Coupling Labor Supply Decisions: An Experiment in India

Matt Lowe* Madeline McKelway†

December 2025

Abstract

Joint household decision-making becomes more common as economies develop. Can external intervention advance this transition? We study this question in the context of female employment in India. We randomized which spouse was given a ticket enabling enrollment in a women's weaving job, and cross-randomized the other to receive no information about the ticket, information, or information and discussion with their spouse. Academic and local experts predict information and discussion should raise enrollment. Instead, information had no effect, and discussion reduced enrollment by 50%. We do not find an effect of giving the wife the job ticket rather than the husband. Our results show that intervention towards joint decision-making about an opportunity can lower take-up. ***JEL codes:*** D13, O12, J22.

*University of British Columbia. E-mail: matt.lowe@ubc.ca.

†Dartmouth College. E-mail: madeline.d.mckelway@dartmouth.edu. Many thanks to Daron Acemoglu, Abhijit Banerjee, Nina Buchmann, Esther Duflo, Marcel Fafchamps, Erica Field, Rema Hanna, Anett John, Nathan Nunn, Claudia Olivetti, Gautam Rao, Ben Roth, Simone Schaner, Frank Schilbach, and Heidi Williams for helpful comments and suggestions. Thanks also to Deivis Angeli, Bakul Chugh, Maddy Dutta, Xiang Li, Aditya Madhusudan, Ieda Matavelli, Daniella Rolle, and Ravichandra Tadigadapa for outstanding research assistance. We are grateful for financial support from the Weiss Family Fund. The field experiment received IRB approval from MIT (#1608659930) and IFMR, and was pre-registered in the AEA registry (#0001678). The expert survey with researchers received IRB approval from UBC (#H21-00189) and Stanford (#60413), while the local expert survey received IRB approval from Dartmouth (STUDY00033137).

1 Introduction

Household decision-making processes transform as economies develop. Couples in more developed countries are less likely to have decisions made by the husband, more likely to have decisions made by the wife, and more likely to make decisions jointly (Figure 1). In tandem, development researchers and policymakers often view it as best practice for interventions to involve couples jointly rather than one spouse alone. For example, we find that 26% of intra-household-related experiments in the AEA RCT Registry include joint interventions.[1] This focus on joint interventions also reflects a large body of evidence from developing countries that finds that spouses not only have different preferences, but that their differing incentives can lead to deviations from efficient household decision-making.[2] A number of barriers to efficient decision-making – like information withholding or bargaining costs – could be mitigated through interventions that target both spouses. More generally, involving both spouses could promote family unity and wellbeing, key objectives of policymakers (Ashraf 2024). However, in this paper we show that interventions toward joint decision-making do not work as predicted.

We study household decision-making about female labor supply in India. India has one of the lowest female labor force participation (FLFP) rates in the world, and raising it is a policy priority (Fletcher et al. 2017). Existing literature suggests household frictions may be partly responsible for low FLFP in India. Employment is not concealable; couples have to come to some sort of joint agreement if the woman is to supply labor. However, men in India voice less support for women's employment than women do (Bernhardt et al. 2018; Field et al. 2021; Bursztyn et al. 2023), which could give husbands an incentive to withhold information about job opportunities for their wives. Indeed, there is evidence from other developing settings of spouses strategically withholding information about money and contraception use from each other (Anderson and Baland 2002; Ashraf 2009; Ashraf et al. 2014; Castilla and Walker 2013; Fiala 2018; Zhang 2024b,a). Further, marriages in India are often arranged, which means spouses in young couples may not know each other well and may face especially high bargaining costs (Coase 1960; Riedl 1995; Anderlini and Felli 2001; Coase 2005). Joint interventions could help by making information symmetric or nudging spouses to discuss female labor supply. This paper presents results from a randomized trial evaluating such interventions. We document near consensus among intra-household researchers that our interventions should raise FLFP, yet we find null and negative effects.

[1] Specifically, we identify the pre-registrations that mention "intra-household", "wife", "husband", or "married", along with spelling variants. Of these experiments, GPT-4 codes 26% as involving joint household interventions. See Appendix B for further details.

[2] See, for example: Udry (1996); Duflo and Udry (2004); Ashraf (2009); De Mel et al. (2009); Robinson (2012); Ashraf et al. (2014); Schaner (2015); Hoff et al. (2017); Schaner (2017); Afzal et al. (2018); Ashraf et al. (2020a); Conlon et al. (2021); Buchmann et al. (2023); Anderson and Baland (2002); Castilla and Walker (2013); Fiala (2018); Nyqvist and Jayachandran (2017); Zhang (2024b,a).

