and 50% of target households had an adolescent boy surveyed

[20] Adolescents' age and gender were the only household-level variables pre-specified for balance checks, since there was only a pre-survey listing in Punjab rather than a full baseline survey due to the security situation. In case there were inconsistencies between adolescents' age at midline compared to their age at baseline (Sindh province) or pre-survey listing (Punjab province), age was verified at midline by the enumerator using National Identity Card or Birth Registration Certificate as applicable. For adolescent age, we report balance checks on this verified midline age minus one year. For 32 adolescents we do not have their age at midline, so we use their baseline age instead. Robustness checks using different ways to address inconsistencies in adolescent age are available from the authors upon request. In Punjab, during the midline survey we asked 'retrospective baseline' questions. For non-pre-specified variables in the table, we use the baseline values for Sindh, and the responses to the retrospective baseline questions for Punjab. These variables are primarily presented for descriptive statistics, and not balance checks.

("boy households"). The average age of the surveyed adolescent was 15.3 years, and consistent with our selection criteria, their ages ranged from 13 to 17 years at baseline. 36% of the surveyed adolescents were promised/engaged to be married at baseline, but as per our selection criteria none were married. 56% of adolescents were in school at baseline, with the percentage of girls who are in school (45%) much lower than that for boys (66%).[21] Parents' average years of schooling is low: 4.5 years for fathers and just 1.2 years for mothers. The average age of (first) marriage of the father was 22.4 years, and the average age of the father's (first) spouse at marriage was 18.4 years, with 47% being child-married. Most of parents' marriages (72%) involved a dowry, i.e., a transfer of money or property by the family of the bride to the family of the groom at the time of marriage. Online Appendix Table O1 highlights that in about 64% of the study villages females can leave their compound of residence unaccompanied by a male family member; and on average sample villages have about 170 households. On average we treated 10% of the households in a village, with a minimum share of 0.4% and a maximum share of 40%.

Online Appendix Table O6 presents a comparison of household assets and composition between the target households and the Pakistan Multiple Indicator Cluster Survey (MICS) 2017–2018 (Punjab Bureau of Statistics and United Nations Children's Fund, 2019; Sindh Bureau of Statistics and United Nations Children's Fund, 2021), disaggregated by Sindh and Punjab provinces. Overall, target households are comparable in assets to the rural subsample in the MICS data. Household composition is similar along a range of characteristics of the household head. A notable difference is that mothers in the target households tend to have slightly lower levels of education compared to those in the MICS data.

[21] By the endline survey in the control group, only 23% girls are still in school, with the highest completed education being less than primary school.

Table 3: Household-level descriptives and balance

	(1) C	(2) F	(3) M	(4) F+M	(5) F vs C	(6) M vs C	(7) F+M vs C	(8) F vs M	(9) F vs F+M	(10) M vs F+M	(11) Observations
Adolecent is female	0.510	0.492	0.476	0.505	-0.018	-0.034	-0.005	0.016	-0.013	-0.029	1687
	(0.501)	(0.501)	(0.500)	(0.501)	(0.601)	(0.341)	(0.918)	(0.660)	(0.674)	(0.394)	
					[-0.035]	[-0.068]	[-0.010]	[0.032]	[-0.026]	[-0.058]	
Adolecent's age	15.332	15.448	15.473	15.408	0.116	0.142	0.077	-0.025	0.040	0.065	1687
	(1.170)	(1.243)	(1.251)	(1.173)	(0.212)	(0.107)	(0.457)	(0.774)	(0.655)	(0.456)	
					[0.096]	[0.117]	[0.065]	[-0.020]	[0.033]	[0.055]	
Adolescent is engaged	0.360	0.279	0.324	0.266	-0.081	-0.035	-0.094	-0.045	0.013	0.059	756
	(0.481)	(0.450)	(0.469)	(0.443)	(0.297)	(0.651)	(0.194)	(0.533)	(0.841)	(0.385)	
					[-0.174]	[-0.075]	[-0.204]	[-0.099]	[0.030]	[0.133]	
Adolescent in school	0.561	0.540	0.553	0.584	-0.021	-0.007	0.023	-0.014	-0.044	-0.030	1622
	(0.497)	(0.499)	(0.498)	(0.494)	(0.663)	(0.891)	(0.661)	(0.771)	(0.346)	(0.559)	
					[-0.042]	[-0.015]	[0.047]	[-0.028]	[-0.089]	[-0.062]	
Years of schooling father	4.464	4.835	5.045	5.052	0.371	0.581	0.588	-0.210	-0.217	-0.007	1630
	(4.821)	(4.759)	(5.091)	(5.104)	(0.333)	(0.158)	(0.229)	(0.650)	(0.669)	(0.957)	
					[0.078]	[0.117]	[0.118]	[-0.043]	[-0.044]	[-0.001]	
Years of schooling mother	1.187	1.274	1.647	1.511	0.087	0.460	0.324	-0.373	-0.237	0.136	1644
	(3.192)	(3.000)	(3.506)	(3.394)	(0.700)	(0.105)	(0.249)	(0.190)	(0.389)	(0.755)	
					[0.028]	[0.137]	[0.098]	[-0.114]	[-0.074]	[0.040]	
Marriage age - father	22.386	22.492	22.726	22.054	0.106	0.341	-0.332	-0.235	0.438	0.672	1639
	(4.923)	(5.176)	(5.882)	(5.114)	(0.718)	(0.467)	(0.439)	(0.688)	(0.288)	(0.186)	
					[0.021]	[0.063]	[-0.066]	[-0.042]	[0.085]	[0.131]	
Marriage age - father's spouse	18.363	18.386	18.540	18.297	0.023	0.177	-0.067	-0.153	0.090	0.243	1639
	(3.467)	(3.707)	(4.042)	(3.967)	(0.855)	(0.552)	(0.831)	(0.677)	(0.731)	(0.489)	
					[0.006]	[0.047]	[-0.018]	[-0.040]	[0.023]	[0.061]	
Parents had dowry	0.723	0.733	0.718	0.691	0.010	-0.006	-0.032	0.015	0.042	0.027	1647
	(0.448)	(0.443)	(0.451)	(0.463)	(0.878)	(0.933)	(0.630)	(0.815)	(0.522)	(0.700)	
					[0.021]	[-0.013]	[-0.071]	[0.034]	[0.093]	[0.058]	

Notes: Columns 1-4 show the mean of the variable in each experimental arm – the Female arm (F), the Male arm (M), the Female+Male arm (F+M), and Control arm (C). Standard deviations are indicated in parentheses. Columns 5-10 show the difference in means for each combination of experimental arms. In Columns 5-10, p-values from robust standard errors clustered at the village level (unit of randomization) are indicated in parentheses, using a logit regression for binary variables, and an OLS regression for continuous variables. Normalized differences are reported in square brackets, calculated as the difference between the sample means of experimental arms divided by the square root of the sum of the sample variances. Stars indicate: *** 1 percent ** 5 percent * 10 percent level of significance. "Adolescent is engaged" uses baseline data from Sindh province only, as we did not have a retrospective measure for this in Punjab province. "Marriage age - father" refers to his first marriage, and "Marriage age - father's spouse" refers to his first spouse."Dowry" is a transfer of money or property by the family of the bride to the family of the groom at the time of marriage.

4.4 Attrition

Online Appendix Tables O2 and O3 present attrition rates by treatment arm for midline and endline respectively; for fathers, mothers and adolescents. In Columns 5-10, we test for differential attrition between each experimental arm. In Sindh, where we conducted a baseline survey, we consider a an individual respondent to have attrited if they participated in the baseline survey but they were not available during the midline or endline survey respectively. In Punjab, since we were not able to conduct a full baseline survey due to security concerns, we consider an individual to have attrited if they were randomly selected at baseline from the household listing to participate in the surveys and intervention, and we were not able to survey them during the midline or endline respectively.

The average attrition rates in the control group are low: 5.3% for fathers, 1.2% for mothers, and 0.7% for adolescents at midline; 13% for fathers, 6% for mothers, and 4.6% for adolescents at endline.[22] We also present the breakdown of the number of attrited respondents in Sindh and Punjab at midline and endline in Appendix Tables O4 and O5. We are able to recover detailed information on outcome variables for adolescents from their parents or other household members even when the adolescent is absent, as long as the whole household has not attrited.

4.5 Compliance and participation

Online Appendix Table O7 shows that the gender composition of treatment arms was perfectly complied with. In villages assigned to the Male arm, there were no women that participated in the intervention, and in villages assigned to the Female arm there were no men that participated. We find no evidence that the gender that was targeted had an influence on the number of individuals participating in the intervention. On average, 18 women from target households participated in the Female arm; and on average 16 men from target households participated in the Male arm; and 15 men and 16 women participated in the Female+Male arm. This suggests that it is unlikely that differential participation of men and women across treatment arms will explain our results.

[22] At midline, even though we find that the p-value of one of the 18 tests is significant at the 10% level and one out of the 18 tests is significant at the 5% level, after correcting for the false discovery rate (FDR) across the 6 experimental arm comparisons at the level of each individual, we find that the q-value is not significant. We also note that for the p-values that are significant, the raw difference in the number of attriters is comparatively low: 13 mothers attrited in the Male arm versus 5 in the control arm; 9 adolescents attrited in the Female arm versus 2 in the Male arm.

5 Results

5.1 Child marriages in target households

Our pre-specified estimating equation for testing impacts on child marriage is a logit regression for the binary outcome "child marriage:"

$$Y_{ivs} = \alpha + \eta T_v + \rho T_v \times G_{ivs} + \theta G_{ivs} + \delta_s + \mu_{ivs}, \tag{1}$$

where Y_{ivs} is the outcome variable of interest at midline or endline for adolescent i, in village v and in stratum s. Y_{ivs} takes on value one if the adolescent reports that she was married below age 18 and zero otherwise. T_v is a vector of dummies for the village being assigned to each of our three treatment arms – *Male*, *Female*, or *Female+Male* – relative to the control group. G_{ivs} is a binary variable that takes on value one if the gender of the adolescent child in the household is a girl. The vector of estimated coefficients η, therefore, represent the intention-to-treat (ITT) effects on households with an adolescent boy, and $\eta + \rho$ represents the ITT effects on households with an adolescent girl. δ_s are fixed effects for randomisation strata. Standard errors are robust to village-level heteroskedasticity, as this was the level of randomization. We also report additional p-values for the treatment coefficients as calculated from randomization inference tests (Young, 2019).

An alternative specification for the child marriage data is a discrete approximation of a censored duration model, where we estimate the average annual hazard rate into child marriage for girls aged 13-17 years, taking the method in Corno et al. (2020) as the starting point. Right-censoring occurs in our data because some girls are still unmarried and still under the age of 18 by the midline or endline, whereby they may still experience an unobserved child marriage. The duration we are interested in is, thus, the time after age 13 when they are still unmarried and at risk of being child-married, and the moment when she turns 18 and is no longer at risk of child marriage.

We convert our data into an adolescent-year panel, where each adolescent contributes at most five observations to the sample, one observation for each at-risk year between 13 and 17 until she is either married and exits the data, or passes the 18-year threshold. We, thus, use the individual marital histories up to the midline or endline survey. Right-censored observations — that is, observations where girls are less than 18 and still unmarried by

midline or endline — are coded as missing.[23] Using the adolescent-year panel we then estimate the probability of adolescent i entering marriage at age t, with current age k, in village v and stratum s using the following equation:

$$Y_{iktvs} = \beta T_v + \phi_t + \gamma_k + \delta_s + \epsilon_{iktvs} \tag{2}$$

The dependent variable, Y_{iktvs}, is a binary outcome of interest. Adolescent i will have one of three possible paths for the binary dependent variable, Y_{iktvs}. For an adolescent who is observed to marry between age 13 and 17 at midline or endline, the dependent variable is coded as zero for the ages prior to her marriage and as one for the age at which she marries. When she marries, she exits the data. For an adolescent observed to reach the age of 18 by the midline or endline while still being unmarried, the dependent variable is zero for all of her observations. Finally, for an adolescent still under the age of 18 and unmarried by the midline or endline survey, the observations prior to this censoring point are coded as zero while the post-censoring observations are coded as missing. Other variables are specified as in Equation (1). The vector β is our coefficient of interest and represents the ITT effect of our interventions on the average annual hazard rate of child marriage. ϕ_t is a vector of age fixed effects accounting for the different probability of marriage at at-risk years $\{13, 14..., 17\}$, and γ_k are fixed effects for girls' current age to account for cohort effects. Thus, our identifying variation comes from within-at-risk-age and within-current-age variation in treatment arms and marriage outcomes. We estimate Equation (2) using OLS. Standard errors are again robust to village-level heteroskedasticity.

The main difference between the specifications is how censored data are accounted for. The child-marriage outcomes in the OLS and pre-speficied logit regression in Equation (1) codes all girls who are under 18 and not married at the time of the survey as unmarried, thereby not accounting for the continued risk of child marriage in their lifetime for these adolescents that the hazard regression in Equation (2) accounts for.

Table 4 presents the estimated ITT effects on the probability of child marriage for girl adolescents at midline and endline. Columns (1) and (2) present the estimated effects from OLS regressions and Columns (3) and (4) from the pre-specified logit regressions in

[23]Note that we observe more completed marital histories at endline than at midline, implying less censoring in the endline analysis. We are, however, interested in short-term and longer-term impacts of our intervention, and therefore conduct the analysis at both midline and endline.

Equation (1). By endline we observe completed child marriage histories for 71% of the sample of adolescent girls. We see that 6% and 12% of girls are child-married and 11% and 22% are married in the control group at midline and endline respectively. By endline, 63% of girls that are unmarried have moved beyond the 18-year threshold, and are therefore no longer at risk of being child-married. In Columns (5) and (6) we present estimated effects on the annual hazard rate into child marriage from estimating the censored duration model in Equation (2) using OLS. The average annual hazard rate of child marriage for girls in the control group is 1.9 percent per year at midline and 3.0 percent per year at endline.

We find that targeting the intervention at women only does not lead to significant impacts on the probability of child marriage for girl adolescents, either in the short- or the long-run, when compared to the control group. We show that targeting the intervention at men only significantly reduces the probability of child marriage for girl adolescents in the short- and in the long-run. At midline, the reduction is 4.1 percentage points, a 66% reduction relative to the control group mean of 6.2 percent. The reduction is similar and significant in the OLS, the pre-specified logit specifications and the hazard model (p-value < 0.05). At endline, 1.5 years after the intervention, the reduction is 5.3 percentage points, relative to a control group mean of 12.3% impliying a 43% reduction (p-value < 0.10). The reduction in the average annual hazard rate is similar (p-value < 0.05). When women and men are jointly treated, there is a 5.2 percentage points reduction in the likelihood of child marriage for girl adolescents at endline, implying a similar 42% reduction (p-value < 0.10). The reduction in the average annual hazard rate is also similar (p-value < 0.10). At midline, however, we do not find a significant reduction in the probability of child marriage in this arm. In the Male arm and Female+Male arm, respectively, relative to the Control group, the effect is not significantly different over time – midline versus endline. Comparing the effect sizes between treatment arms, we find that at midline the reduction in the probability of child marriage is significantly larger in the Male arm as compared to the Female+Male arm but only in the hazard model specification (p-value < 0.10), and also significantly larger as compared to the Female arm in both specifications (p-value < 0.05). At endline, we cannot reject the null that the estimates in all three treatment arms are the same. Our results are robust to including all balance variables as well as parent's education level as controls.

Our results are robust to using alternative assumptions for estimating the hazard rate model. Column (1) and Column (2) in Appendix Table O8 present results where the dura-

tion of interest begins from the individual's baseline age, which is just before the intervention was implemented. Column (3) and (4) present results where censored observations that are still at risk of being child-married are coded as the cohort-specific, age-specific, and treatment arm-specific probability of marriage between their current age until they reach the 18-year threshold, thereby accounting for the continued risk of girl child marriage.

Appendix Table O9 shows the estimates from our pre-specified model in Equation (1) for the additional pre-specified outcome – "marriage." The control group mean in Column (1) shows that 11% and 22% of girls are married at midline and endline. We observe significant reductions in marriages (not necessarily child marriages) in the Male arm at midline and endline, and in the Female+Male arm at endline. We can not reject the null hypothesis that estimates between each treatment arm are the same.[24] The average age of marriage is 16.6 years at midline and 17.0 years at endline in the control group.

Robustness of marriage reporting: We conduct extensive checks on reporting of marriage outcomes. Our main specifications in Table 4 and Appendix Table O10 use the reports from the adolescents in the households, and thus not the report from the treated parent. This is common practice in the child marriage literature, as adolescents are assumed to have weaker incentives to misreport. Nevertheless, we triangulate responses from the mother, father and the adolescent in the household and find a rate of consistency above 95%. As presented in Online Appendix Table O10, our results on marriage are robust to using reports by the father, the mother, or any of the three respondents in the household. Furthermore, in our context in Pakistan the dominant practice is patrilocality, i.e., girls move out of the household after marriage. We conducted a verification exercise of the reported marriage status by observing where the girl resides during our midline and endline. We are able to trace 87% of the girls who are reported as being married at endline, Appendix Table O9 shows impacts on marriages are consistent in the verified sample.[25] We also show that in kernel density plots of age of marriage in the control versus treatment arms (Online Appendix Figure O1) we do not see bunching at the 18 year threshold, which is the legal

[24]For adolescent boys, we find no impact of the intervention on marriage or child marriage, at midline or endline (Online Appendix Table O11). Boy adolescents aged 13-17 are much less likely to get married than girls: just 3% of boys in the Control arm are married at midline, and 7% at endline. Boys are also much less likely to be child-married: 1% and 4% of boys in the Control arm at midline and endline respectively, compared to 7% and 12% of girls at midline and endline respectively.

[25]This verification was made possible by the fact that 90% of marriages happen between spouses from the same community

age of marriage. The modal age of marriage remains below 18. These patterns suggest that our results are not driven by demand effects wherein treated individuals begin reporting the legal age of marriage after receiving the intervention.

Table 4: Target households: Child marriage outcomes for adolescent girls

	OLS		Pre-specified Logit		Annual Hazard Rate	
	(1)	(2)	(3)	(4)	(5)	(6)
	Midline	Endline	Midline	Endline	Midline	Endline
Female	0.013	-0.025	0.013	-0.025	0.003	-0.006
	(0.672)	(0.492)	(0.671)	(0.493)	(0.671)	(0.542)
	[0.709]	[0.544]	[0.705]	[0.550]	[0.723]	[0.567]
Male	-0.041	-0.053	-0.041	-0.053	-0.013	-0.015
	(0.029)**	(0.085)*	(0.027)**	(0.083)*	(0.016)**	(0.038)**
	[0.048]**	[0.124]	[0.048]**	[0.131]	[0.012]**	[0.065]*
Female+Male	-0.015	-0.051	-0.015	-0.052	-0.005	-0.013
	(0.500)	(0.086)*	(0.498)	(0.081)*	(0.447)	(0.077)*
	[0.490]	[0.094]*	[0.488]	[0.092]*	[0.504]	[0.087]*
Observations	828	798	828	798	2755	3239
Control Mean	0.062	0.123	0.062	0.123	0.019	0.030
p-val M $\neq$ F	0.058	0.412	0.056	0.407	0.023	0.263
p-val FM $\neq$ F	0.363	0.433	0.361	0.424	0.318	0.383
p-val FM $\neq$ M	0.140	0.945	0.138	0.948	0.099	0.728
p-val F $<$ C	0.664	0.246	0.665	0.247	0.992	0.271
p-val M $<$ C	0.014	0.042	0.014	0.042	0.008	0.019
p-val FM $<$ C	0.250	0.042	0.249	0.041	0.223	0.039
p-val M $<$ F	0.028	0.206	0.028	0.204	0.011	0.131
p-val M $<$ FM	0.069	0.473	0.069	0.474	0.049	0.364
p-val FM $<$ F	0.181	0.216	0.180	0.212	0.159	0.192

Note: The table presents treatment effects on the probability of child marriage at midline and endline for girl adolescents from our target households. Marriage outcomes are as reported by the adolescent themselves as pre-specified in the pre-analysis plan. Columns 1-2 report marginal treatment effects for girl adolescents ($\eta + \rho$) from the pre-specified logit regression in Equation 1 on child marriage, i.e, a binary variable that takes on value one if the adolescent was less than 18 years old at the time of marriage, for midline and endline respectively. Fixed effects for randomisation strata are included at endline and not at midline due to perfect prediction. Columns 3-4 report the average hazard rate into child marriage, using a discrete approximation of a censored duration model, estimated from Equation 2. Data is an adolescent-year panel, where each adolescent contributes at most five observations to the sample, one observation for each at-risk year between 13 and 17 until she is either married or passes the 18-year threshold after which she exits the data. The dependent variable is a binary variable that is coded as zero for the ages prior to her marriage, and as unity for the age at which she marries. Right censored observations – that is, observations between the current age at the corresponding survey-round and age 18 for girls who are aged less than 18 and still unmarried by midline or endline – are coded as missing. Fixed effects for each age at-risk of marriage, current age and randomisation strata are included. *P*-values for marginal treatment effects from standard errors clustered at village level (unit of randomisation) are indicated in parentheses. Exact *p*-values for marginal treatment effects from randomization inference tests based on 1000 permutations are provided in square brackets. Stars indicate: *** 1 percent ** 5 percent * 10 percent level of significance. The row "Control Mean" indicates the probability of child marriage (Columns 1-2) and the average annual probability of child marriage (Columns 3-4) in the control group. *P*-values for all comparisons of treatment effects between each experimental arm as well as one-sided tests are reported at the bottom of the table: Male arm (M), Female arm (F) and Female+Male arm (FM).

Finally, we consider the effect on engagements, as households may misreport marriages as an engagement. This could be due to the fact that while marriage under 18 is illegal, engagements are not. Reassuringly, we show that for girl adolescents who are unmarried at midline and endline, there is no significant increase in the likelihood of being engaged or newly engaged between survey rounds (Online Appendix Table O12).[26]

5.2 Child marriages at the village level

We use our village-level observational data to estimate impacts on all marriages in our sample villages. We estimate a village-fixed-effect regression at the level of marriages of the following form:

$$Y_{mvst} = \tau + \kappa T_v + \psi T_v \times post_t + \lambda post_t + \delta_s + \delta_s \times post_t + \xi_{mvst} \tag{3}$$

where Y_{mvst} is the outcome variable of interest at midline or endline for marriage m in village v in stratum s at village-month t. T_v is a vector of dummies for the village being assigned to each of our treatment arms — *Male*, *Female*, or *Female+Male* — relative to the control group. $post_t$ is a binary variable that indicates that a marriage falls in the period after completion of the intervention, i.e. after and including July 2019. ψ therefore combines the treatment effect of the intervention on treated households in the village, as well as spillover effects on other households in the village. We estimate ψ for the period up to midline (January 2020) and endline (March 2021). δ_s are strata fixed effects, and $\delta_s \times post_t$ is included for inference since randomization was blocked on strata (Bruhn and McKenzie, 2009). Standard errors are robust to village-level heteroskedasticity, as this was the level of randomization (Abadie et al., 2023).

In Table 5, the variable "child marriage" is a binary variable that takes on value one if marriage m in village v involved a bride below the age of 18, and zero if the observed marriage involved a bride 18 and above. The continuous variable "age of marriage" corresponds to the age of the bride at the time of marriage.

We see that 61% of marriages at midline and 35% of marriages at endline in the control

[26] We focus on early marriage instead of early engagement, since most of the documented severe negative consequences to girls and to future children arise from practices that take place only after the actual ritual of marriage, such as conjugal living, the girl leaving her birth family's household, and consummation. We note that in this context, girls and boys can be promised or engaged by their families at a very early age or even at birth, while they continue to live with their birth family until the time of marriage.

group are child marriages. The average age of marriage for adolescent girls at midline is 17.4 and at endline is 18.4 in the control group.[27] Column 1 shows that, at midline, we find a significant reduction in child marriages of girls at the village level across all treatment arms. At endline (Column 2), we see that reductions only remain significant at endline when women are treated alone or jointly, i.e. in the Female and Female+Male arm. We estimate a 21 percentage points (p-value < 0.01) reduction in the likelihood that at least one girl is married below 18 in each month in the Female arm. In the Female+Male treatment arm, the effect is 25 percentage points (p-value < 0.01). For both arms this reduced likelihood is more than a 60% reduction in child marriages as compared to the control mean. The effects at endline in the Female arm are significantly smaller than at midline (p-value < 0.05) and are not statistically different across midline and endline in the Female+Male arm. At endline, we find no significant impacts on girl child marriage when the intervention is targeted at men only. Between midline and endline the estimated effects in the Male arm are significantly different from each other (p-value < 0.01). The evidence, thus, shows that reductions in child marriage in the Male arm fade out over time, but are sustained in the Female and Female+Male arm. At endline the reduction in the Male arm is significantly smaller than the reduction in the Female+Male arm (p-value < 0.05). We see qualitatively similar effects if we consider effects on the age of marriage (Columns 3 and 4).

Village-level effects cannot be solely driven by the target household marriages (N=188), as the number of marriages observed at the village level (N=1383) substantially exceeds this. Our results are robust to converting our data into monthly summary statistics at village-level. We find qualitatively similar effects on the likelihood of child marriage, average age of marriage, and minimum age of marriage in village-months that have a marriage (Appendix Table O13). We also show in Online Appendix Table O14 that our results are robust to a specification where months without child marriages are coded as zero and a count variable of the number of child marriages in a month, where again months without marriages are zero.

[27] We attribute this downward trend to the fact that 1) the government of Pakistan instituted the Child Marriage Restraint (Amendment) Bill 2018, which was passed by the senate in 2019, which increased the minumum age of marriage for females to 18; 2) COVID-19 happened between the midline survey and endline survey, which led to lockdowns happening between March and May 2020, covering an important part of the wedding season. The lockdowns implied that large gatherings were banned, marriage halls and event venues were closed and outdoor weddings were only allowed with limited attendance. Given that our endline survey ran from September 2020 to March 2021 this declining trend may be particularly strong over the period the endline covers.

Table 5: Village-marriage level: (Child) marriage outcomes for girls

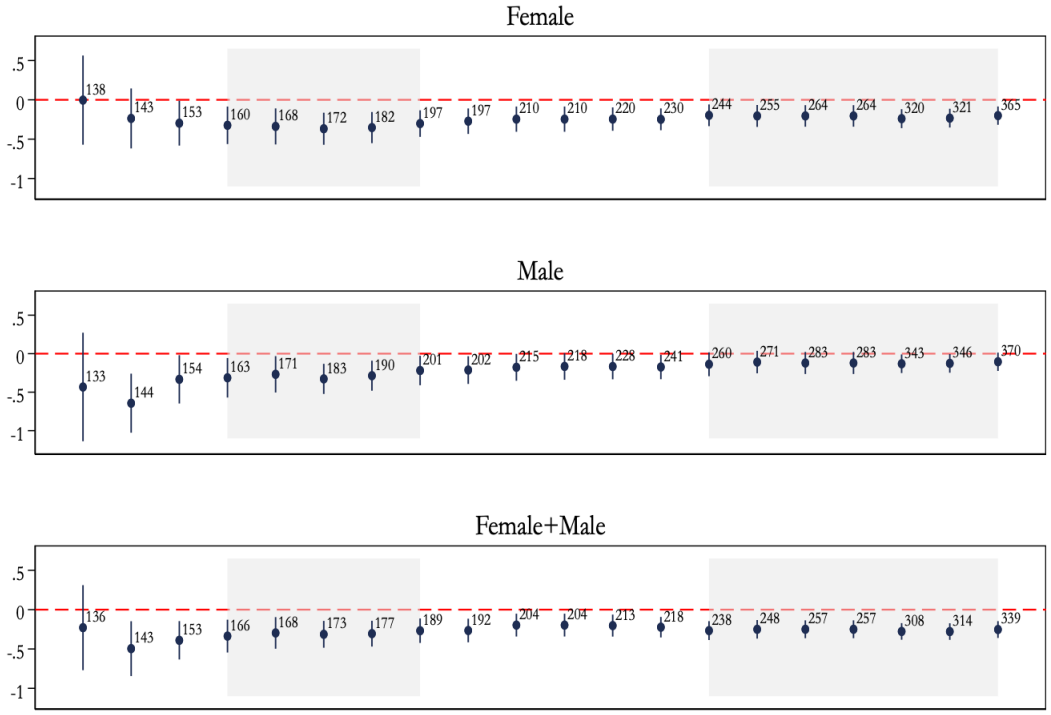
	Child Marriage		Marriage Age	
	Midline	Endline	Midline	Endline
	(1)	(2)	(3)	(4)
Post X Female	-0.407	-0.206	1.293	0.676
	(0.002)***	(0.005)***	(0.027)**	(0.050)*
Post X Male	-0.337	-0.107	0.790	0.350
	(0.004)***	(0.140)	(0.089)*	(0.311)
Post X Female+Male	-0.262	-0.251	0.678	0.949
	(0.025)**	(0.000)***	(0.182)	(0.003)***
Observations	697	1383	697	1383
Control Mean	0.614	0.348	17.386	18.429
p-val M $\neq$ F	0.608	0.193	0.368	0.380
p-val FM $\neq$ F	0.269	0.527	0.271	0.430
p-val FM $\neq$ M	0.530	0.044	0.808	0.085
p-val M $>$ F	0.304	0.096		
p-val M $>$ FM	0.735	0.022		

Notes: The table presents estimates for girls of ψ from Equation 3 using village fixed effects regressions with data at the level of marriages. In Columns 1 and 2, the binary dependent variable takes value one if the marriage involved a bride below the age of 18. In Columns 3 and 4, the dependent variable is the age of marriage of the bride. "Midline" counts observations between the pre-treatment months and January 2020, and "Endline" counts observations between the pre-treatment months and March 2021. The variable "post" is a binary variable that takes on the value zero if the marriage lies in the period before (not including) July 2019, the period before or during the intervention; and takes on value one if the marriage lies in the period after the intervention was completed. Fixed effects for randomisation strata and their interaction with "post" are included. *P*-values from standard errors clustered at village level (unit of randomisation), are indicated in parentheses. Stars indicate: *** 1 percent ** 5 percent * 10 percent level of significance. The row "Control Mean" reports the likelihood of child marriage and average marriage age of girls during the pre-intervention period in the control group. *P*-values for all comparisons of treatment effects between each experimental arm as well as one-sided tests are reported at the bottom of the table: Male arm (M), Female arm (F) and Female+Male arm (FM).

The above results are also visually illustrated in Figure 1. The figure decomposes the effect over time by plotting cumulative marriages in each of our post-intervention study months. Immediately after the intervention we observe significant reductions in child marriages in all arms. These initial reductions are, however, only sustained over time in the Female and Female+Male arm. The 95% confidence intervals for the estimated coefficient in each village-month are always below the horizontal red line (zero threshold). We also see that the impact fades out for the Male arm over time, with confidence intervals touching the horizontal red line soon after the intervention.

Finally, the impacts observed at village level are also consistent with effects on marriage proposals received by girl adolescents in target households as measured in the household panel survey. Online Appendix Table O15 shows a significant reduction in marriage proposals for all girls in the Female+Male arm, as compared to the Control arm; and a significant reduction for girls below the age of 18 in the Female and Female+Male arm, as compared to the Control. Effects are, however, not significantly different between treatment arms. It seems unlikely that effects are entirely driven by proposals from target households: first, in most villages, our households represent only a small fraction of households in the village with adolescents of marriageable age; and second, proposals normally come from boys older than the age of our target adolescent girls. It therefore seems likely that the reduction in proposals to the girls in our target households reflects a shift in behaviour by other households in the village; consistent with the village-level effects observed above.

Figure 1: Village-marriage level: Cumulative child marriage impacts by study month



Notes: Cumulative treatment effects on child marriage for girls using village fixed effects regressions at the level of marriages estimated at each post-intervention observation month. Fixed effects for randomisation strata and their interaction with "post" are included. The binary dependent variable takes value one if the marriage involved a bride below the age of 18, and zero otherwise. The marker represents the point estimate in that month, and the vertical lines above and below the marker the 95% confidence interval based on standard errors clustered at village level (unit of randomisation). The number above the marker indicates the cumulative number of marriages upto that month in the respective treatment arm. The red horizontal line represents the zero threshold. The shaded region on the left corresponds our midline household survey months (November 2019 until March 2020) and the shaded region on the right corresponds to our endline household survey months (September 2020 and March 2021). The panel presents the treatment effects over time – on top is the Female arm versus the Control, in the middle is the Male arm versus the control arm, and on the bottom is the Female+Male arm versus the control arm.

6 Conceptual Framework

In this section, we first show how a model of intra-household communication and decision-making based on Bayesian persuasion can rationalize our household-level results. This framework, which models spouses who make important household decisions, is formally

developed in this paper's companion paper, Anderberg et al. (2024) and applied to the decision of when to marry a daughter (see Section B.1 to B.4 in Appendix B for a formal presentation). Second, to explain differences in child marriage outcomes across treatment arms in both target households and at the village level, we extend the model by assuming gender-differential information transmission from targeted spouses to others in the community and by modeling village equilibrium outcomes.

6.1 Household decision in target households

In the model in Anderberg et al. (2024), fathers and mothers make a joint household decision to either delay their daughter's marriage or marry their daughter early. In their decision, they trade off potential returns from delaying marriage, for example in terms of improved health of the daughter or increased quality of the daughter's future spouse, with potential costs (caused by stigma) of deviating from a community age-of-marriage norm that favors early marriage. The probability that the return to delay materializes can be either high or low, dependent on an unknown state of the world, over which the parents have a common prior belief. The probability of stigma is increasing in the deviation from the community norm.