Our field experiment took place in rural Uttar Pradesh, where we partnered with India's largest carpet manufacturer. The firm, Obeetee, offers a program to train and employ women as carpet weavers. The program entails four months of paid training in carpet weaving followed by employment. Both training and employment are full-time and take place in all-female weaving centers located in the villages where participants live. We conducted an experiment with married couples in which the wife was aged 18-30. Baseline data reveal female employment is just 13%, over 50% of couples disagree with each other about the appropriateness of women working as weavers, this disagreement more often takes the form of wives being more supportive than husbands being more supportive, and husbands' preferences are more predictive of program take-up than wives, consistent with high male bargaining power.

We randomized how the female employment opportunity was presented to couples. We first printed enrollment tickets for a randomly-chosen 90% of couples identified in a census. Women could not enroll in the program without their unique job ticket. We used a two-by-three design among the couples with tickets, cross-randomizing (i) which spouse was given the ticket and information about the program, and (ii) which of three information sets the non-ticketed spouse was assigned to. Non-ticketed spouses in the first group ($NoInfo$) received no information about the ticket or the program. The second group ($Info$) received information but received it separately from the ticketed spouse. The final group (*Discuss*) received information at the same time as the ticketed spouse, and the two were encouraged to discuss the opportunity for three minutes at the end of the meeting. In all cases, the ticketed spouse was told what information about the opportunity would be given to their spouse. Moreover, any spouse who was given information about the ticket was told that some couples (the 10%) would be surveyed but not receive tickets. As a result, ticketed spouses in $NoInfo$ had plausible deniability – they could plausibly deny the existence of the ticket to their spouse. We analyze only couples with tickets, and our sample includes 495 couples.

Women enrolled in the program by going to their local women's weaving center on enrollment day, about five days after tickets were given. They were required to present their enrollment ticket and to come with their husbands, ensuring that enrollment, like typical labor supply decisions, was a joint household decision that both partners had to condone. Pooling all treatment groups, 15% of ticketed wives enrolled – a low enrollment rate, but non-trivial given that only 13% of wives were employed at baseline.

Our experiment was motivated by a model of household bargaining with frictions. In this world, the wife can only enroll if the household bargains and decides jointly for her to do so, but two frictions may prevent bargaining. The first is information withholding: if only one spouse learns about the opportunity, the informed spouse could prevent the wife from taking it by not telling the uninformed spouse about it. Withholding will happen when the informed spouse is personally opposed to the wife enrolling and when revealing the information would lead the household to decide to enroll. The second friction is

a bargaining cost: we assume that if spouses disagree about whether the wife should enroll, one must pay a cost to start the bargaining process. The cost could represent the disutility from striking up an uncomfortable conversation. What does this model predict in terms of effects of our interventions? First, $Info$ should raise enrollment relative to $NoInfo$ because it prevents a spouse from stymying enrollment by withholding information. This effect should be larger when the ticket recipient is less supportive of women weaving than the non-ticketed spouse. Second, *Discuss* should increase enrollment relative to $Info$. We view *Discuss* as eliminating the bargaining cost, by giving individuals an excuse to bring up the job, for example. The discussion effect should be larger when couples disagree – in the model, bargaining is costless when both spouses want the same thing. Finally, the framework has an ambiguous prediction for the effect of giving the wife the ticket ($WifeTicket$). There would be no effect in $Info$ or *Discuss*, but there could be a positive or negative effect in $NoInfo$; husbands in $NoInfo$ are more likely to want to withhold information given their preferences, but wives who are opposed to enrollment are more likely to need to withhold information given their lower bargaining power. Thus the overall effect of $WifeTicket$ is ambiguous, but it should be larger in couples where the wife is more supportive of enrollment than the husband.[3]

We solicited predictions of enrollment rates in each of our treatment arms from intra-household researchers, focusing largely on individuals who had done intra-household research in developing country settings. Almost all 70 experts we surveyed agree with the predictions of the bargaining frictions model. On average, the experts predicted that enrollment would be 5.5 percentage points higher in $Info$ than in $NoInfo$. There was near-consensus: 90% of experts predicted a positive effect. Likewise, experts predicted that enrollment would be 6.1 percentage points higher in *Discuss* than in $Info$, with near-consensus again: 86% of experts predicted a positive effect. Consistent with our model's ambiguous predictions for the effect of $WifeTicket$, the average expert predicted an effect of just 1.7 percentage points, and there was little consensus on whether the effect would be positive or negative, with just 56% of experts predicting a positive effect.

We also solicited predictions from 241 local experts: male and female weavers working in Obeetee loom centers. We asked these individuals to rank the effectiveness of three different strategies for informing couples about Obeetee's female weaving program in terms of their likelihood of getting women to apply: informing the husband alone, the wife alone, or the two together with discussion. As with researchers, discussion ranked highest among locals, with 68% predicting discussion would be the best of the three approaches and 8% predicting it would be the worst. Even so, locals were less positive about

[3]The primary purpose of the ticket randomization was to generate variation in incentives to withhold information for heterogeneous tests of $Info$. It is possible $WifeTicket$ could have had an effect in $Info$ or *Discuss* through a mechanism outside of our model (e.g., ticketed spouses prevent enrollment by destroying the ticket), but this seemed less likely ex ante.

discussion than researchers: the corresponding statistics for researchers are 76% and 1%. The locals also predicted informing the husband alone would outperform informing the wife, potentially reflecting norms of male authority.

Our findings are contrary to our priors and to those of the experts. First, we find no evidence that preventing information withholding increases enrollment. The effect of $Info$ is -4 percentage points off a base of 18% in $NoInfo$ and is statistically insignificant. We reject the mean academic expert prediction of a 5.5 percentage point increase ($p = 0.01$). Consistent with these results, we find direct evidence of information diffusion: over 70% of non-ticketed spouses in $NoInfo$ reported knowledge of the job ticket on an endline survey done in the week following enrollment. $Info$ did increase ticket awareness, but given the null effects on enrollment, this increase occurs for a set of couples that would not enroll even when information is symmetric.

Second, turning to the effects of *Discuss*, we first note that this treatment was successful in kickstarting a decision process – it increased the total number of discussions about the job couples reported having at endline. The effect was one discussion from a base of about three in the other arms, and roughly 80% of this effect can be explained by discussions couples had during the job information meeting. However, discussion lowered enrollment – by 10 percentage points (or 56%) relative to $NoInfo$, and by 6 percentage points (or 40%) relative to $Info$. We reject the mean academic expert prediction of *Discuss* relative to $NoInfo$ ($p < 0.01$) and relative to $Info$ ($p < 0.01$).

Third, we see no effect of $WifeTicket$, though as noted above our ex ante model had an ambiguous prediction for this effect. We estimate enrollment rates to be 15 and 16 percentage points in $HusbTicket$ and $WifeTicket$, respectively, and the difference is not statistically significant. We cannot reject the mean academic expert prediction of 1.7 percentage points for the effect of $WifeTicket$. In addition, when interacting the treatments, we find no differences in the effects of $Info$ or *Discuss* by $WifeTicket$.

Further, we see limited evidence for the patterns of heterogeneity predicted by the ex ante model. The effect of $Info$ is no different when the less supportive spouse was given the ticket as opposed to the equally or more supportive one. Likewise, the effect of *Discuss* is not more positive – if anything it is more negative – among couples who disagree. In $NoInfo$, the effect of $WifeTicket$ is larger when the wife is more supportive of female weavers than the husband, as the ex ante model predicts. However, this is unlikely to work through the mechanism of hiding posited by the model; the corresponding heterogeneous effect on whether the non-ticketed spouse knows about the ticket is not significant.

Having rejected the bargaining-with-frictions model, we then explore alternative mechanisms that might explain our effects, noting that this part of our paper is more speculative. We find suggestive, though mixed, evidence for an alternative "veto power" model that posits (i) couples enroll if neither spouse vetoes enrollment, (ii) spouses incur a cost if they exercise their veto (e.g. from being disagreeable

to their partner), and (iii) our treatments varied the cost by varying who felt they had a right to make the decision, with ticketed spouses always feeling entitled (meaning a zero cost of vetoing) but joint decision-making interventions like *Discuss*, and to a lesser extent, $Info$, make the non-ticketed spouse feel more entitled. This alternative model rationalizes the lack of information withholding (the ticket-holder in *NoInfo* always feels comfortable vetoing so has no incentive to withhold), and the negative effect of discussion (enrollment is less likely when two people might veto rather than one). Furthermore, it implies a specific pattern of heterogeneity that is the near opposite of what our ex ante model predicts: interventions toward joint decision-making should only reduce enrollment among the couples in which the ticketed spouse is more supportive of women weaving than the non-ticketed one.[4] The data confirm this prediction. Additionally, we find that among couples who disagreed in their support for women weaving, assigning the ticket to the more supportive spouse increased enrollment in $NoInfo$ but had no effect in $Info$ or *Discuss* – just what we would expect if veto power is unequal in $NoInfo$, but closer to equal in $Info$ and *Discuss*.

We note, however, that our evidence for the veto power model is only suggestive. Our experiment was not designed to test for veto decision-making directly, and the heterogeneity results provide only indirect evidence. In addition, we do not see evidence in our endline for another prediction of the veto power model: that non-ticketed spouses should be reported as having more influence over the enrollment decision in *Info* and *Discuss*.