While parents both value the potential return equally, they may update their belief about its likelihood individually if they receive new information, for example from an intervention. Furthermore, in line with the stylized facts in Table 2 and prior literature on social norms – Blackwell (2000); Eagly and Carli (1981); Grasmick et al. (1993); Kuang et al. (2020); Nolen-Hoeksema (2004); Steinberg and Monahan (2007); Tibbetts and Herz (1996), we assume that fathers are less stigma-averse then mothers, and thus place less weight on sanctions from deviating from the norm than mothers do. Parents trade off these returns and costs. They will prefer to delay marriage if the expected returns to delay, which are conditional on the beliefs about the likelihood of the return materializing, outweigh the expected cost due to stigma aversion from deviation from the community norm. Spouses may disagree about the preferred timing of marriage. In that case, fathers prefer delaying marriage and mothers prefer to marry their daughter early, and the decision that is implemented depends on the respective spouses' bargaining weights. This set-up thus implies that the probability of delay is highest when the probability of a return is a high, lowest when the probability of the return is low, with an intermediate outcome for the probability of return based on the prior. It also implies that fathers are more likely to want to delay marriage than mothers.

Suppose now that one spouse receives new information about the likelihood of a return from delay, for example from an intervention. They can then send a signal to their spouse that there is a high or a low likelihood of the return, which may or may not be a truthful signal. Based on this communication, the spouse may update their beliefs about the likelihood of return using Bayes' rule, anticipating that their spouse's signal is potentially untruthful. If the informed spouse is the father (mother), *and* they always wants to delay marriage (marry early) irrespective of the likelihood of the return, but the mothers (fathers) wants to delay (marry early) only with a high (low) probability of the return – i.e., there is a likelihood that the parents will disagree after communication – the mother (father) is *persuadable* and the father (mother) may engage in *Bayesian persuasion*. In this case it is privately optimal for the father (mother) to communicate that there is a high (low) probability of the return, with a probability just high enough to make the mother (father) prefer to delay (marry early). For any other situation, truthful communication about the actual probability is optimal for the father (mother). In contrast, if spouses are jointly informed, neither is provided with an informational advantage and no persuasion is possible.

Anderberg et al. (2024) use this model to consider what effect information and communication have on the household decision to delay a daughter's marriage. Given that the edutainment intervention provided information on the benefits of delaying marriage in order to reduce child marriage, we focus on the case where the newly received information is that the likelihood of a return to delay is high. When this information is provided to fathers or both spouses jointly, it will increase the likelihood of households delaying marriage compared to the prior. These effects should both be larger than when mothers are targeted alone. The effect of providing information to mothers alone is ambiguous compared to the prior.

So to summarize, in the *short-run*, when community age-of-marriage norms are fixed, the model predicts the highest marriage delay when fathers are informed alone or jointly, compared to the prior and compared to when mothers are informed alone. In the *long-run*, however, the community age-of-marriage norm can adjust, and the stigma associated with delaying marriage might reduce or disappear. In that case the likelihood of marriage delay should be higher compared to the prior, and similar when parents are informed alone or jointly.

6.2 Empirical results – predictions for target households

Our research design exogenously varies who within a target household receives information on the true likelihood of the benefit from delaying marriage: no-one (Control arm, decision on the prior), fathers only (Male arm), mothers only (Female arm), or both father and mothers jointly (Female+Male arm).

Effect on belief updating on benefits from delay. First, we empirically test whether parents update their beliefs on *health benefits* as a result of our intervention. Online Appendix Tables O17 and O18 show that there is no consistent pattern of updating on beliefs about health costs, both mental and physical, to the child or grandchild as a result of our treatments. At baseline, however, both fathers and mothers already seem aware of the most extreme risks of early childbearing – namely the risk of death to young mothers and to their children – suggesting limited scope for interventions to make such costs even more salient. Therefore, we argue that our intervention did not provide them with *new* information about health benefits of delaying marriage. With respect to *spousal quality benefits* of delaying marriage, Anderberg et al. (2024) show that parents update their beliefs in response to treatment in all arms. This suggests that the intervention provided *new* information about these benefits.[28] Table O19 shows impacts of the intervention on these beliefs if a girl's marriage is delayed until she is 18.[29] We measure spousal quality returns to delaying marriage of a girl adolescent as the expectation that the spouse will have completed secondary (grade 9 or 10) or high school education (grade 11 or 12), conditional on the adolescent girl being married at age 18.It is important to mention that the share of fathers and mothers that expect a future spouse with secondary or high school education increases with the conditional age of marriage of the daughter, suggesting a positive correlation in expectations about the age of marriage of a girl, and the quality of the spouse in terms of his education level. We observe that targeted parents in all treatment arms update their beliefs and expect a higher education level of the future groom of their adolescent daughter if they were to delay their

[28] In the literature, the education of grooms is seen as one of the most meaningful indicators of spousal quality that can be readily collected from parents in a short survey. The groom's education as a dimension of the quality of a marriage match for girls is also used in Buchmann et al. (2023).

[29] These relationships between marriage age and the expectation of the education level of the future spouse are merely correlations, and we can not distinguish whether delaying marriage causes matching with spouses with better education level, or whether spouses with better education level prefer delayed marriages.

daughter's marriage.

Effect on marriage delay. We observe updating of beliefs towards a high likelihood of spousal quality returns from delaying marriage in all arms. Consistent with this we find significant reductions in child marriage in the Male arm at midline and endline, and in the Female+Male arm at endline. As predicted, we also observe in the relevant one-sided test at midline, that reductions in child marriage in the Male arm and the Female+Male arm are larger than reductions in the Female arm. At endline there are, however, no significant differences across the arms, consistent with our prediction that sustained reductions in child marriages in the Female+Male and the Female arm at village level imply a lower of the cost of stigma due to a reduction in child marriages at the village-level, and a weakening of the community age-of-marriage norm, for which we provide evidence below.

6.3 Model of village-level equilibrium

The model in Anderberg et al. (2024) focuses only on the household, and does not consider village-level impacts. However, in this paper we have shown that the reductions in child marriage at village level imply the existence of spillovers from households targeted by the edutainment intervention to other households in the village.[30] Our village-level results demonstrate a reduction in the probability of child marriage in villages where women are treated, either alone or jointly, in the long-run, 18 months after our intervention. In the Male arm, effects fade out over time. This is in contrast to the patterns observed for target households, where reductions in the likelihood of child marriage occur when fathers are treated alone or jointly. If spillovers had simply followed the treatment effects in target households we should have observed the same patterns of results at village level. This is not the case, and the diverging pattern of results in target households versus the village implies that simple household-level spillovers are unlikely to be the mechanism driving village results.

Instead, these patterns suggest a mechanism of spillovers that depends on the gender of the spouse that is informed through the intervention. We already presented descriptives in Section 4 that suggest that female respondents participate more in community meetings than males. We also asked respondents that participated in the intervention who they talked

[30]In our context, mobility is limited across villages and 90% of marriages happen between brides and grooms from the same village. Thus, the spread of information is likely to be contained within villages.

to about the intervention. 34% of mothers that participated in the intervention report that they discussed the intervention with relatives and neighbours that are not in their household, as compared to 25% of fathers, and this difference is statistically significant. Consistent with these statistics, we assume gender-differential information transmission from the informed spouse to other members in the village. To further corroborate this assumption, we also provide evidence of treatment heterogeneity, that shows that the village-level impacts in the Female and Female+Male arm are driven by villages where the women in our target households have "high agency" in their communities. We proxy this by their ability to leave their compounds of residence, their education and their attendance in community meetings. Online Appendix Table O16 shows that for all three proxies, village-level reductions appear driven by villages where directly targeted women have "high agency." This supports an interpretation where women play an active role in transmitting information about the intervention, including *new* information about benefits from delaying marrying in the community.

Therefore, we extend the model of household decisions by assuming that parents who receive new information from the intervention transmit this information to other individuals in the community through gender-segregated networks.[31] We also assume that mothers – given their relatively stronger hesitancy to deviate from age-of-marriage-norms – experience stronger incentives to transmit information about potential benefits of marriage delay to other women in the community.

A standard feature of transmission processes – applicable both to information transmission and transmission of infections – is that if anyone who receives the information passes it on to *more than one person* on average, then the information will eventually reach the full population. In contrast, if anyone who receives the information passes it on to *less than one person* on average the transmission process will die out.

Gender difference in transmission of information. Our model rationalizes the observed results as follows. We can choose the rate of information transmission such that

[31]The assumption of gender-segregated networks is highly plausible in our context. Women or men rarely engage or are allowed to engage in conversations with individuals of the other gender who are not their spouse, children or close relatives. Public programs, such as by health extension workers, are also typically organized for men and women separately. Similarly, in our Female+Male arm since both genders could not be in the public space jointly, we organized separate simultaneous sessions for men and women.

(i) any father who receives information that the likelihood of a return from marriage delay is high, passes this information on to fewer than one other father on average, and (ii) any mother who receives information that the likelihood of a benefit is high passes this on to more than one other mother on average. Then the transmission of the information will fade out in the Male arm, and the probability of marriage delay at the village level will revert to the equilibrium based on the prior. In contrast, the information transmission will eventually reach all households in the Female arm and the Male+Female arm, and the probability of marriage delay at the village level will revert to a new equilibrium, where the likelihood of a benefit is high.

As a final step, we compare the likelihood of marriage delay for the equilibrium without new information and the equilibrium where mothers are targeted and the new information about benefits to delaying marriage reaches all households in the community (equivalent to the Female and Female+Male arm). We formalize this in Section B.6 in Appendix B. We find that for all but one type of household, the predicted delay frequency is either the same or lower when the community has no new information about these benefits, as compared to the delay frequency with new information disseminated by mothers, making it plausible that marriage delay is observed, consistent with our village-level results in arms where women are treated.

Hence, gender differences in information transmission are consistent with the following patterns at village level: (i) marriage delay decisions in the long-run reverting back to the pre-intervention level in the Male arm, (ii) marriage delay decisions converging to a new equilibrium with more frequent marriage delays in the long-run in the Female arm and the Male+Female arm.

6.4 Empirical results – predictions at village level

To provide additional empirical support for the mechanisms that generate our village-level results in our model, we leverage the dynamics of our empirical results over time. These show that when only men are treated (Male arm) the reductions in child marriage at village-level fade out, while when women are treated (Female arm and Female+Male arm) reductions in child marriage at village-level are sustained. In arms with sustained village-level effects, we observe that the likelihood of child marriage delay in target households becomes stronger over time. In our model, the lower willingness of mothers to deviate from age-of-marriage norms due to higher stigma aversion is, ceteris paribus, the driver of differences in predictions

of child marriage delay in target households across treatment arms. Because child marriage becomes less likely at the village-level in arms where mothers are treated and mothers are more stigma averse, especially mothers become more willing to delay marriage.

It is, therefore, natural to consider any changes in the beliefs of mothers about age-of-marriage norms in the village across treatment arms. Table 6 shows that mothers in the Female and Female+Male arm consistently update their beliefs about the attitudes of other men and women in the community. It is important to note that we do not find a consistent pattern of updating by fathers regarding the beliefs about the attitudes of other men and other women in the community in any treatment arm. Specifically, mothers are less likely to believe that other community members find early marriage desirable or acceptable, consistent with the changes in village-level reductions in child marriages as observed in Figure 1.[32]

[32]In Section 2 we refer to the components of a social norm, as defined by Bicchieri (2006), and village-level changes in the likelihood of child marriage (empirical expectation) as well as beliefs of mothers' about community members' beliefs about the acceptable age of marriage (contingency) are two of the five components of a norm, supporting our assumption that indeed norms shift. The pattern of updating is not driven by a mechanisms whereby our intervention corrects a misperceived social norm, that we measure following the method used by Bursztyn et al. (2023, 2020), by facilitating a platform to share information about attitudes (Online Appendix Table O20). Instead, it seems likely that mothers and fathers update their second-order beliefs in line with real changes in behaviour that they observe at the village level.

Table 6: Mothers' beliefs about attitudes of **other men and women** in the community towards child marriage for girls

	Best Age < 14		Best Age 14-15		Best Age 16-17		Accept 12-15		Accept 16-17	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Female	-0.255	-0.388	-0.613	-0.753	-0.247	-0.409	-0.646	-0.591	-0.102	-0.461
	(0.205)	(0.090)*	(0.005)***	(0.000)***	(0.253)	(0.053)*	(0.003)***	(0.010)***	(0.613)	(0.040)**
	[0.253]	[0.090]*	[0.016]**	[0.001]***	[0.253]	[0.080]*	[0.007]***	[0.020]**	[0.614]	[0.041]**
Male	-0.185	-0.220	-0.228	-0.046	-0.135	-0.074	-0.302	-0.279	-0.097	-0.072
	(0.283)	(0.262)	(0.323)	(0.839)	(0.531)	(0.739)	(0.157)	(0.206)	(0.656)	(0.740)
	[0.484]	[0.787]	[0.484]	[0.840]	[0.532]	[0.840]	[0.315]	[0.413]	[0.656]	[0.741]
Female+Male	-0.350	-0.408	-0.488	-0.376	-0.217	-0.192	-0.592	-0.478	-0.221	-0.162
	(0.039)**	(0.027)**	(0.031)**	(0.084)*	(0.294)	(0.392)	(0.005)***	(0.032)**	(0.273)	(0.471)
	[0.059]*	[0.083]*	[0.059]*	[0.127]	[0.295]	[0.392]	[0.011]**	[0.065]*	[0.274]	[0.471]
Observations	1649	1650	1649	1650	1649	1650	1649	1649	1648	1649
Control Mean	2.034	2.179	3.618	3.598	5.037	5.152	3.613	3.556	5.098	5.279
M ≠ F	0.706	0.448	0.065	0.003	0.619	0.135	0.111	0.189	0.983	0.076
M ≠ FM	0.275	0.285	0.237	0.189	0.710	0.621	0.171	0.394	0.561	0.684
F ≠ FM	0.606	0.927	0.549	0.103	0.893	0.336	0.803	0.642	0.546	0.190
M > F	0.353	0.224	0.032	0.002	0.310	0.067	0.056	0.095	0.491	0.038
M > FM	0.138	0.143	0.119	0.094	0.355	0.31	0.086	0.197	0.281	0.342
FM < F	0.303	0.463	0.726	0.949	0.554	0.832	0.598	0.679	0.273	0.905

Notes: The table presents intent-to-treat effects from OLS regressions at midline. The dependent variable is the mother's belief about the number of other men out of 10 in their community and other women out of 10 in their community who find less than 14 the best age (Columns 1-2), 14-15 the best age (Columns 3-4), 16-17 the best age (Columns 5-6), 12-15 an acceptable age (Columns 7-8), and 16-17 an acceptable age (Columns 9 and 10) to marry off a girl. It takes a value from 0-10. Fixed effects for randomisation strata are included. *P*-values from standard errors clustered at village level (unit of randomisation) are indicated in parentheses. *Q*-values correcting for false discovery rate within each family using the Benjamini–Hochberg procedure are indicated in square brackets. The family over which we correct reflects the three categories of best age per respondent per treatment arm and the two categories of acceptability per respondent per treatment arm. Stars indicate: *** 1 percent ** 5 percent * 10 percent level of significance. The row "Control Mean" multiplied by 10 indicates the average percentage of other men and women in the community that mothers in the control group believe would agree or strongly agree with the specific statement. *P*-values for comparison of treatment effects between experimental arms for various one-sided and two-sided alternative hypothesis are reported at the bottom of the table: Male arm (M), Female arm (F) and Female+Male arm (FM).

7 Conclusion

Norms that prescribe behaviour of women and girls exist in many contexts. Even though private benefits from deviating from these norms may be largest for women, if concerns about social repercussions outweigh these private benefits, women may be more reluctant to engage in the behaviour that deviates from the norms. In a context of early age-of-marriage norms for girls, we investigate to what extent targeting men, women, or both with the same *new* information about potential benefits from delaying marriage, affects outcomes for girls in target households and at the village level.

We show that targeting the intervention at men, either alone or jointly with women, reduces the probability of child marriage for girls in *target households*. We find no significant reductions in target household when women alone are targeted. Consistent with our evidence, these results are rationalized by the relatively stronger hesitancy of mothers to deviate from early age-of-marriage norms. Fathers, being less hesitant to deviate from norms – when they receive new information about potential benefits from delaying marriage – persuade mothers to delay marriage, while such persuasion is not possible when fathers and mothers jointly receive new information.

In contrast, targeting the intervention at women, either alone or jointly, leads to sustained reductions in the probability of child marriage for girls *at village level*. However, short-term reductions fade out over time when targeting men alone. We rationalize these results by extending the model, incorporating gender-segregated information transmission from the targeted spouses to other households in the village, where women have a stronger incentive to communicate potential benefits of marriage delay in the community. We demonstrate that our results are consistent with an interpretation where a new equilibrium of delayed marriage is established at village level when women are targeted.