Given the mixed evidence for the veto power model, we also consider alternative explanations for our results. The null effect of $Info$ could be explained by spouses having internalized norms of honesty, or through spouses not believing that they had plausible deniability. Otherwise, we explore the evidence for and against a range alternative explanations for the negative effects of discussion, caveating that our empirical tests are often under-powered. We find four of these alternative explanations to be plausibly operative, either due to a lack of evidence ruling them out or some evidence ruling them in: (i) *Discuss* gave spouses in favor of enrollment too little time to prepare their arguments, (ii) forcing discussion prevented supportive spouses from bringing up the topic at an optimal time, (iii) the presence of a surveyor increased adherence to the norm that women not work, and (iv) *Discuss* made it more difficult for the ticket recipient to (positively) misrepresent information about the job.

Regardless of the exact mechanism in our experiment, how generalizable is our main result that discussion reduces female employment? In a more speculative exercise, we explore the correlation between spousal discussion and female employment in data from the Indian Human Development Survey (IHDS).

[4] In the ex ante model, there is no subgroup for which information or discussion can have a negative effect on enrollment. In addition, in that model the positive effects of discussion are largest among couples that disagree – this is because these are the couples that may avoid bargaining, preventing enrollment. By eliminating the bargaining cost, discussion causes some of these couples to enroll.

We document a robust negative correlation between how often couples discuss work-related topics and whether the wife is employed, conditional on a host of controls including discussion unrelated to work and even household fixed effects. We interpret this as tentative evidence of the external validity of our main experimental result.

The main contribution of our paper is to provide evidence that prompting household discussion about an opportunity can lower take-up. This is contrary to the predictions of experts, who nearly unanimously predicted that *Discuss* would have positive effects. It also has implications for policy. Many anti-poverty policies introduce new opportunities to households – immunizations, agricultural technologies, workfare programs, to name a few – with the hope that households will take them up (World Bank Group 2015; Kremer et al. 2019). We find that nudging households towards joint decision-making can produce a status quo bias against take-up. Whether *Discuss* made decisions worse for households is more ambiguous. In the particular case of our veto model where discussion removes both spouses' veto costs, enrollment will be inefficiently low in *Discuss* relative to efficient bargaining models. But even so, it is ambiguous whether discussion reduces or increases efficiency relative to $NoInfo$ or $Info$. The calculus could change under extensions of the model; if, for instance, there were uncertainty about the net gains from female employment, the status quo bias produced by *Discuss* would reduce experimentation and learning.

Our discussion treatment is unique in the literature in that it was narrowly designed to prompt discussion. Other interventions that encourage household discussion are more intensive, in that they tend to change the nature of those discussions as well – e.g., by providing marital counseling (Boyer et al. 2022), developing interpersonal skills (Ashraf et al. 2020b; Björkman-Nyqvist et al. 2023; Kala and McKelway 2025b), or encouraging joint planning (Donald et al., 2022). These more intensive interventions have found favorable results, at least from the perspective of the policymaker. That said, our more negative results are not without precedent. Most similar is Dean and Jayachandran (2019), who ran a field experiment concurrently to ours with female teachers in India, randomizing whether a surveyor guided a conversation between a teacher and her family members on the pros and cons of her working. While Dean and Jayachandran (2019) cannot reject the null hypothesis of no effect on retention in their smaller sample of 171 teachers, their estimated effect is negative and, at -6 percentage points, similar in magnitude to our *Discuss* effect. There are also parallels with Anderberg et al. (2024), who document increased violence from delivering an edutainment intervention to both spouses rather than either alone, and with Gazeaud et al. (2023), who find a financial training for women is less effective when their husbands can attend. Taken together, the evidence suggests policies that simply nudge households towards joint decision-making can backfire; more intensive interventions may ensure that joint decision-making is productive and shifts outcomes towards policymakers' objectives.

We also contribute to evidence on information withholding in the household. In contrast to existing

literature (Anderson and Baland 2002; Ashraf 2009; Ashraf et al. 2014; Castilla and Walker 2013; Fiala 2018; Zhang 2024b,a), we do not find evidence of strategic information withholding and see direct evidence of information diffusion. One explanation is that in our case the spouse who generally has an incentive to hide (the husband) has more control over the eventual decision and thus less need to hide. However, $Info$ effects are not more positive among couples where the ticketed spouse is less supportive, which would include couples with less supportive women. Our results suggest information withholding may be constrained by factors such as veto decision-making, norms of honesty, or a fear of getting caught.

Our paper also contributes to work on the determinants of female labor force participation in developing countries (see Heath and Jayachandran (2018) for a summary). We build especially on work that explores household constraints to women's labor supply (Bursztyn et al. 2020; Dean and Jayachandran 2019; Heath and Tan 2020; Field et al. 2021; Kala and McKelway 2025b; McKelway 2025a,b), showing that efforts to encourage joint decision-making can lower labor supply. Finally, our work introduces expert and local prediction to work on household intervention. In doing so, we build on the growing effort to predict results in economics (DellaVigna and Pope 2018; DellaVigna et al. 2019; Thomas et al. 2020).

2 Background on Female Labor Supply in Uttar Pradesh

2.1 Gender Norms in Uttar Pradesh

Our study takes place in rural villages in eastern Uttar Pradesh, India. Uttar Pradesh is one of India's poorest states. Reflecting this, the median husband in our experimental sample earns only INR 4,500 ($68) per month. Our setting also features strong adherence to traditional gender norms. Purdah is an important feature of local culture – women veil their faces and stay out of sight of men outside of the family. In a pilot survey of 50 women, 82% said that they practiced purdah, and 86% said that the practice is important. Arranged marriage and patrilocality are the norm: at the time of marriage, women generally leave their natal villages and become a part of a family they do not know well in a new village. Married women face constraints on their physical mobility and are responsible for many household chores, including child-rearing, cooking, tending to livestock, and household cultivation. Whereas 82% of husbands in our experimental sample had done activities to earn income in the previous three months, only 13% of their wives had done so.[5] This norm is strongly ingrained in cultural mindsets, with 87% of

[5]Low levels of women's employment are seen across India, not just in Uttar Pradesh. The country's female labor force participation rate is among the lowest in the world, and low even among countries with similar per capita income (21% among females aged 15+ in 2016, according to the World Bank's World Development Indicators).

men and 84% of women believing that husbands should earn more income than their wives.

2.2 Partner Firm and Women's Job Opportunity

We partnered with Obeetee, one of India's largest carpet manufacturing and exporting firms.[6] Carpet weaving has been an important industry in the region since the time of British rule, when the British set up the industry to take advantage of low labor costs. Weaving is generally considered a low-caste occupation and, as with most formal-sector employees in this setting, weavers are predominantly male.[7]

Obeetee imports wool, usually from Rajasthan, India, to its factory in eastern Uttar Pradesh, where it is converted to yarn. The yarn is then supplied to hundreds of loom owners located in villages in the region. These loom owners operate small loom centers in their villages and employ local men to weave carpets by hand. Obeetee purchases completed carpets from the loom owners, packages them, and ships them to buyers globally.

In recent years, Obeetee has begun a program to train and employ women as weavers. They had several motivations for doing so. A larger pool of labor allows Obeetee to more easily take advantage of periods of high export demand for carpets, while also helping to offset local labor shortages driven by increasing rural-urban migration of male weavers. Further, payments made to women during training count towards government-required corporate social responsibility, which means there is implicit government support for the program.

Each center recruits women living in the center's village. Status quo recruitment tended to involve the firm advertising the job opportunity to existing male weavers, and asking them to spread the information to interested women in their households. In contrast, we experimented with alternative recruitment styles that ensure the diffusion of information to eligible women.

As women usually have no prior experience in weaving, the job begins with a four-month, well-paid training period.[8] By many objective measures, the job is desirable. It is near potential employees' homes, involves safe and comfortable work, requires reasonable hours, and demands no prior training. To respect gender norms, many steps are taken to ensure that women would interact only with other women while at

[6] This is the first experiment conducted as part of an ongoing partnership with Obeetee. Subsequent experiments have also studied Obeetee's female weaving program, but in separate villages from the ones studied here (Kala and McKelway, 2025b; McKelway, 2025b,a, 2022). A final ongoing experiment studies workplace rewards in Obeetee's male and female weaving centers (Kala and McKelway, 2025a).

[7] For example, 76% of Uttar Pradesh textile workers in the 2011 Indian Human Development Survey are male, and 92% are Other Backwards Castes (OBC) or Scheduled Castes (SC). Lower castes are over-represented given that only 72.4% of all Uttar Pradesh respondents are OBCs or SCs.

[8] A woman who attended training every day and met performance targets would earn 87% as much as the median husband in our sample at baseline. Initially, the women were to be paid monthly in cash. However, India's demonetization occurred in the middle of our study, resulting in a rushed transition toward paying directly into bank accounts.

work – only female weavers work in the women's centers and the centers are owned by females, though in practice the husband of the owner is involved with managing the center and the trainer is often male. Qualitative evidence suggests that participants also see the job as desirable: at baseline, 88% of our experimental sample say that workers in this job would be completely safe, and only 11% say that the job is low status.[9] Nevertheless, as we will see below, enrollment and retention rates are low, consistent with low levels of female employment in this setting.

2.3 Preferences Regarding Women's Employment

We use data from our baseline survey ($N = 495$ couples) to describe systematic differences in preferences towards women's employment within the household. Before mentioning the actual weaving position, we asked both husbands and wives, separately, how appropriate they thought it would be for men or women in their household to hold a full-time job outside of the house in three different occupations: construction, weaving, and teaching. Construction and teaching were chosen to reflect the least and most "women-appropriate" occupations in this area,[10] while weaving was chosen to match the actual job opportunity we offer in the experiment. We record answers on a zero to two scale: inappropriate, somewhat appropriate, or completely appropriate. We use this data to establish three core facts that motivate our experiment.