Our findings indicate that targeting women to change outcomes for women and girls may lead to seemingly counter-intuitive outcomes in their own households versus at village level. We show, however, that these results are generated by the same mechanism whereby women – in their own households – may be reluctant to change outcomes because of costs of social repercussions. Those same costs may, however, lead them to experience incentives to communicate about benefits that may induce changes in outcomes at the village level. Thereby, to ultimately "push" for another equilibrium that also reduces their costs to change

outcomes in their own households. A further key implication of our results is that it is important to measure impacts at both household and community levels, as measuring just one (typically the household level) may lead to an incomplete picture of effects, and sub-optimal policy design.

References

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. M. (2023). When should you adjust standard errors for clustering? *Quarterly Journal of Economics*, 138(1):1–35.

Adams, A. and Andrew, A. (2024). Revealed beliefs and the marriage market return to education. Technical report, University of Oxford, Job Market Paper, https://drive.google.com/file/d/1rDSYmeME8u6OP82ys1x4sfufo9NErgBz/view.

Aizer, A. (2010). The gender wage gap and domestic violence. *American Economic Review*, 100(4):1847–1859.

Anderberg, D., Cassidy, R., Dam, A., Janssens, W., Morsink, K., and van Veldhoven, A. (2024). Keeping the peace while getting your way: Information, persuasion and intimate partner violence. *IFS Working Paper No. 24/22.* Institute for Fiscal Studies.

Anderson, S. (2007). The economics of dowry and brideprice. *Journal of Economic Perspectives*, 21(4):151–174.

Andrew, A., Krutikova, S., Smarrelli, G., and Verma, H. (2024). Empowering adolescent girls: Does it take a village? Technical report, https://drive.google.com/file/d/1Rg0aeLzXL7A4cdCl5JTYE0-l0sXlaV-F/view.

Armand, A., Atwell, P., and Gomes, J. F. (2020). The reach of radio: Ending civil conflict through rebel demobilization. *American Economic Review*, 110(5):1395–1429.

Ashraf, N., Bau, N., Low, C., and McGinn, K. (2020). Negotiating a better future: How interpersonal skills facilitate intergenerational investment. *Quarterly Journal of Economics*, 135(2):1095–1151.

Baird, S., Chirwa, E., McIntosh, C., and Özler, B. (2010). The short-term impacts of a schooling conditional cash transfer program on the sexual behavior of young women. *Health Economics*, 19(S1):55–68.

Baird, S., McIntosh, C., and Özler, B. (2011). Cash or condition? Evidence from a cash transfer experiment. *Quarterly Journal of Economics*, 126(4):1709–1753.

Baird, S., McIntosh, C., and Özler, B. (2019). When the money runs out: Do cash transfers have sustained effects on human capital accumulation? *Journal of Development Economics*, 140:169–185.

Banerjee, A., La Ferrara, E., and Orozco, V. (2019). Entertainment, education, and attitudes toward domestic violence. *American Economic Review Papers and Proceedings*, 109:133–137.

Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., and Topalova, P. (2009). Powerful women: does exposure reduce bias? *Quarterly Journal of Economics*, 124(4):1497–1540.

Bellemare, M. F., Novak, L., and Steinmetz, T. L. (2015). All in the family: Explaining the

persistence of female genital cutting in West Africa. *Journal of Development Economics*, 116:252–265.

Bernhardt, A., Field, E., Pande, R., Rigol, N., Schaner, S., and Troyer-Moore, C. (2018). Male social status and women's work. *American Economic Review, Papers and Proceedings*, 108:363–67.

Bertrand, M., Kamenica, E., and Pan, J. (2015). Gender identity and relative income within households. *Quarterly Journal of Economics*, 130(2):571–614.

Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms.* Cambridge University Press.

Bicchieri, C., Jiang, T., and Lindemans, J. W. (2014). A social norms perspective on child marriage: The general framework. *UNICEF*.

Blackwell, B. S. (2000). Perceived sanction threats, gender, and crime: A test and elaboration of power-control theory. *Criminology*, 38(2):439–488.

Breen, R. and Cooke, L. P. (2005). The persistence of the gendered division of domestic labour. *European Sociological Review*, 21(1):43–57.

Bruhn, M. and McKenzie, D. (2009). In pursuit of balance: Randomization in practice in development field experiments. *American Economic Journal: Applied Economics*, 1(4):200–232.

Buchmann, N., Field, E., Glennerster, R., Nazneen, S., and Wang, X. Y. (2023). A signal to end child marriage: Theory and experimental evidence from Bangladesh. *American Economic Review*, 113(10):2645–2688.

Bursztyn, L., Cappelen, A. W., Tungodden, B., Voena, A., and Yanagizawa-Drott, D. H. (2023). How are gender norms perceived? *National Bureau of Economic Research. Working Paper no. 31049.*

Bursztyn, L., Fujiwara, T., and Pallais, A. (2017). 'Acting Wife:' Marriage market incentives and labor market investments. *American Economic Review*, 107(11):3288–3319.

Bursztyn, L., González, A. L., and Yanagizawa-Drott, D. (2020). Misperceived social norms: Women working outside the home in Saudi Arabia. *American Economic Review*, 110(10):2997–3029.

Bursztyn, L. and Jensen, R. (2017). Social image and economic behavior in the field: Identifying, understanding, and shaping social pressure. *Annual Review of Economics*, 9:131–153.

Chakraborty, P. and Serra, D. (2023). Gender and leadership in organisations: The threat of backlash. *The Economic Journal*, page uead110.

Chari, A., Heath, R., Maertens, A., and Fatima, F. (2017). The causal effect of maternal age at marriage on child wellbeing: Evidence from India. *Journal of Development Economics*, 127:42–55.

Collin, M. and Talbot, T. (2017). Do age-of-marriage laws work? Evidence from a large sample of developing countries. *CGD Working Paper No. 458.* Center for Global Development.

Corno, L., Hildebrandt, N., and Voena, A. (2020). Age of marriage, weather shocks, and the direction of marriage payments. *Econometrica*, 88(3):879–915.

Corno, L. and La Ferrara, E. (2022). Norms Replacement: A field experiment on female

genital cutting. Technical report, Extended Abstract, https://events.bse.eu/live/files/3752-cornolaferraraextendedabstractpdf.

Corno, L. and Voena, A. (2023). Child marriage as informal insurance: Empirical evidence and policy simulations. *Journal of Development Economics*, page 103047.

Dhar, D., Jain, T., and Jayachandran, S. (2022). Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in India. *American Economic Review*, 112(3):899–927.

Donati, D., Orozco-Olvera, V., and Rao, N. (2022). Using social media to change gender norms : An experiment within facebook messenger in India. *World Bank Policy Research Working Paper No. 10199*.

Duflo, E., Dupas, P., and Kremer, M. (2015). Education, HIV, and early fertility: Experimental evidence from Kenya. *American Economic Review*, 105(9):2757–97.

Eagly, A. H. and Carli, L. L. (1981). Sex of researchers and sex-typed communications as determinants of sex differences in influenceability: a meta-analysis of social influence studies. *Psychological Bulletin*, 90(1):1.

Edmonds, E., Feigenberg, B., and Leight, J. (2021). Advancing the agency of adolescent girls. *Review of Economics and Statistics*, 105(4):1–46.

Efferson, C., Vogt, S., and Fehr, E. (2020). The promise and the peril of using social influence to reverse harmful traditions. *Nature Human Behaviour*, 4(1):55–68.

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

Fors, H. C., Isaksson, A.-S., and Lindskog, A. (2021). Harmful Norms: Can Social Convention Theory Explain the Persistence of Female Genital Cutting in Africa? *IFN Working Paper No. 1417*. Research Institute of Industrial Economics.

Gage, A. and Van Rossem, R. (2006). Attitudes toward the discontinuation of female genital cutting among men and women in Guinea. *International Journal of Gynecology & Obstetrics*, 92(1):92–96.

Glennerster, R., Murray, J., and Pouliquen, V. (2023). Media, social pressure, and combating misinformation: Experimental evidence on mass media and contraception use in Burkina Faso. Technical report, University of Oxford, Job Market Paper, https://drive.google.com/file/d/1Jf_863MxVkBdS5rpQ0A-ssIYUYxuxTml/view.

Grasmick, H. G., Blackwell, B. S., and Bursik, R. J. (1993). Changes in the sex patterning of perceived threats of sanctions. *Law & Society Review*, 27(4):679–705.

Green, D. P., Wilke, A. M., and Cooper, J. (2020). Countering violence against women by encouraging disclosure: A mass media experiment in rural Uganda. *Comparative Political Studies*, 53(14):2283–2320.

Gulesci, S., Jindani, S., La Ferrara, E., Smerdon, D., Sulaiman, M., and Young, H. (2021). A stepping stone approach to understanding harmful norms. *CEPR Discussion Paper No. DP15776*. Centre for Economic Policy Research.

Hicks, J. and Hicks, D. L. (2019). Lucky late bloomers? Consequences of delayed marriage for women in rural western Kenya. *SSRN No. 2622189*. Available at SSRN: http://dx.doi.org/10.2139/ssrn.2622189.

Imbens, G. W. and Rubin, D. B. (2015). *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press.

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

Jayachandran, S. (2021). Social norms as a barrier to women's employment in developing countries. *IMF Economic Review*, 69(3):576–595.

Jensen, R. and Oster, E. (2009). The power of TV: Cable television and women's status in India. *Quarterly Journal of Economics*, 124(3):1057–1094.

Jensen, R. and Thornton, R. (2003). Early female marriage in the developing world. *Gender & Development*, 11(2):9–19.

Karing, A. (2024). Social signaling and childhood immunization: A field experiment in Sierra Leone. Technical report, University of Chicago, https://drive.google.com/file/d/137MDtXeZbrHU-xXUJWlJ-rs9J-e0I1k9/view.

Kuang, J., Ashraf, S., Shpenev, A., Delea, M. G., Das, U., and Bicchieri, C. (2020). Women are more likely to expect social sanctions for open defecation: Evidence from tamil nadu india. *Plos one*, 15(10):e0240477.

La Ferrara, E., Chong, A., and Duryea, S. (2012). Soap operas and fertility: Evidence from Brazil. *American Economic Journal: Applied Economics*, 4(4):1–31.

Leibbrandt, A., Wang, L. C., and Foo, C. (2018). Gender quotas, competitions, and peer review: Experimental evidence on the backlash against women. *Management Science*, 64(8):3501–3516.

Macmillan, R. and Gartner, R. (1999). When she brings home the bacon: Labor-force participation and the risk of spousal violence against women. *Journal of Marriage and the Family*, pages 947–958.

Malhotra, A. and Elnakib, S. (2021). 20 years of the evidence base on what works to prevent child marriage: A systematic review. *Journal of Adolescent Health*, 68(5):847–862.

McGavock, T. (2021). Here waits the bride? The effect of Ethiopia's child marriage law. *Journal of Development Economics*, 149:102580.

Nolen-Hoeksema, S. (2004). Gender differences in risk factors and consequences for alcohol use and problems. *Clinical psychology review*, 24(8):981–1010.

Punjab Bureau of Statistics and United Nations Children's Fund (2019). Pakistan - Punjab Multiple Indicator Cluster Survey 2017-2018. https://mics.unicef.org/surveys?display=card&keys=pakistan. Ref. PAK_2017_MICS-PUN_v01_M.

Roy, S., Hidrobo, M., Hoddinott, J., and Ahmed, A. (2019). Transfers, behavior change communication, and intimate partner violence: Post-program evidence from rural Bangladesh. *Review of Economics and Statistics*, 101(5):865–877.

Sagna, M. L. (2014). Gender differences in support for the discontinuation of female genital cutting in Sierra Leone. *Culture, Health & Sexuality*, 16(6):603–619.

Siddique, A., Vlassopoulos, M., and Zenou, Y. (2024). Leveraging edutainment and social networks to foster interethnic harmony. *SSRN No. 4810983.* Available at SSRN: https://dx.doi.org/10.2139/ssrn.4810983.

Sindh Bureau of Statistics and United Nations Children's Fund (2021). Pakistan - Sindh Multiple Indicator Cluster Survey 2018-2019. https://mics.unicef.org/surveys?display=card&keys=pakistan. Survey Findings Report.

Steinberg, L. and Monahan, K. C. (2007). Age differences in resistance to peer influence. *Developmental psychology*, 43(6):1531.

Tapsoba, A. (2023). Polygyny, timing of marriage and economic shocks in sub-saharan africa. Technical report, Toulouse School of Economics, Job Market Paper, https://www.dropbox.com/scl/fi/2lkcsq8ve6cruoifj5wpv/Polygyny_and_Economic_determinants_of_FFO_Nov2023.pdf?rlkey=r50yrzymfx2ldl3oczhisegr0&e=1&dl=0.

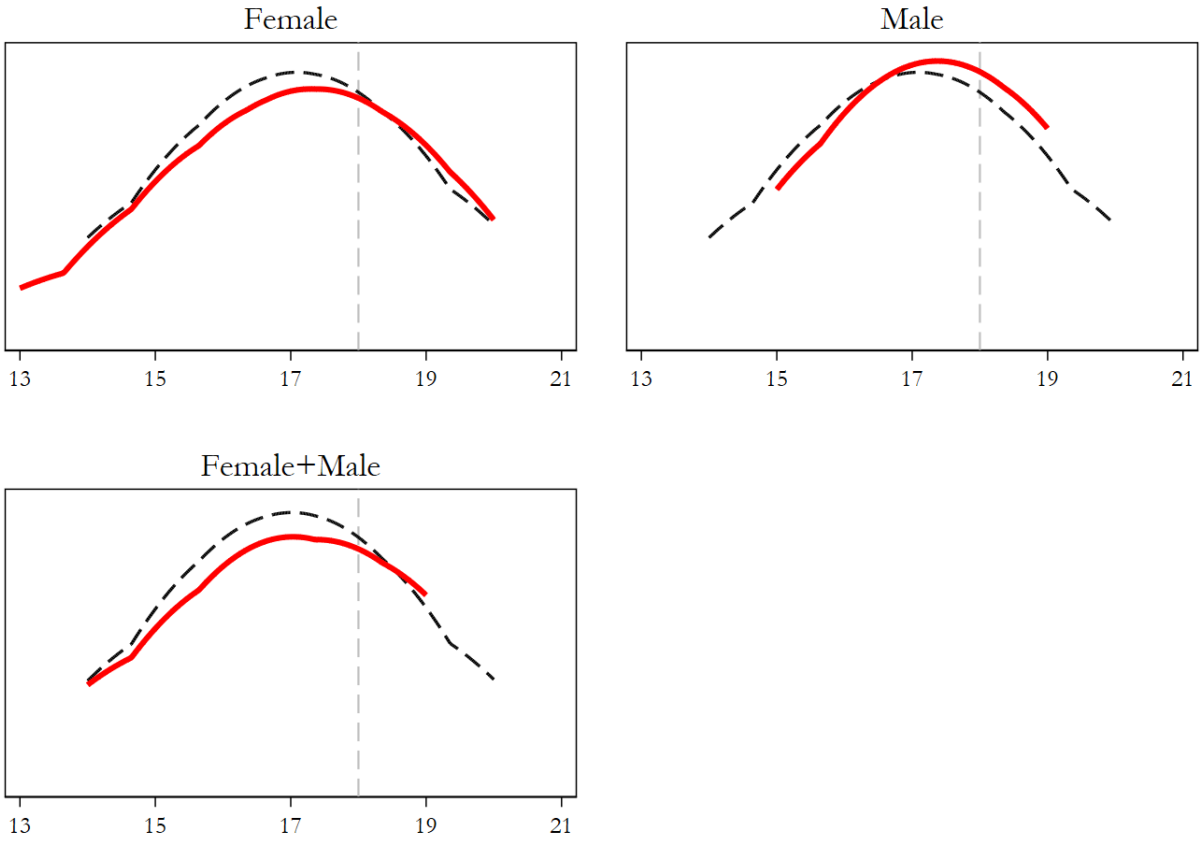
Tibbetts, S. G. and Herz, D. C. (1996). Gender differences in factors of social control and rational choice. *Deviant Behavior*, 17(2):183–208.

Wilson, N. et al. (2022). Child marriage bans and female schooling and labor market outcomes: Evidence from natural experiments in 17 low-and middle-income countries. *American Economic Journal: Economic Policy*, 14(3):449–477.

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

A Online Appendix

Figure O1: Target households: Kernel density plot of girls' marriage age



Notes: The figure presents kernel density plots of age at the time of marriage for all girl adolescents our from target households at endline. Age at the time of marriage is plotted on the x-axis. Black dashed line represents the variable's distribution in the control group. Red line represents the distribution of the variable in each treatment arm, i.e., Female arm (top-left), Male arm (top-right), Female+Male arm (bottom-left). Vertical grey line is the 18 years threshold, below which a marriage is considered a child marriage.