First, wives are more supportive of women working as weavers than husbands. Specifically, while there are no gender differences in the perceived appropriateness of men working in each of the three occupations (columns 1 to 3, Table 1), there is a systematic divergence in the perceived appropriateness of women working (columns 4 to 6). Wives think it would be significantly more appropriate for women in their homes to work in all three jobs than their husbands do, with an effect of 38% of the husband mean for weaving (column 5).[11] Wives also show more interest in the specific job opportunity we advertised than their husbands (column 7). The level of interest is high, with 57% of women reporting being very interested and 20% being somewhat interested. Reflecting the gender gap in preferences, non-enrollment is often explained ex-post as due to opposition from husbands and their family members (Figure A1).[12] Our staff partners at Obeetee also frequently mentioned opposition from husbands as a key constraint to

[9] These questions were only asked to those in the treatment groups that were provided details about the job opportunity at baseline.

[10] Among Uttar Pradesh-based respondents to the 2011 Indian Human Development Survey, 91% of construction workers and 53% of teachers are male.

[11] The divergence is smaller for teaching, at 6% of the husband mean, consistent with it being a less male-dominated profession. The husband mean for each type of job held by a woman is notably lower than when men would hold the job, making the perceived appropriateness of women working on average lower than that for men.

[12] The most common reason provided for not enrolling is that there is no-one else to do household chores. This reason may also reflect the bargaining position of the husband and the husband's lack of support for the job. In particular, had the husband been willing to take on more household chores or ask his family to do so, the wife may have been able to work.

enrollment.

Second, couples often disagree with each other about whether women should weave. 58% of couples give different answers to each other when asked about the appropriateness of women weaving (panel (a), Figure 2). Consistent with the first result, this disagreement takes the form of women being more supportive most (67%) of the time, though this still leaves a sizeable percent of the disagreement (33%) taking the opposite form.

Third, the preferences of both husbands and wives are highly predictive of enrollment, but husbands' preferences are more predictive. Job enrollment is higher when husbands and wives deem women weaving to be appropriate (panel (b), Figure 2), with enrollment at 37% when both spouses consider women weaving to be "completely appropriate," and 1% when both answer "inappropriate." While the preferences of both spouses matter, those of the husband matter more. For example, enrollment is over twice as high when the husband answers "completely appropriate" and the wife answers "somewhat appropriate" than the opposite case. More generally, husbands' preferences are statistically significantly more predictive of enrollment than wives' preferences in specifications with and without controls (Table 2). These results are consistent with husbands having high bargaining power.

Summarizing, (i) wives are more supportive of women weaving than husbands, (ii) spouses often disagree with each other about the appropriateness of women weaving, and (iii) both husbands' and wives' preferences are predictive of enrollment but husbands' preferences are more predictive. These facts motivate our experiment design.

3 Experiment Design

3.1 Timeline

We pre-registered the experiment in the AEA RCT Registry, and summarize deviations in Appendix C. The one important deviation is that we were not able to study the effects of the jobs themselves, due to a lack of over-subscription.

We conducted the experiment in conjunction with the opening of six new weaving centers, each of which had slots for 20 women weavers. The firm gave permission for us to run all recruitment activities for these centers. Recruitment and center openings occurred sequentially from September 2016 to January 2017 (see the timeline in Figure 3).

Census and Randomization. For each center, we first conducted a census of the catchment area. The catchment area was defined by the loom owner as the area from which the firm would have recruited

women in the absence of the study. These areas typically consisted of the entire village surrounding the loom, but excluded high-caste hamlets.[13] Surveyors visited each home in the catchment area and surveyed the household head, asking him or her to list all adults in the household along with their gender, age, marital status, caste, and contact information. A catchment area's census typically took four to seven days. Using the census data, we identified all women in the firm's eligible age range (18 to 30) along with a "pair" for the woman. If the woman was married, the pair was her husband. If not, the pair was her household head. We dropped pairs whenever either the woman or her pair were not available for surveys in the next month. While we included eligible unmarried women in recruitment, our analysis focuses on decision-making in married couples and excludes unmarried pairs. We then randomly assigned treatment for the catchment area at the couple-level, stratifying by hamlet and an indicator for Other Backwards Castes (OBC).[14]

Baseline and Intervention. Following the randomization, we implemented a baseline survey containing our experimental intervention. This period lasted for six to eight days for each center and occurred one to six weeks after the census ended. The baseline involved individual surveys of all eligible women, and separate individual surveys of their pairs. These surveys began with a female surveyor meeting the participant at their home.[15] Before beginning the survey, the surveyor and the participant moved to a private place where they could not be overheard. Each surveyor was randomly assigned a group of participants in a randomly ordered list,[16] subject to the constraint that the two members of each pair were assigned to different surveyors. This constraint served to avoid the possibility of surveyors undoing "plausible deniability" by revealing that a spouse had received a job ticket. Both when setting appointments for the surveys and just before the surveys began, surveyors confirmed that both women and pairs were likely to be available to be surveyed, and only proceeded with the surveys if this was the case.

The baseline survey itself had two parts. The first part was a questionnaire that asked about demographics, employment, and attitudes towards women's employment, as already discussed in Section 2.3. The second part of the survey was our experimental intervention. The intervention varied according to whether, and if so how, the job opportunity was presented to the respondent. We describe each treatment in detail in Section 3.2 below.

[13] High castes tend to not see weaving as a job appropriate for their class and are particularly opposed to women working outside of the home.

[14] The omitted category includes Scheduled Castes (SC), Scheduled Tribes (ST), and "Don't Know". In our experimental sample of 495 couples 54% are OBCs, 44% are SCs, 1% are STs, and 1% don't know.

[15] All surveyors were female because households in this setting would not have been comfortable with women meeting individually with male surveyors.

[16] This randomness ensured treatment assignment was not correlated with surveyor. Indeed, we see in Table A1 that the main results (from Table 3) are robust to the inclusion of surveyor fixed effects.

Enrollment Day. On the day after the baseline survey ended, we hosted an enrollment day. Any woman wishing to enroll in the job was required to come with her pair to the weaving center between 7am and 7pm on that day. We chose the long 12-hour period to help ensure that husbands' working hours did not present a barrier to employment.[17]

Those that enrolled were also required to present unique enrollment tickets given to one member of the pair during the baseline survey, as described further below. The requirement for women to attend with their pairs was supported by our partner firm, and is important, for both practical and conceptual reasons. Practically, it eliminated scenarios in which a woman would enroll without her husband's permission, thereby reducing dropout and subsequent intra-household discord. Conceptually, it means that we can interpret enrollment as a decision made jointly by the household that both partners had to condone. They also had to condone it *publicly* given the loom centers were within villages and not hidden, but this also means women's participation in the job – like most jobs in the setting – was not concealable and hence the public nature of enrollment helps to make it indicative of labor supply decisions.

Endline. Finally, we conducted an endline survey in the three to five days following enrollment. The purpose of this survey was to help us understand how enrollment decisions had been made. To this end, the survey included a quiz about job information to assess participants' knowledge of the job along with questions about the pair's decision-making process. 91% of respondents completed the survey by phone, while the remaining 9% completed the survey in person after we were unable to reach them by phone. For the final three centers we added several questions to the endline survey and amended our AEA pre-registration to reflect this change. As a result, we are missing data on the added questions for 45% of our endline respondents.

3.2 Treatments

Plausible Deniability. We began by printing enrollment tickets. Each ticket had the names of a particular pair written on it along with a unique identification number (Figure A2). No woman could enroll without her unique ticket.[18] Crucially, we only printed tickets for a randomly chosen 90% of eligible women.

[17] Consistent with this, when we regress enrollment on treatments, modified strata, and controls, the coefficient on the indicator for whether the husband is employed at baseline is -1.3 pp and statistically insignificant ($p = 0.76$). This suggests that work constraints do not reduce take-up, particularly given that we might expect this coefficient to be negative for other reasons, like income effects.

[18] The enrollment ticket requirement was strictly enforced. The survey form used to record enrollment would not proceed to confirm couples' intention to enroll unless they presented an enrollment ticket with names and ID clearly visible. Prior to

The remaining 10% of pairs still completed baseline surveys, but following the survey, a randomly chosen member of the pair was told that the eligible woman in the pair had not received a ticket. Any participants in the 90% assigned to receive information about the ticket (via randomizations detailed below) were also told that some participants would be surveyed and not receive tickets. It was therefore common knowledge that we had not printed tickets for all eligible women and that being surveyed did not signal that a ticket had been printed. This system ensured plausible deniability: whenever only one spouse was informed about and given the ticket, he or she could plausibly deny having received the ticket. Non-ticketed pairs exist only for this purpose and our analysis focuses on pairs that actually received tickets.

Assigning a Ticketed Spouse. The delivery of the ticket and job information to married couples varied according to two, cross-randomized couple-level treatments (Figure 4).[19] In some cases, there are multiple treated couples per household – our analysis sample of 495 couples reside in 444 households.[20]

The first treatment determined which spouse would receive the ticket: 50% of the time the ticket was given to the wife, and 50% of the time to the husband. Anyone who received the ticket received full information about the job details, the enrollment process, and what their pair would be told by a surveyor.

Information Given to the Non-Ticketed Spouse. The second, cross-randomized treatment determined what information the non-ticketed spouse was told by a surveyor about the ticket and job opportunity.