Table O1: Village-level descriptives and balance

	(1) C	(2) F	(3) M	(4) F+M	(5) F vs C	(6) M vs C	(7) F+M vs C	(8) F vs M	(9) F vs F+M	(10) M vs F+M
Boys primary school	0.227	0.311	0.364	0.364	0.084	0.136	0.136	-0.053	-0.053	0.000
	(0.424)	(0.468)	(0.487)	(0.487)	(0.372)	(0.157)	(0.157)	(0.601)	(0.601)	(1.000)
					[0.188]	[0.299]	[0.299]	[-0.110]	[-0.110]	[0.000]
Girls primary school	0.182	0.200	0.273	0.273	0.018	0.091	0.091	-0.073	-0.073	0.000
	(0.390)	(0.405)	(0.451)	(0.451)	(0.828)	(0.307)	(0.307)	(0.419)	(0.419)	(1.000)
					[0.046]	[0.216]	[0.216]	[-0.170]	[-0.170]	[0.000]
Mixed primary school	0.432	0.400	0.409	0.500	-0.032	-0.023	0.068	-0.009	-0.100	-0.091
	(0.501)	(0.495)	(0.497)	(0.506)	(0.761)	(0.829)	(0.522)	(0.931)	(0.342)	(0.391)
					[-0.064]	[-0.046]	[0.135]	[-0.018]	[-0.200]	[-0.180]
Girls can leave compound	0.636	0.733	0.750	0.773	0.097	0.114	0.136	-0.017	-0.039	-0.023
	(0.487)	(0.447)	(0.438)	(0.424)	(0.324)	(0.246)	(0.157)	(0.858)	(0.667)	(0.803)
					[0.207]	[0.245]	[0.299]	[-0.038]	[-0.090]	[-0.054]
Teashop in village	0.432	0.444	0.455	0.477	0.013	0.023	0.045	-0.010	-0.033	-0.023
	(0.501)	(0.503)	(0.504)	(0.505)	(0.905)	(0.831)	(0.669)	(0.924)	(0.757)	(0.831)
					[0.025]	[0.045]	[0.090]	[-0.020]	[-0.065]	[-0.045]
Female teachers in girls' school	0.636	0.711	0.727	0.750	0.075	0.091	0.114	-0.016	-0.039	-0.023
	(0.487)	(0.458)	(0.451)	(0.438)	(0.457)	(0.364)	(0.251)	(0.867)	(0.683)	(0.811)
					[0.158]	[0.194]	[0.245]	[-0.036]	[-0.087]	[-0.052]
Distance to nearest primary girls' school (km)	2.703	2.453	2.690	2.767	-0.250	-0.013	0.064	-0.237	-0.314	-0.078
	(1.486)	(1.499)	(1.520)	(1.634)	(0.504)	(0.972)	(0.873)	(0.542)	(0.437)	(0.852)
					[-0.168]	[-0.009]	[0.041]	[-0.157]	[-0.200]	[-0.047]
Distance to nearest primary girls' school (min)	14.741	13.964	17.038	15.704	-0.776	2.298	0.963	-3.074	-1.739	1.335
	(9.197)	(7.560)	(10.570)	(9.603)	(0.734)	(0.401)	(0.707)	(0.225)	(0.458)	(0.632)
					[-0.092]	[0.232]	[0.102]	[-0.335]	[-0.201]	[0.139]
Distance to nearest secondary girls' school (km)	8.128	6.993	6.695	7.219	-1.135	-1.433*	-0.909	0.298	-0.226	-0.524
	(3.901)	(3.524)	(3.043)	(3.711)	(0.157)	(0.062)	(0.273)	(0.676)	(0.773)	(0.482)
					[-0.305]	[-0.410]	[-0.239]	[0.091]	[-0.062]	[-0.141]
Distance to nearest secondary girls' school (min)	30.000	30.182	32.325	31.667	0.182	2.325	1.667	-2.143	-1.485	0.658
	(15.392)	(11.384)	(12.799)	(13.098)	(0.950)	(0.454)	(0.591)	(0.420)	(0.576)	(0.818)
					[0.013]	[0.164]	[0.117]	[-0.177]	[-0.121]	[0.050]
Distance to nearest town (min)	30.523	32.267	28.227	28.500	1.744	-2.295	-2.023	4.039	3.767	-0.273
	(14.788)	(13.223)	(12.115)	(11.744)	(0.559)	(0.427)	(0.478)	(0.135)	(0.157)	(0.915)
					[0.124]	[-0.170]	[-0.151]	[0.319]	[0.301]	[-0.023]
Number of households	172.000	198.689	220.909	204.500	26.689	48.909	32.500	-22.220	-5.811	16.409
	(214.764)	(369.649)	(340.191)	(187.771)	(0.677)	(0.421)	(0.451)	(0.768)	(0.925)	(0.780)
					[0.088]	[0.172]	[0.161]	[-0.063]	[-0.020]	[0.087]
Joint Significance(p-value)					0.855	0.180	0.312	0.541	0.826	0.983

Notes: Column 1-4 show the mean of the variable in each experimental arm – the Female arm (F), the Male arm (M), the Female+Male arm (F+M), and Control arm (C). Standard deviations are indicated in parentheses. Column 5-10 shows the difference in means for each combination of experimental arms. In Columns 5-10, p-values from robust standard errors clustered at the village level (unit of randomization) are indicated in parenthesis using a logit regression for binary variables, and an OLS regression for continuous variables. Normalized differences are reported in square brackets, calculated as the difference between the sample means of experimental arms divided by the square root of the sum of the sample variances. Row "Joint Significance (p-value)" reports the p-value on the chi-squared test that coefficients and p-values from all 12 regressions on balance variables are jointly unrelated to the treatment assignment. Stars indicate: *** 1 percent ** 5 percent * 10 percent level of significance.

Table O2: Attrition by treatment arms at midline

	(1) M	(2) F	(3) F+M	(4) C	(5) M vs C	(6) F vs C	(7) F+M vs C	(8) F vs M	(9) F+M vs M	(10) F+M vs F	(11) Observations
Father	0.061	0.062	0.073	0.053	0.008	0.009	0.020	0.002	0.012	0.010	1687
	25	*27*	*31*	*22*	(0.682)	(0.587)	(0.354)	(0.931)	(0.600)	(0.633)	
					[0.818]	[0.818]	[0.818]	[0.932]	[0.818]	[0.818]	
Mother	0.032	0.016	0.014	0.012	0.020	0.004	0.002	-0.015	-0.017	-0.002	1687
	13	*7*	*6*	*5*	(0.098)*	(0.616)	(0.799)	(0.217)	(0.158)	(0.818)	
					[0.435]	[0.818]	[0.818]	[0.435]	[0.435]	[0.818]	
Adolescent	0.022	0.005	0.012	0.007	0.015	-0.003	0.005	-0.017	-0.010	0.007	1687
	9	*2*	*5*	*3*	(0.107)	(0.616)	(0.530)	(0.048)**	(0.315)	(0.293)	
					[0.321]	[0.617]	[0.617]	[0.290]	[0.473]	[0.473]	

Notes: The table presents tests for differential attrition at midline. Panel consists of 1687 target households, with 756 households in Sindh province, and 931 households in Punjab province. In each household, three respondents were to be surveyed, i.e. father, mother and adolescent. In Column 1-4, the attrition rate for each individual respondent by experimental arm - the Female arm (F), the Male arm (M), the Female+Male arm (F+M), and Control arm (C), are presented. An individual respondent is considered an attritor if we were unable to survey him/her at midline. In italics, the raw number of individuals that attrited are reported. In Column 5-10, the difference in attrition rate between experimental arms from logit regressions are presented. Fixed effects for randomisation strata are not included, due to perfect prediction given our low attrition rate. *P*-values for marginal treatment effects from standard errors clustered at village level (unit of randomisation) are indicated in parentheses. *Q*-values correcting for false discovery rate within each family using the Benjamini–Hochberg procedure are indicated in square brackets. Stars indicate: *** 1 percent ** 5 percent * 10 percent level of significance.

Table O3: Attrition by treatment arms at endline

	(1) M	(2) F	(3) F+M	(4) C	(5) M vs C	(6) F vs C	(7) F+M vs C	(8) F vs M	(9) F+M vs M	(10) F+M vs F	(11) Observations
Father	0.143	0.122	0.155	0.130	0.013	-0.007	0.025	-0.021	0.012	0.033	1687
	59	*53*	*66*	*54*	(0.637)	(0.782)	(0.387)	(0.455)	(0.697)	(0.254)	
					[0.782]	[0.782]	[0.782]	[0.782]	[0.782]	[0.782]	
Mother	0.080	0.046	0.080	0.060	0.020	-0.014	0.020	-0.034	-0.000	0.034	1687
	33	*20*	*34*	*25*	(0.381)	(0.434)	(0.385)	(0.101)	(0.991)	(0.102)	
					[0.521]	[0.521]	[0.521]	[0.306]	[0.991]	[0.306]	
Adolescent	0.063	0.035	0.061	0.046	0.017	-0.011	0.015	-0.028	-0.002	0.026	1687
	26	*15*	*26*	*19*	(0.391)	(0.466)	(0.444)	(0.131)	(0.928)	(0.155)	
					[0.559]	[0.559]	[0.559]	[0.465]	[0.929]	[0.465]	

Notes: The table presents tests for differential attrition at endline. Panel consists of 1687 target households, with 756 households in Sindh province, and 931 households in Punjab province. In each household, three respondents were to be surveyed, i.e. father, mother and adolescent. In Column 1-4, the attrition rate for each individual respondent by experimental arm - the Female arm (F), the Male arm (M), the Female+Male arm (F+M), and Control arm (C), are presented. An individual respondent is considered an attritor if we were unable to survey him/her at endline. In italics, the raw number of individuals that attrited are reported. In Column 5-10, the difference in attrition rate between experimental arms from logit regressions are presented. Fixed effects for randomisation strata are not included, due to perfect prediction given our low attrition rate. *P*-values for marginal treatment effects from standard errors clustered at village level (unit of randomisation) are indicated in parentheses. *Q*-values correcting for false discovery rate within each family using the Benjamini–Hochberg procedure are indicated in square brackets. Stars indicate: *** 1 percent ** 5 percent * 10 percent level of significance.

Table O4: Number of attriters by treatment arms in Sindh at midline and endline

	Midline				Endline				
	(1) M	(2) F	(3) M+F	(4) C	(5) M	(6) F	(7) M+F	(8) C	(9) Observations
Father	0.070	0.053	0.042	0.042	0.124	0.063	0.099	0.069	756
	13	*10*	*8*	*8*	*23*	*12*	*19*	*13*	
Mother	0.059	0.021	0.026	0.016	0.070	0.005	0.026	0.026	756
	11	*4*	*5*	*3*	*13*	*1*	*5*	*5*	
Adolescent	0.049	0.011	0.026	0.016	0.059	0.005	0.021	0.016	756
	9	*2*	*5*	*3*	*11*	*1*	*4*	*3*	

Notes: The table presents the attrition rate and numbers at midline and endline for Sindh. The panel consists of 756 households in Sindh province. The number in normal font represents the attrition rate and the number in italics presents the raw number of individuals that attrited. In each household, three respondents were surveyed, i.e. the father, the mother and the adolescent. In Columns 1-4 for midline and Columns 5-8 for endline, the attrition rate and numbers for each individual respondent by experimental arm - the Female arm (F), the Male arm (M), the Female+Male arm (F+M), and Control arm (C), are provided. An individual respondent is considered attrited if we were unable to survey him/her in the respective surveyround. Regression analyses for the sub-sample cannot be carried out due to low number of attrited individuals.

Table O5: Number of attriters by treatment arms in Punjab at midline and endline

	Midline				Endline				
	(1) M	(2) F	(3) M+F	(4) C	(5) M	(6) F	(7) M+F	(8) C	(9) Observations
Father	0.053	0.070	0.098	0.062	0.159	0.169	0.201	0.181	931
	12	*17*	*23*	*14*	*36*	*41*	*47*	*41*	
Mother	0.009	0.012	0.004	0.009	0.088	0.078	0.124	0.088	931
	2	*3*	*1*	*2*	*20*	*19*	*29*	*20*	
Adolescent	0.000	0.000	0.000	0.000	0.066	0.058	0.094	0.070	931
	0	*0*	*0*	*0*	*15*	*14*	*22*	*16*	

Notes: The table presents the attrition rate and numbers at midline and endline for Punjab. The panel consists of 931 households in Punjab province. The number in normal font represents the attrition rate and the number in italics presents the raw number of individuals that attrited. In each household, three respondents were surveyed, i.e. the father, the mother and the adolescent. In Columns 1-4 for midline and Columns 5-8 for endline, the attrition rate and numbers for each individual respondent by experimental arm - the Female arm (F), the Male arm (M), the Female+Male arm (F+M), and Control arm (C), are provided. An individual respondent is considered attrited if we were unable to survey him/her in the respective surveyround. Regression analyses for the sub-sample cannot be carried out due to low number of attrited individuals.

Table O6: Comparison of target households with administrative data

	Baseline		MICS	
	(1) Sindh	(2) Punjab	(3) Sindh	(4) Punjab
Assets				
Electricity	93.5	97.7	72.1	97.8
Refrigerator	15.2	15.2	16.8	–
Computer	6.0	6.6	2.7	10.0
Television	36.1	57.9	33.5	61.6
Mobile Phone	85.3	96.3	83.2	94.2
Household composition				
HH head has no education	36.9	41.0	46.0	38.3
Age of HH head	46.5	47.7	46.9	49.3
Male HH head	100.0	97.6	89.9	89.6
Mother has no education	81.1	77.6	69.1	53.2

Notes: Column 1-2 show the mean of household assets and composition in target household. Data is used from the baseline survey in Sindh and retrospective baseline questions in Punjab. Column 3-4 show the corresponding mean from the Pakistan Multiple Indicator Cluster Survey (MICS) 2017-2018 in Sindh and Punjab provinces. This survey wave corresponds closest to the time of our data collection. Household assets descriptives use the rural sample.

Table O7: Participation in screening and group discussion by treatment arm

	Male attendees	Female attendees
Female		18.09
		(2.58)
Male	15.77	
	(2.38)	
Female+Male	14.94	16.15
	(2.71)	(3.75)

Notes: The table presents the average number of participants in each treatment arm, disaggregated by the gender of the attendees. Standard deviation are reported in parentheses. In the Female+Male intervention, two separate sessions were conducted - one for males and one for females from the same household. In this treatment arm participants were explicitly informed that the other gender was participating in the same intervention.

Table O8: Target households: Child marriage outcomes for adolescent girls

	Annual Hazard Rate			
	Alternative 1		Alternative 2	
	Midline	Endline	Midline	Endline
	(1)	(2)	(3)	(4)
Female	0.021	-0.011	0.004	-0.013
	(0.004)***	(0.238)	(0.721)	(0.280)
	[0.003]***	[0.294]	[0.767]	[0.341]
Male	-0.030	-0.019	-0.018	-0.017
	(0.000)***	(0.024)**	(0.036)**	(0.107)
	[0.000]***	[0.046]**	[0.048]**	[0.150]
Female+Male	-0.011	-0.024	-0.007	-0.017
	(0.049)**	(0.004)***	(0.440)	(0.105)
	[0.060]*	[0.010]**	[0.461]	[0.119]
Observations	2990	2843	1641	2159
Control Mean	0.044	0.053	0.029	0.042
p-val M $\neq$ F	0.000	0.314	0.032	0.691
p-val FM $\neq$ F	0.000	0.097	0.323	0.695
p-val FM $\neq$ M	0.000	0.465	0.187	0.996

Notes: The table presents treatment effects on the average hazard rate into child marriage for girl adolescents from our target households, using a discrete approximation of a censored duration model estimated from Equation 2 at midline and endline. Marriage outcomes are as reported by the adolescent themselves as pre-specified in the pre-analysis plan. Data is an adolescent-year panel, where each adolescent contributes at most five observations to the sample, one observation for each post-baseline at-risk year between 13 and 17 until she is either married or passes the 18-year threshold after which she exists the data. The dependent variable is coded as zero for the ages prior to her marriage, as unity for the age at which she marries. Columns 1-2 report the average hazard rate into child marriage, where right censored observations – that is, observations between the current age at the corresponding surveyround and age 18 for girls who are aged less than 18 and still unmarried by midline or endline – that is are still at risk of being child-married are coded as the cohort-specific, age-specific, and treatment arm-specific probability of marriage. Columns 3-4 report the average hazard rate into child marriage, where right censored observations – that is, observations between the current age at the corresponding survey-round and age 18 for girls who are aged less than 18 and still unmarried by midline or endline – are coded as missing. Fixed effects for age cohort, current age and randomization strata are included. *P*-values for marginal treatment effects are based on standard errors clustered at village level (unit of randomisation), and are indicated in parentheses. Exact *p*-values for marginal treatment effects from randomization inference tests based on 1000 permutations are provided in square brackets. Stars indicate: *** 1 percent ** 5 percent * 10 percent level of significance. The row "Control Mean" indicates the average annual probability of child marriage over cohorts in the control group. *P*-values for all comparisons of treatment effects between each experimental arm are reported at the bottom of the table: Male arm (M), Female arm (F) and Female+Male arm (FM).

Table O9: Target households: Marriage outcomes reported and verified adolescent girls

	Reported		Verified	
	(1) Midline	(2) Endline	(3) Midline	(4) Endline
Female	−0.001	−0.065	−0.010	−0.051
	(0.969)	(0.175)	(0.772)	(0.259)
Male	−0.053**	−0.097**	−0.053**	−0.090**
	(0.039)	(0.022)	(0.039)	(0.019)
Female+Male	−0.029	−0.084**	−0.034	−0.058
	(0.285)	(0.042)	(0.203)	(0.125)
Observations	828	798	825	780
Control Mean	0.100	0.222	0.100	0.194

Notes: The table presents marginal treatment effects on marriage outcomes for girl adolescents $(\eta + \rho)$ from the pre-specified logit regression for marriage in Equation 1. Marriage outcomes are as reported by the adolescent themselves as pre-specified in the pre-analysis plan. Reported is the estimation for midline and endline. In Column 1-2, the binary dependent variable takes on value one if the adolescent was married at midline and endline respectively. In Column 3-4, the binary dependent variable takes on value one if the adolescent was traced and responded to the marriage verification module at midline and endline respectively. Fixed effects for randomisation strata are included at endline and not at midline due to perfect prediction. *P*-values for marginal treatment effects from standard errors clustered at village level (unit of randomisation) are indicated in parentheses. Exact *p*-values for marginal treatment effects from randomization inference tests based on 1000 permutations are provided in square brackets. Stars indicate: *** 1 percent ** 5 percent * 10 percent level of significance. The row "Control Mean" indicates the probability of marriage in the control group.*P*-values for all comparisons of treatment effects between each experimental arm are reported at the bottom of the table: Male arm (M), Female arm (F) and Female+Male arm (FM).

Table O10: Target households: Marriage for adolescent girls reported by father, mother, any household member

	Father's response		Mother's response		Any respondent	
	(1) Midline	(2) Endline	(3) Midline	(4) Endline	(5) Midline	(6) Endline
Female	0.005	−0.050	−0.006	−0.056	−0.006	−0.052
	(0.900)	(0.328)	(0.873)	(0.248)	(0.868)	(0.283)
Male	−0.047	−0.098	−0.054	−0.088	−0.057	−0.089
	(0.063)*	(0.034)**	(0.040)**	(0.047)**	(0.036)**	(0.042)**
Female+Male	−0.039	−0.106	−0.030	−0.099	−0.034	−0.086
	(0.148)	(0.019)**	(0.281)	(0.020)**	(0.257)	(0.054)*
Observations	775	722	820	784	828	793
Control Mean	0.096	0.220	0.101	0.216	0.114	0.228

Notes: The table presents marginal treatment effects on marriage outcomes for girl adolescents ($\eta+\rho$) from the pre-specified logit regression in Equation 1. Reported is the estimation for midline and endline. Marriage outcomes of the adolescent is reported separately by the father, mother or any respondent. That is, Column 1-2 if the father said the adolescent was married, Column 3-4 if the mother said the adolescent was married, and Column 5-6 if any respondent (father, mother or adolescent) said that the adolescent was married. Fixed effects for randomisation strata are included at endline and not at midline due to perfect prediction. P-values for marginal treatment effects from standard errors clustered at village level (unit of randomisation) are indicated in parentheses. Stars indicate: *** 1 percent ** 5 percent * 10 percent level of significance. Row "Control Mean" indicates the marriage rate for girl adolescents in the control group.

Table O11: Target households: (Child) marriage and marriage age outcomes for adolescent boys

	Married		Marriage age		Child Marriage	
	(1) Midline	(2) Endline	(3) Midline	(4) Endline	(5) Midline	(6) Endline
Female	-0.002	-0.006	-1.000	-0.130	0.008	0.006
	(0.904)	(0.813)	(0.357)	(0.826)	(0.445)	(0.757)
	[0.808]	[0.824]	[0.482]	[0.860]	[0.650]	[0.794]
Male	0.009	0.012	0.000	0.456	0.014	0.009
	(0.636)	(0.637)	(1.000)	(0.387)	(0.254)	(0.668)
	[0.732]	[0.659]	[1.000]	[0.446]	[0.370]	[0.747]
Female+Male	0.009	-0.021	-0.250	0.177	0.019	-0.006
	(0.628)	(0.335)	(0.714)	(0.742)	(0.144)	(0.768)
	[0.705]	[0.359]	[0.771]	[0.757]	[0.226]	[0.765]
Observations	840	800	28	55	840	800
Control Mean	0.030	0.073	17.000	16.929	0.010	0.036

Notes: The table presents marginal treatment effects on marriage outcomes for boy adolescents (η) from the pre-specified logit regression for (child) marriage and OLS regression for marriage age in Equation 1. Marriage outcomes are as reported by the adolescent themselves as pre-specified in the pre-analysis plan. Reported is the estimation for midline and endline. In column 1-2, the binary dependent variable takes on value one if the adolescent was married at midline and endline respectively. In Column 3-4, the dependent variable is the age of the girl at the time of marriage, conditional on being married at midline and endline. In Column 5-6, the binary dependent variable takes on value one if the adolescent was less than 18 years old at the time of marriage for midline and endline. Fixed effects for randomisation strata are included at endline and not at midline due to perfect prediction. P-values for marginal treatment effects from standard errors clustered at village level (unit of randomisation) are indicated in parentheses. Exact p-values for marginal treatment effects from randomization inference tests based on 1000 permutations are provided in square brackets. Stars indicate: *** 1 percent ** 5 percent * 10 percent level of significance. The row "Control Mean" indicates the probability of marriage, average age of marriage and probability of child marriage in the control group.

Table O12: Target households: Engagement outcomes for adolescent girls

	Engaged		Newly Engaged	
	(1) Midline	(2) Endline	(3) Midline	(4) Endline
Female	−0.020	−0.016	0.033	0.009
	(0.682)	(0.779)	(0.420)	(0.802)
Male	0.042	0.094	0.058	0.044
	(0.417)	(0.112)	(0.159)	(0.167)
Female+Male	−0.015	0.008	0.012	0.012
	(0.776)	(0.884)	(0.764)	(0.696)
Observations	746	665	736	665
Control Mean	0.283	0.293	0.177	0.076

Notes: The table presents marginal treatment effects on engagement outcomes for girl adolescents ($\eta + \rho$) from the pre-specified logit regression in Equation 1. Reported is the estimation for midline and endline. In Column 1-2, the dependent variable is a binary variable that takes on value one (zero otherwise) if the adolescent is engaged/promised. Column 3-4 presents effects on new engagements, that is, adolescents who were not engaged at baseline but engaged at midline (Column 3: midline) and those who were not engaged at midline but engaged at endline (Column 4). For engagement between baseline and midline, for Sindh province we use responses from the baseline survey. For Punjab where security concerns prevented us from having a baseline survey, we estimate newly engaged as those adolescents who are engaged at midline, and their age at engagement is one year less than or equal to their age at midline, as these adolescents are most likely to be engaged between our survey rounds. Fixed effects for randomisation strata are included. *P*-values for marginal treatment effects from standard errors clustered at village level (unit of randomisation) are indicated in parentheses. Stars indicate: *** 1 percent ** 5 percent * 10 percent level of significance. Row "Control Mean" indicates the probability of being engaged and probability of being newly engaged in the control group.

Table O13: Village-month level: (Child) marriage outcomes for girls

	Child Marriage		Marriage Age	
	Midline	Endline	Midline	Endline
	(1)	(2)	(3)	(4)
Post X Female	-0.402	-0.208	1.277	0.664
	(0.003)***	(0.012)**	(0.029)**	(0.050)*
Post X Male	-0.340	-0.103	0.747	0.271
	(0.004)***	(0.204)	(0.108)	(0.434)
Post X Female+Male	-0.255	-0.244	0.742	0.826
	(0.028)**	(0.001)***	(0.149)	(0.014)**
Observations	677	1136	677	1136
Control Mean	0.643	0.413	17.333	18.458
p-val M $\neq$ F	0.651	0.218	0.333	0.266
p-val FM $\neq$ F	0.263	0.643	0.335	0.641
p-val FM $\neq$ M	0.481	0.070	0.991	0.116
p-val M $>$ F	0.326	0.109		
p-val M $>$ FM	0.759	0.035		

Notes: The table presents estimates for girls of ψ from Equation 3 using village fixed effects regressions with marriage data aggregated at the village-month level. In Columns 1-2, the binary dependent variable takes on value one if in the village-month at least one observed marriage involved a bride below the age of 18, and zero if all observed marriages involved a bride 18 and above. In Columns 3-4, the dependent variable is the average age of all brides who got married in the village-month. "Midline" counts observations between the pre-treatment months and January 2020 and "endline" counts observations between the pre-treatment months and March 2021. The variable "post" is a binary variable that takes on the value zero if the village-month lies in the period before (not including) July 2019, the period before or during the intervention; and takes on value one if the observation month lies in the period after the intervention was completed. Fixed effects for randomisation strata and their interaction with "post" are included. *P*-values from standard errors clustered at village level (unit of randomisation), are indicated in parentheses. Stars indicate: *** 1 percent ** 5 percent * 10 percent level of significance. The row "Control Mean" reports the likelihood of child marriage and average marriage age of girls, aggregated to a village-month observation, during the pre-intervention period in the control group. *P*-values for all comparisons of treatment effects between each experimental arm as well as one-sided tests are reported at the bottom of the table: Male arm (M), Female arm (F) and Female+Male arm (FM).

Table O14: Village-month level: Number of child marriages

	Any Marriage		Number of Marriages	
	(1) Midline	(2) Endline	(3) Midline	(4) Endline
Post × Female	−0.073	−0.050	−0.075	−0.047
	(0.017)**	(0.078)*	(0.014)**	(0.116)
Post × Male	−0.034	−0.012	−0.034	−0.010
	(0.286)	(0.659)	(0.288)	(0.715)
Post × Female+Male	−0.042	−0.049	−0.051	−0.062
	(0.162)	(0.070)*	(0.093)*	(0.025)**
Observations	2451	4740	2451	4740
Control Mean	0.106	0.072	0.106	0.078

Notes: The table presents estimates for girls of ψ from Equation 3 using village fixed effects regressions with data at the level of marriages aggregated at the village-month level. In Column 1-2, the binary dependent variable takes value one if there was at least one marriage in the village-month that involved a bride below the age of 18. It takes on a value 0 if there was no marriage or all marriages involved a bride aged 18 and above in a village in a given month. In Column 3-4, the dependent variable is a count variable of the number of marriages in the village-month that involved a bride below the age of 18. The count variable is 0 if there was no marriage or all marriages involved a bride aged 18 and above in a village in a given month. "Midline" counts observations between the pre-treatment months and January 2020, and "Endline" counts observations between the pre-treatment months and March 2021. The variable "post" is a binary variable that takes on the value zero if the village-month lies in the period before July (not including) 2019, the period before or during the intervention; and takes on value one if the observation month lies in the period after the intervention was completed. Fixed effects for randomisation strata and their interaction with "post" are included. *P*-values from standard errors clustered at village level (unit of randomisation), are indicated in parentheses. Stars indicate: *** 1 percent ** 5 percent * 10 percent level of significance. The row 'control mean' reports the specific variable in the control group.

Table O15: Proposals coming in for girl adolescents in target households

	(1) All	(2) Age <18	(3) Age >=18
Female	-0.058	-0.127	0.131
	(0.283)	(0.026)**	(0.176)
Male	-0.077	-0.109	0.021
	(0.189)	(0.104)	(0.844)
Female+Male	-0.114	-0.141	-0.019
	(0.020)**	(0.010)**	(0.865)
Observations	722	548	174
Control Mean	0.324	0.336	0.289
p-val M $\neq$ F	0.729	0.774	0.267
p-val FM $\neq$ F	0.213	0.761	0.152
p-val FM $\neq$ M	0.464	0.590	0.725

Notes: The table presents marginal treatment effects ($\eta + \rho$) on marriage proposals coming in for girl adolescents from target households using the pre-specified logit regression in Equation 1 at endline. The dependent variable is a binary variable that takes on value one (zero otherwise) if the adolescent received a marriage proposal. Column 1 shows effects on proposals coming in for all girl adolescents. In Column 2, effects are estimated on the sub-sample of girl adolescents who are below 18 prior to the endline and in Column 3 for those who are 18 and over. Fixed effects for randomisation strata are not included due to perfect prediction. *P*-values for marginal treatment effects from standard errors clustered at village level (unit of randomisation) are indicated in parentheses. Stars indicate: *** 1 percent ** 5 percent * 10 percent level of significance. Row "Control Mean" indicates the probability of marriage proposals coming in in the control group. *P*-values for all comparisons of treatment effects between each experimental arm are reported at the bottom of the table: Male arm (M), Female arm (F) and Female+Male arm (FM).

Table O16: Village-month level: (Child) marriage outcomes for girls by village-level agency of targeted women

	Leave compound		Education		Community meetings	
	(1) No	(2) Yes	(3) No	(4) Yes	(5) No	(6) Yes
Female	-0.019	-0.300	-0.042	-0.263	-0.230	-0.171
	(0.894)	(0.011)**	(0.834)	(0.011)**	(0.230)	(0.083)*
		[0.094]		[0.281]		[0.763]
Male	-0.038	-0.158	0.095	-0.177	-0.040	-0.134
	(0.821)	(0.167)	(0.601)	(0.090)*	(0.809)	(0.210)
		[0.513]		[0.148]		[0.598]
Female+Male	-0.134	-0.325	-0.175	-0.290	-0.247	-0.237
	(0.202)	(0.003)***	(0.294)	(0.002)***	(0.069)*	(0.024)**
		[0.160]		[0.506]		[0.952]
Observations	311	825	300	836	434	702

Notes: The table presents estimates for girls of ψ from Equation 3using village fixed effects regressions with data at the level of marriages aggregated at the village-month level. The binary dependent variable takes on value one if in village-month at least one observed marriage involved a bride below the age of 18, and zero if all observed marriages involved a bride 18 and above. The variable "post" is a binary variable that takes on the value zero if the village-month lies in the period before (not including) July 2019, the period before or during the intervention; and takes on value one if the observation month lies in the period after the intervention was completed. Column 1-2 presents impacts on child marriage in the sample of villages where women *cannot* or *can* leave the compound of the household, respectively. Column 3-4, presents impacts on child marriage in the sample of villages where *at least one woman* or *none of the women* among our target households has any education, respectively. Column 5-6, presents impacts on child marriage in the sample of villages where *at least one woman* or *none of the women* among our target households attended a community meeting, respectively. Fixed effects for randomisation strata and their interaction with "post" are included. *P*-values from standard errors clustered at village level (unit of randomisation), are indicated in parentheses. In square brackets, the *p*-value of the *chi-squared* test, that estimated treatment effects from the two seemingly unrelated regressions (SUR) are significantly different from each other. Stars indicate: *** 1 percent ** 5 percent * 10 percent level of significance.

Table O17: Fathers' beliefs about health risks to a girl in case of child marriage

	Mother			Child		
	(1) Depression	(2) Poor health	(3) Mortality	(4) Low birthweight	(5) Poor health	(6) Mortality
Female	−0.017	0.049	0.015	0.021	0.094	0.033
	(0.576)	(0.244)	(0.714)	(0.399)	(0.046)**	(0.247)
Male	−0.008	−0.002	−0.007	0.017	0.049	0.025
	(0.808)	(0.969)	(0.856)	(0.456)	(0.293)	(0.404)
Female+Male	0.007	0.038	−0.003	0.013	0.054	0.023
	(0.837)	(0.367)	(0.931)	(0.592)	(0.245)	(0.396)
Observations	1545	1545	1545	1545	1545	1545
Control Mean	0.207	0.590	0.308	0.112	0.355	0.112

Notes: The table presents intent-to-treat effects from Logit regressions at midline. The dependent variable takes on value one if fathers, mention that the expected risks to a girl (column 1-3) and the girl's resulting child (column 4-6) in case of early marriage or childbearing are: depression to mother (column 1), poor health of mother from early childbearing (column 2), mortality risk to mother from early child bearing (column 3), low birth weight of child from early childbearing (column 4), poor health of child from early childbearing (column 5) or mortality risk to child from early childbearing (column 6). Fixed effects for randomisation strata are included. *P*-values from standard errors clustered at village level (unit of randomisation) are indicated in parentheses. Stars indicate: *** 1 percent ** 5 percent * 10 percent level of significance. Row "Control Mean" indicates the expectations of fathers in the control group.

Table O18: Mothers' beliefs about health risks to a girl in case of child marriage

	Mother			Child		
	(1) Depression	(2) Poor health	(3) Mortality	(4) Low birthweight	(5) Poor health	(6) Mortality
Female	0.015	0.051	0.014	−0.017	0.035	0.003
	(0.594)	(0.210)	(0.762)	(0.685)	(0.475)	(0.935)
Male	−0.044	0.033	0.023	−0.041	−0.021	−0.022
	(0.061)*	(0.468)	(0.610)	(0.286)	(0.609)	(0.495)
Female+Male	0.006	0.080	0.006	−0.063	−0.029	−0.055
	(0.814)	(0.042)**	(0.912)	(0.073)*	(0.502)	(0.118)
Observations	1650	1650	1650	1650	1650	1650
Control Mean	0.122	0.501	0.335	0.149	0.213	0.193

Notes: The table presents intent-to-treat effects from Logit regressions at midline. The dependent variable takes on value one if mothers, mention that the expected risks to a girl (column 1-3) and the girl's resulting child (column 4-6) in case of early marriage or childbearing are: depression to mother (column 1), poor health of mother from early childbearing (column 2), mortality risk to mother from early child bearing (column 3), low birth weight of child from early childbearing (column 4), poor health of child from early childbearing (column 5) or mortality risk to child from early childbearing (column 6). Fixed effects for randomisation strata are included. *P*-values from standard errors clustered at village level (unit of randomisation) are indicated in parentheses. Stars indicate: *** 1 percent ** 5 percent * 10 percent level of significance. Row "Control Mean" indicates the expectations of mothers in the control group.

Table O19: Mothers' and fathers' beliefs about spousal quality benefits from marriage delay to 18

	Secondary School		High School	
	(1)	(2)	(3)	(4)
	Father	Mother	Father	Mother
Female	0.077	0.104	0.054	0.115
	(0.111)	(0.063)*	(0.284)	(0.034)**
	[0.395]	[0.127]	[0.395]	[0.127]
Male	0.115	0.103	0.072	0.141
	(0.019)**	(0.068)*	(0.152)	(0.010)***
	[0.076]*	[0.091]*	[0.203]	[0.020]**
Female+Male	0.082	0.109	0.077	0.116
	(0.082)*	(0.053)*	(0.124)	(0.031)**
	[0.246]	[0.071]*	[0.246]	[0.062]*
Observations	769	814	769	814
Control Mean	0.597	0.566	0.265	0.254
p-val M $\neq$ F	0.447	0.978	0.717	0.646
p-val FM $\neq$ F	0.931	0.928	0.641	0.984
p-val FM $\neq$ M	0.481	0.907	0.923	0.658

Notes: The table presents marginal treatment effects for girl adolescents ($\eta + \rho$) from the pre-specified logit regression in Equation 1 at midline. The binary dependent variable takes on value one if the respondent - mother or father - states that they expect that the future spouse of their adolescent girl will have atleast completed secondary (Grade 10 and above) or high school (Grade 12 and above), if she is married at 18. Fixed effects for randomisation strata are included. *P*-values for marginal treatment effects from standard errors clustered at village level (unit of randomisation) are indicated in parentheses. *Q*-values correcting for false discovery rate within each family using the Benjamini–Hochberg procedure are indicated in square brackets. The family over which we correct reflects the the four schooling levels per respondent per treatment arm. Stars indicate: *** 1 percent ** 5 percent * 10 percent level of significance. The row "Control Mean" indicates the percentage in the control group stating the above expectation at the given age.

Table O20: Over- and underestimation of other men's and women's attitudes towards acceptability of child marriage for girls

	Mother				Father			
	Accept 12-15		Accept 16-17		Accept 12-15		Accept 16-17	
	(1) Men	(2) Women	(3) Men	(4) Women	(5) Men	(6) Women	(7) Men	(8) Women
Female	−0.086	−0.011	0.030	0.014	−0.019	0.033	0.042	0.028
	(0.004)***	(0.689)	(0.410)	(0.684)	(0.474)	(0.292)	(0.236)	(0.509)
Male	−0.058	−0.010	0.054	0.045	−0.014	0.001	0.076	0.038
	(0.067)*	(0.741)	(0.157)	(0.229)	(0.619)	(0.975)	(0.020)**	(0.433)
Female+Male	−0.030	−0.021	0.053	−0.032	−0.011	0.006	0.041	−0.046
	(0.317)	(0.520)	(0.182)	(0.412)	(0.691)	(0.869)	(0.240)	(0.372)
Observations	1649	1649	1648	1649	1559	1558	1559	1558
Control Mean	0.219	0.129	-0.190	-0.068	0.162	0.098	-0.146	-0.031

Note: The table presents intent-to-treat effects from OLS regressions at midline. Responses of mothers are presented in Column 1-4 and fathers in column (5-8). The dependent variable is difference between the percentage of mothers (fathers) in the corresponding village in at midline who find that 14-15 (Columns 1-2 & Columns 5-6) or 16-17 (Columns 3-4 & Columns 7-8) an acceptable age to marry off a girl, and the mother's (father's) belief about how many other men (women) in the village agree with the respective statements. The dependent variable takes on a value from 0-1. Fixed effects for randomisation strata are included. *P*-values from standard errors clustered at village level (unit of randomisation) are indicated in parentheses. Stars indicate: *** 1 percent ** 5 percent * 10 percent level of significance. The row "Control Mean" multiplied by 100, indicates the average percentage point under- or over-estimation in the control group.

A.1 Group Discussion Script and Questions

A.1.1 Introduction by gender specialist

Enumerator instructions: This is to set the tone for the session. We'll introduce ourselves for the session: "*My name is [insert]. Thank you for taking time out of your day to attend our activity! I am here on behalf of [insert] to host a session for you today. We will watch a short film and hold a short discussion afterwards about the film. We hope you enjoy this session.*"

A.1.2 Discussion questions post-street theatre

"*Now we're going to talk about some things related to what we just watched. This is an informal discussion so please know that there are no right or wrong answers.*" Enumerator instructions: Follow the questions in the script closely. If conversation shifts off topic, gently redirect it back by repeating the question or asking the next question from the script. Do not participate in the discussion. Merely stick to asking the questions.

1. What did you think of the story?

2. What parts of the story resembled real-life happenings in your community? Were there aspects that did not hold true?
3. Which character did you like the most? What about this character was likeable?
4. Which character did you dislike the most? What about this character did you dislike?
5. From what you have seen today, what are the merits of early child marriage?
6. From what you have seen today, what are the costs of early child marriage?
7. What does Pakistani Law say about early child marriage?
8. How can you stop early child marriage?
9. Conclusion: Is there anything else you would like to say about today's activity?

B Model Appendix

Section B.1 to B.5 are copied from Anderberg et al. (2024). The model extension in Section B.6 is developed in this paper.

B.1 Set-up

Consider a set J of households drawn from the same local community. In each household $j \in J$ there is a mother and a father who need to make a decision about when to make their daughter available for marriage. They can either opt for an "early marriage" or a "delayed marriage" and we denote their choice by $d_j \in \{0, 1\}$ where $d_j = 1$ indicates "delay".

Delaying marriage comes with a *potential* benefit, $r_j \in \{0, 1\}$. The realization of a benefit is, however, uncertain and all parents have incomplete information about how likely a benefit to delay is: the probability of a benefit, denoted $\pi_s \in (0, 1)$, depends on a "state of the world" s which is either "low" or "high."All parents hold a common prior over s.

Assumption 1:

(Return to delay, state of the world, and prior). There is a state of the world $s \in S \equiv \{L, H\}$ over which parents share a common prior belief, $\mu \equiv \Pr(s = H) \in (0, 1)$. Any household $j \in J$ may experience a benefit from delayed marriage, $r_j \in \{0, 1\}$, and the probability of a benefit occurring takes the form,

$$\Pr(r_j = 1 | d_j, s) = d_j \pi_s, \tag{4}$$

where $0 < \pi_L < \pi_H < 1$.

For simplicity, we will use P to denote the prior. Consistent with this notation, let

$$\pi_P \equiv (1 - \mu)\pi_L + \mu\pi_H, \tag{5}$$

denote the perceived likelihood of a benefit to delayed marriage based on the common prior.

There is a pre-existing local social norm for the timing of marriage which we can assume reflects a commonly perceived frequency of delayed marriage in the community, $\widehat{d} \in (0, 1/2)$. That is, early marriages are relatively common in the community, defining expected behavior. Household j may experience stigma $z_j \in \{0, 1\}$, and the risk of stigma is larger the more the household's decision is at odds with the norm.

Assumption 2:

(Social norm and probability of stigma). Any household $j \in J$ may experience stigma, $z_j \in \{0,1\}$, and the probability of stigma occurring takes the form,

$$\Pr\left(z_j = 1 | d_j, \widehat{d}\right) = \rho\left(d_j - \widehat{d}\right)^2, \tag{6}$$

where $\rho > 0$ and where the pre-existing norm satisfies $\widehat{d} \in (0, 1/2)$.

As the parents face a binary decision, we assume simple additive preferences over the potential delay benefit r_j and over stigma z_j, and we normalize the utility from benefit. Hence let the preferences of spouse $i \in \{1,2\}$ in household $j \in J$ be given by

$$u_{ij}(r_j, z_j) = r_j - \sigma_{ij} z_j. \tag{7}$$

Parents' preferences are assumed to vary both *across* and *within* households. This variation is captured by the individual stigma aversions $\sigma_{ij} \geq 0$. Hence $\boldsymbol{\sigma}_j \equiv (\sigma_{1j}, \sigma_{2j})$ characterizes the preferences of the two spouses in household j.

Within-household preference heterogeneity implies that, in every household, one spouse will be no more stigma-averse than their partner. Without loss of generality, we can label the least stigma-averse spouse in each household as spouse 1, whereby $\sigma_{1j} \leq \sigma_{2j}$ for all $j \in J$. In line with general patterns of the data (see Table 2), we will refer to the less stigma-averse parent 1 as the *father* and the more stigma-averse parent 2 as the *mother*. This of course is *not* without loss of generality. For this reason, when we outline model predictions below, we will pay close attention to this being a potential misclassification in some households and focus only on *robust* predictions.

Specifically, let the stigma aversions in household j be made up of a household component γ_j and a spouse 2 specific component λ_j, both of which are positive, random and independent.

Assumption 3:
(Stigma aversion) The stigma aversions $\boldsymbol{\sigma}_j$ of the parents in household $j \in J$ take the form $\sigma_{1j} = \gamma_j$ and $\sigma_{2j} = \gamma_j + \lambda_j$ where γ_j and λ_j are independent random variables with cumulative distribution function Γ and Λ respectively, each with support $[0, \infty)$.

Assumption 3 implies that σ_{1j} and σ_{2j} are positively correlated via the common household component.

Whilst all parents hold the common prior belief μ over the state s, individual beliefs may change in light of information provided to them, either directly or via their spouse. To this end, let μ_{ij} denote the belief held by spouse i in household j that $s = H$. Given μ_{ij} and

σ_{ij}, spouse i in household j will want to delay marriage if and only if their expected benefit to delay exceeds any associated increase in expected disutility from stigma,

$$E\left[r_j|d_j=1,\mu_{ij}\right] \geq \sigma_{ij}\left\{E\left[z_j|d_j=1\right]-E\left[z_j|d_j=0\right]\right\}. \tag{8}$$

Using (4) and taking expectation also over the state to rewrite the left hand side using μ_{ij}, and using (6) and (7) to rewrite the right hand side, equation (8) can be written as

$$\left(1-\mu_{ij}\right)\pi_L+\mu_{ij}\pi_H \geq \sigma_{ij}\rho\left(1-2\widehat{d}\right), \tag{9}$$

where we note that, by Assumption 2, $\rho\left(1-2\widehat{d}\right)>0$. Since early marriage is perceived to be the more common choice in the community, any parent will perceive that delaying marriage increases the risk of stigma.

The parents in any given household may disagree on the marriage timing decision. But even then, a decision has to be made. The following assumption captures the notion of "decision power" in the simplest possible way.

Assumption 4:
(Decision power) The marriage timing decision, $d_j \in \{0,1\}$, implemented in household $j \in J$ is the option preferred by spouse 1 with probability α and the option preferred by spouse 2 with probability $1-\alpha$, where $\alpha \in (0,1)$.

This simple formulation allows for the possibility that fathers, for instance, have a larger "final say."But even then, they will want their spouse to agree with them as the mother has a positive say.

B.2 Household preferences

Consider the parents' preferred choices based on the prior belief P or based on full knowledge of the state s being either L or H. Specifically, define the following indicator for whether spouse i in household j prefers delay or not,

$$d_{ij}\left(k\right) \equiv \begin{cases} 1 & \text{if} \quad \sigma_{ij} \leq \underline{\sigma}\left(k\right) \\ 0 & \text{if} \quad \sigma_{ij} > \underline{\sigma}\left(k\right) \end{cases}, \quad i=1,2,\ j\in J \text{ and } k=L,P,H, \tag{10}$$

where, using (9), the threshold stigma aversions are defined by

$$\underline{\sigma}\left(k\right) \equiv \frac{\pi_k}{\rho\left(1-2\widehat{d}\right)} > 0 \text{ for } k=L,P,H. \tag{11}$$

Note that, for any household $j \in J$, the following monotonicities apply,

$$d_{ij}(L) \leq d_{ij}(P) \leq d_{ij}(H) \text{ for } i = 1,2 \quad \text{and} \quad d_{1j}(k) \geq d_{2j}(k) \text{ for } k \in L, P, H. \tag{12}$$

The first simply says that, if a parent prefers to delay at P they also prefer to delay at H, and if they prefer to delay at L they also prefer to do so at P and H, and vice versa. This follows from the fact that, by Assumption 1, $\underline{\sigma}(L) < \underline{\sigma}(P) < \underline{\sigma}(H)$. The second says that, if there is disagreement between the spouses then spouse 1 prefers delay and spouse 2 prefers early marriage. This follows since, by Assumption 3, $\sigma_{1j} \leq \sigma_{2j}$ for all $j \in J$.[33]

B.3 Communication between spouses

Consider now the case where, in each household $j \in J$, one spouse learns the true state s. After learning s, the informed spouse i can communicate with their partner, denoted $-i$, by sending a signal $\widetilde{s}_{ij} \in S$. A *communication strategy* for spouse i is a mapping from the set of (true) states S to the probability of sending the signal $\widetilde{s}_{ij} = H$. This can be represented as a vector $\boldsymbol{\phi}_{ij} \equiv \left(\phi_{ij}^L, \phi_{ij}^H\right)$ where $\phi_{ij}^s \equiv \Pr(\widetilde{s}_{ij} = H|s)$. A natural interpretation of the signal $\widetilde{s}_{ij} = s$ is as spouse i claiming that the state is s, which may be truthful or not. The set of possible communication strategies is $\Phi \equiv [0,1] \times [0,1]$.

Upon receiving the realized signal $\widetilde{s}_{ij} \in S$ from spouse i the partner, $-i$, updates their belief using Bayes' rule and in line with $\boldsymbol{\phi}_{ij}$. Hence the posterior beliefs held by $-i$ that the true state is H will satisfy

$$\Pr(s = H|\widetilde{s}_{ij} = H; \boldsymbol{\phi}_{ij}) = \frac{\mu\phi_{ij}^H}{\mu\phi_{ij}^H + (1-\mu)\phi_{ij}^L}, \tag{13}$$

$$\Pr(s = H|\widetilde{s}_{ij} = L; \boldsymbol{\phi}_{ij}) = \frac{\mu\left(1-\phi_{ij}^H\right)}{\mu\left(1-\phi_{ij}^H\right) + (1-\mu)\left(1-\phi_{ij}^L\right)}. \tag{14}$$

The order of events is as follows. Spouse i first chooses a communication strategy $\boldsymbol{\phi}_{ij} \in \Phi$. The true state of the world s is then revealed to i who sends a signal $\widetilde{s}_{ij}$ to $-i$ based on s and in accordance with $\boldsymbol{\phi}_{ij}$. $-i$ then updates their beliefs as in (13) and (14). Finally, the spouse who gets to have the final say (Assumption 4) implements their preferred option as

[33] One interpretation of the perceived delay frequency $\widehat{d}$ defining the local social norm is as the equilibrium frequency based on the prior. If so, $\widehat{d}$ would satisfy $\alpha \Pr(\sigma_{1j} \leq \underline{\sigma}(P)) + (1-\alpha)\Pr(\sigma_{2j} \leq \underline{\sigma}(P)) = \widehat{d}$. Given the assumed preferences for conforming with the social norm, this equation will generally have multiple solutions. For instance, all households delaying marriage is always an equilibrium. But an equilibrium with $\widehat{d} < 1/2$ is also possible and a community may find itself in such a low delay frequency equilibrium for reasons of path dependency. For instance, if in the past, parents did not perceive any direct benefit from delaying marriage.

the household decision d_j.

We can now consider what communication strategy will be privately optimal for the informed spouse, maximizing their own expected utility, knowing that their partner will update their belief based on it.[34] The following Claim describes the optimal communication strategies for each spouse.[35] It states that spouse i, when privately informed, will report non-truthfully – that is, engage in persuasion – when they themselves prefer one option irrespective of the state of the world, but their non-informed partner's preferred choice depends on their belief, making them "persuadable".

Claim 1:

(Privately optimal communication).

(i) ***Informed Spouse 1 (Father)****: If* $\sigma_{1j} < \underline{\sigma}(L)$*, but* $\underline{\sigma}(L) < \sigma_{2j} \leq \underline{\sigma}(H)$*, then the privately optimal communication strategy for spouse* 1 *is* $\boldsymbol{\phi}^*_{1j} = \left(\phi^*_{1j}, 1\right)$ *where,*

$$\phi^*_{1j} = \min\left\{\frac{\mu}{(1-\mu)}\frac{\left(\pi_H - \sigma_{2j}\rho\left(1-2\widehat{d}\right)\right)}{\left(\sigma_{2j}\rho\left(1-2\widehat{d}\right) - \pi_L\right)}, 1\right\}. \tag{15}$$

For all other $\boldsymbol{\sigma}_j$*, the truthful strategy,* $\boldsymbol{\phi}^*_{1j} = (0,1)$*, is privately optimal for spouse* 1*.*

(ii) ***Informed Spouse 2 (Mother)****: If* $\sigma_{2j} > \underline{\sigma}(H)$*, but* $\underline{\sigma}(L) \leq \sigma_{1j} < \underline{\sigma}(H)$*, then the privately optimal communication strategy for spouse* 2 *is* $\boldsymbol{\phi}^*_{2j} = \left(0, \phi^*_{2j}\right)$ *where,*

$$\phi^*_{2j} = \max\left\{1 - \frac{(1-\mu)}{\mu}\frac{\left(\sigma_{1j}\rho\left(1-2\widehat{d}\right) - \pi_L\right)}{\left(\pi_H - \sigma_{1j}\rho\left(1-2\widehat{d}\right)\right)}, 0\right\}. \tag{16}$$

For all other $\boldsymbol{\sigma}_j$*, the truthful strategy,* $\boldsymbol{\phi}^*_{2j} = (0,1)$*, is privately optimal for spouse* 2*.*

The less stigma-averse parent (spouse 1) will – if anything – persuade their partner to agree to delay marriage. Claim 1 states that spouse 1 will use persuasion when they

[34] It is worth noting that the informed spouse's choice of communication strategy will not depend on the distribution of decision-power within the household. Irrespective of the value of α, an informed spouse will always have an incentive to choose a communication strategy that will make their partner agree with them as often as possible, even if, in the end, the informed spouse themself has the final say.

[35] Note that we do not include any direct costs of disagreement which in the current context could also be interpreted as a disutility from IPV. Adding such a cost would not impact on the equilibrium as the informed spouse is already choosing a communication strategy to maximize the likelihood of their partner agreeing with them. We are also not including any direct cost of misreporting the state ("lying"). As the state is a probability of a return to marriage delay, the realization of the return never reveals the true state.

themself prefer to delay marriage irrespective of the state, but their partner would only want to delay at state H, not at state L. If the partner would prefer to delay marriage based on the prior, then spouse 1 will optimally always claim that the state is H, that is $\phi^*_{1j} = 1$. Doing so conveys no information and ensures that the partner does not update their beliefs and hence continues to support delaying marriage. If, on the other hand, the partner's stigma aversion is stronger – to the point that they would prefer early marriage based on the prior – then spouse 1 cannot make their partner always agree to delay. In this case, spouse 1 reveals state L with the lowest probability, $\phi^*_{1j} \in (0, 1)$, that makes the partner prefer to delay upon receiving the signal H.

By corresponding argument, the more stigma-averse parent (spouse 2) will – if anything – persuade their partner to agree to early marriage. Claim 1 states that spouse 2 will use persuasion when they themself prefer early marriage irrespective of the state, but their partner would only want an early marriage at state L, not at state H. If the partner prefers early marriage on the prior, then spouse 2 would always claim L. If not, they would falsely claim L with the highest probability (and hence reveal H with the lowest probability) that makes the partner prefer early marriage upon receiving the signal L.[36]

In the above Sections B.2 and B.3 we have discussed the mechanics of the model in the context of household preferences and communication over benefits to marriage delay of daughters. The same logic can be applied to other household decisions where spouses have incomplete information benefits and/or costs, where they weigh consequences differently, and where one or both spouses may receive decision-relevant information. Examples may include decisions related to relocation, employment, large purchases and investments etc. The persuasion framework thus generally has many possible applications to household decisions, with or without a focus on disagreements and IPV.

B.4 Effects of information on marriage timing

We now consider what effect information has on marriage timing decisions. By the "effect of information" we mean the change in the probability of a given household delaying marriage compared to when basing their decision on the prior. This depends on what state s is revealed, which spouse receives this information, and on the household's preferences. Formally, the effect of information s, received by spouse $\{i\}$ (possibly both), on household

[36] The scope for persuasion naturally depends on the particular value of their partner's stigma aversion. Specifically, ϕ^*_{1j}, is decreasing in σ_{2j}: the *larger* is spouse 2's stigma aversion, the *less* spouse 1 can persuade their partner to delay marriage. Conversely, $1 - \phi^*_{2j}$, is increasing in σ_{1j}: the *larger* is spouse 1's stigma aversion, the *more* spouse 2 can persuade their partner to opt for an early marriage.

j with stigma aversions $\boldsymbol{\sigma}_j$, is given by,

$$\Delta d\left(s,\{i\},\boldsymbol{\sigma}_j\right) \equiv E\left[d_j|s,\{i\},\boldsymbol{\sigma}_j\right] - E\left[d_j|P,\boldsymbol{\sigma}_j\right], \tag{17}$$

where the expectations are over potential signals sent by the informed spouse and over the final say. As before, we are interested in the population aggregate. Hence, let $\Delta d\left(s,\{i\}\right) \equiv E\left[\Delta d\left(s,\{i\},\boldsymbol{\sigma}_j\right)\right]$ denote the aggregate effect where the expectation is taken over $\boldsymbol{\sigma}_j$.

From Claim 1 we know that, in some households, spouse 2, when privately informed, will persuade their partner not to delay marriage when the state is H. As a result, the aggregate effect of information H is more positive when received by spouse 1 only than when received by spouse 2 only, that is $\Delta d\left(H,\{1\}\right) > \Delta d\left(H,\{2\}\right)$. Indeed, information H, when received by spouse 1 only, will only ever increase the probability of any given household delaying marriage (weakly for all and strictly for some). If the same information is received by spouse 2 only, persuasion efforts by some make the aggregate effect ambiguous. As a consequence, $\Delta d\left(H,\{1\}\right) > 0$ but $\Delta d\left(H,\{2\}\right)$ is ambiguous. Conversely, from Claim 1 we know that, in some households, spouse 1, when privately informed, will persuade their partner to delay marriage when the state is L, leading to $\Delta d\left(L,\{2\}\right) < \Delta d\left(L,\{1\}\right)$. Also, by a parallel argument, $\Delta d\left(L,\{2\}\right) < 0$ whilst $\Delta d\left(L,\{1\}\right)$ is ambiguous.

What if *both* spouses learn the true state? When both are informed about s, each would, if they have the final say, implement their preferred option at that state. As spouse 1 – the less stigma-averse parent – only engages in persuasion at state L, the effect of both learning that the state is H is the same as when spouse 1 only receives this information. Correspondingly, the effect of both learning that the state is L is the same as when spouse 2 only receives this information.[37]

The following claim summarizes the above effects.

Claim 2:

(Effect of information.) The effect of information on marriage delay when:
(i) There is a high probability of a benefit to delayed married, $s = H$, satisfies

$$\Delta d\left(H,\{1\}\right) = \Delta d\left(H,\{1,2\}\right) > \Delta d\left(H,\{2\}\right), \tag{18}$$

where $\Delta d\left(H,\{1\}\right) = \Delta d\left(H,\{1,2\}\right) > 0$ but $\Delta d\left(H,\{2\}\right)$ is ambiguous.

[37]Even though the communication strategies are chosen non-cooperatively, the marriage timing decision is always ex post efficient: the option chosen is always that preferred by at least one spouse given the true state of the world. An information advantage held by one spouse will, in effect, make it more likely that their preferred option is chosen when the spouses' preferred options differ.

(ii) There is a low probability of a benefit to delayed marriage, $s = L$, satisfies

$$\Delta d\left(L, \{2\}\right) = \Delta d\left(L, \{1,2\}\right) < \Delta d\left(L, \{1\}\right), \tag{19}$$

where $\Delta d\left(L, \{2\}\right) = \Delta d\left(L, \{1,2\}\right) < 0$ *but* $\Delta d\left(L, \{1\}\right)$ *is ambiguous.*

B.5 Model robustness and summary of model predictions

The model makes predictions about beliefs and marriage decisions. So far we have duly considered three possible states, L, P, and H, making the assumption that both the information structure and communication structure are ternary. In reality, however, both information and communication can be continuous, and we need to consider the robustness of our predictions relative to non-ternary information and communication structures. Furthermore, we have assumed that fathers are more stigma averse than mothers, to generate the gender differences in preferences that lead to different marriage delay outcomes depending on which spouse is treated. In reality, however, (some) fathers may be more stigma averse than mothers, which will have implications for the robustness of marriage predictions. Finally, the edutainment intervention was designed to emphasize the benefits to delaying marriage, so we need to focus our predictions on the case where the information that is revealed to directly informed parents that the state of the world is H.

Marriage decisions: The equality between the case where the father alone is informed and the case where both are informed is generally not a robust prediction. To see this, consider the possibility that, in some households, the father (mother) is the more (less) stigma-averse parent. If some fathers (mothers) are actually spouse 2 (1), there would be some fathers (mothers) who, when privately informed, would persuade their partners towards early marriage (delayed marriage). This would imply that the effect of positive (negative) information should be no larger when the father alone (mother alone) is informed compared to when both are informed.

In contrast, however, if we allow for a non-ternary information structure, fathers (mothers) – in addition to misreporting that the state is H (L) when in reality they are informed the state is L (H) – would be able to "exaggerate" positive (negative) information when privately informed. This would make the effect of positive (negative) information potentially larger when the father alone (mother alone) is informed compared to when they both are. The robust predictions are then that, in case the state is H, marriage delays increase when fathers are targeted alone, $\Delta d\left(M\right) > 0$, and also when both are targeted, $\Delta d\left(F + M\right) > 0$. These positive effects should both be larger than when mothers are targeted alone, where $\Delta d\left(F\right)$ is ambiguous.

Beliefs: The robust prediction with respect to parental belief updating is that parents should update the same when directly treated, either alone or jointly. There are, however, no robust predictions with respect to the belief updating when parents are indirectly informed, through their spouse. To see this, consider as above the possibility that, in some households, the father (mother) is the more (less) stigma-averse parent. If some fathers (mothers) are actually spouse 2 (1), there would be some fathers (mothers) who would communicate that the state is L (H) when in fact the state is H (L). This implies that belief updating of the mother (father) should be no larger when she (he) is indirectly informed, than when she (he) is privately informed. Similarly, under a binary information and communication structure, misreporting by the father (mother) can only imply communicating that the state is L (H) or P when the intervention informs the father directly (mother directly) that the state is H (L). This implies that the mothers' (fathers') belief updating should also be no larger when she (he) is indirectly informed, than when she (he) is privately informed. In contrast, however, when we allow for the information and communication structure to be continuous, and there is scope for exaggeration, belief updating can be larger when the parents is indirectly informed, then when they are directly informed, either alone or jointly.

B.6 Equilibrium probability of delay

We start by working out the probability of a household choosing to delay for each preference profile, $C_j \in \cup_{k=0}^{9} C^k$, after a mother has been informed that $s = H$ (See Table A1 in Anderberg et al. (2024) for characterization of possible preference profiles).

- **Profiles C^0 to C^1**: At C^0 neither parents wants to delay. At C^1, the mother adopts the uninformative strategy, so this means the father retains his prior at whichher also prefers early marriage. Thus, no marriage delay will occur at C^0 and C^1.

- **Profiles C^2**: At C^2 the mother will not support delay, but the father will do so if she indicates that the state is H. She does with probability ϕ_2^*. We can express the probability that $C_j = C^2$ in terms of the primitive components:

$$\begin{aligned} & \Pr\left(\underline{\sigma}\left(L\right) \leq \sigma_{1j} \leq \underline{\sigma}\left(P\right) \text{ and } \sigma_{2j} > \underline{\sigma}\left(H\right)\right) \\ = \; & \left[\Gamma\left(\underline{\sigma}\left(P\right)\right) - \Gamma\left(\underline{\sigma}\left(L\right)\right)\right]\left[1 - E_{\gamma_j}\left[\Lambda\left(\underline{\sigma}\left(H\right) - \gamma_j\right) | \underline{\sigma}\left(L\right) \leq \gamma_j \leq \underline{\sigma}\left(P\right)\right]\right] \end{aligned}$$

  In this case, delay will happen with the probability $\alpha\overline{\phi}_2^*$, that is, based on this bargaining weight and optimal communication strategy.

- **Profiles C^3**: At profile C^3 the mother and father will disagree irrespective of any information, with the father preferring delay and the mother preferring early marriage.

We can express the probability that $C_j = C^3$ in terms of the primitive components:

$$\begin{aligned} & \Pr\left(\sigma_{1j} \leq \underline{\sigma}\left(L\right) \text{ and } \sigma_{2j} > \underline{\sigma}\left(H\right)\right) \\ = \quad & \Gamma\left(\underline{\sigma}\left(L\right)\right)\left[1 - E_{\gamma_j}\left[\Lambda\left(\underline{\sigma}\left(H\right) - \gamma_j\right)|\gamma_j \leq \underline{\sigma}\left(L\right)\right]\right] \end{aligned}$$

In this case delay will happen with the probability α, that is, based on this bargaining weight since spouses will disagree.

- **Profiles C^4 to C^9**: From Claim 1, at preference profiles $C^4, ..., C^9$ the mother communicates truthfully. Hence both spouses learn that the state is H. In these preference profiles the mother (who is fully informed that the state is H) will support delay. The father at these profiles will also support delay since from Assumption 3 we know that $\sigma_{1j} \leq \sigma_{2j}$ whereby also $\sigma_{1j} \leq \underline{\sigma}\left(H\right)$. We can express the probability that $C_j \in \cup_{k=4}^{9} C^k$ in terms of the primitive components:

$$\Pr\left(\sigma_{2j} \leq \underline{\sigma}\left(H\right)\right) \quad = \quad E_{\lambda_j}\left[\Gamma\left(\underline{\sigma}\left(H\right) - \lambda_j\right)\right]$$

In these 6 cases, delay will happen for certain since both mothers and fathers support delay as they learn at $s = H$.

The new equilibrium probability of a random household choosing delay, $\widehat{d}_{new}$, is the summation of the above cases:

$$\begin{aligned} \widehat{d}_{new} = \quad & E_{\lambda_j}\left[\Gamma\left(\underline{\sigma}\left(H\right) - \lambda_j\right)\right] + \alpha\Gamma\left(\underline{\sigma}\left(L\right)\right)\left[1 - E_{\gamma_j}\left[\Lambda\left(\underline{\sigma}\left(H\right) - \gamma_j\right)|\gamma_j \leq \underline{\sigma}\left(L\right)\right]\right] \\ & + \alpha\overline{\phi}_2^*\left[\Gamma\left(\underline{\sigma}\left(P\right)\right) - \Gamma\left(\underline{\sigma}\left(L\right)\right)\right]\left[1 - E_{\gamma_j}\left[\Lambda\left(\underline{\sigma}\left(H\right) - \gamma_j\right)|\underline{\sigma}\left(L\right) \leq \gamma_j \leq \underline{\sigma}\left(P\right)\right]\right] \end{aligned} \tag{20}$$

B.6.1 Equilibrium under the prior

The equilibrium probability of a random household choosing delay based on the prior was simply

$$\alpha \Pr\left(\sigma_{1j} \leq \underline{\sigma}\left(P\right)\right) + \left(1 - \alpha\right)\Pr\left(\sigma_{2j} \leq \underline{\sigma}\left(P\right)\right)$$

that characterizes $\widehat{d}$ in the baseline no-information case.[38]

Writing this also in terms of primitives, this is equivalent to

$$\alpha\Gamma\left(\underline{\sigma}\left(P\right)\right) + \left(1 - \alpha\right)E_{\gamma_j}\left[\Lambda\left(\underline{\sigma}\left(P\right)\right) - \gamma_j\right] \tag{21}$$

[38]Parents share the same prior belief. Thus, deriving the equilibrium on the prior does not depend on who is informed.

B.7 Comparing Equilibria

Comparing the the new equilibrium (Equation 20), after the mother has learned that $s = H$, to the prior (Equation 21) we now have that delay:

- Will occur for certain at $C_j = C^4$ whereas at the prior it would not occur.
- Will occur for certain at $C_j \in \left\{C^5, C^6\right\}$ whereas on the prior it only occurred with probability α.
- Will occur with probability $\alpha\overline{\phi}_2^*$ at $C_j = C^2$ whereas at the prior it would occur at α (note $\alpha\overline{\phi}_2^* \leq \alpha$).

The *opposing* effect at $C_j = C^2$ that is obtained as the mother is partially successful at dissuading the father, implies that the comparison of the equilibrium is, by nature, ambiguous. It is of course plausibly positive and consistent with the decreased frequency of early marriages in the local population in the long-run.