No Information. In one third of couples, the non-ticketed spouse was told nothing about the job during the baseline survey. Importantly, ticketed spouses in $NoInfo$ could withhold information about job eligibility if they desired; a surveyor would never tell their non-ticketed spouses that a ticket had been given and the ticketed spouse could plausibly deny having received a ticket because 10% of women did not have tickets printed. Because tickets were required to enroll, such withholding was a means by which one spouse could prevent enrollment.

confirming enrollment intention, the survey form also verified the couple who had come to enroll matched the couple whose ID was on the ticket, either by asking the surveyor to confirm a photo taken of the husband at baseline matched the man who had come to enroll, or by ensuring phone numbers provided at enrollment matched phone numbers on record. None of the submitted enrollment forms indicate a couple tried to enroll without a ticket or with another couple's ticket, suggesting the ticket requirement was well-understood.

[19]We followed a simpler procedure for unmarried couples as they were to be excluded from the analysis. In a randomly chosen 50% of pairs, the eligible woman was notified of whether a ticket had been printed for her. If so, she also received the ticket and job details. In the other 50%, the household head received this information. The non-ticketed member of the pair received a baseline survey but no further information. This corresponds to the *NoInfo* row in Figure 4.

[20]Our main results are similar, though noisier, when we restrict to the 397 couples who are the only couples from their households in our analysis sample (Table A2).

Information. In another third of couples, the non-ticketed spouse was told that their pair had or would receive an enrollment ticket along with details of the job and enrollment process. A priori, we expected any effect of this group would be driven by knowledge of the ticket's existence rather than knowledge of the job and enrollment details as ticket information is specific to an individual couple, the unit at which treatment was assigned, whereas job and enrollment details could spread across households in a village. However, we used an additional randomization to allow for the possibility that providing job and enrollment details would have a treatment effect beyond the effect of providing information about the ticket alone. In particular, we split the $Info$ treatment in two: in 50% of $Info$ couples (one-sixth of the full sample) the non-ticketed spouse was told only that their spouse had or would receive an enrollment ticket ($TicketInfo$), while the rest of non-ticketed spouses in $Info$ couples received this information plus job and enrollment details ($FullInfo$). Non-ticketed spouses in $FullInfo$ have greater endline knowledge of the job than those in $TicketInfo$ and view the job as more desirable at endline, but the effect on enrollment is not statistically significant and points in the opposite direction as the effect on job desirability. These effects of $FullInfo$ tend to be more negative under $WifeTicket$, but not significantly so (Table A3).[21] Given this, to increase the power of our tests for information withholding, we combine these two subgroups into a single $Info$ group for the analysis (though for completeness, we also report the expert predictions and core enrollment results without any pooling in Appendix Figures A3 and A8). In contrast to $NoInfo$ couples, both spouses in $Info$ couples knew that the wife could enroll in the job and this information could not be withheld. That said, spouses in $Info$ were told about the opportunity separately and the intervention did not directly affect how they interacted with one another.

Discussion. The final third of couples were assigned to the *Discuss* group. Here, the non-ticketed spouse was present while the job details were given. Surveyors then paused the survey for a full three minutes and encouraged the couple to discuss the opportunity together (exact script in Appendix D). Two surveyors were present during the discussion but remained silent and did not provide any additional information about the job. After the three minutes were up, the surveyor handed the enrollment ticket to the ticketed spouse. The goal of the discussion treatment was to alleviate possible costs to bargaining. While couples in both $Info$ and *Discuss* had full information about the opportunity, only the *Discuss* treatment directly affected how spouses interacted with one another about the opportunity.[22]

[21] With 84 couples in $TicketInfo$ and 87 couples in $FullInfo$, a limitation here is that we lack power to detect small effects on enrollment.

[22] A key logistical difference between this treatment and the others is that husbands and wives had to be together when the job information was given, but separate when taking the baseline questionnaire. Initially, we scheduled husband and wife surveys simultaneously, separated the spouses to individually take the baseline questionnaire, and then brought them back together for the job information and discussion. However, this procedure could have introduced selective attrition as

According to surveyor evaluations, 80% of *Discuss* couples discussed the opportunity during this time, with the other couples remaining silent or discussing something else. Reflecting patriarchal norms, the husband was more likely to talk the most (39% husband, 12% wife, 49% both equally), and more likely to have initiated the conversation (50% husband, 20% wife, 30% both equally). Surveyors saw no clear gender difference in who showed more job interest during the discussion (28% husband, 26% wife, 47% both equally), suggesting husbands veiled their opposition in the surveyors' presence. Surveyors did not report any of the conversations as being argumentative.

3.3 Conceptual Framework

Our experiment was motivated by a simple model of household bargaining with frictions. We formalize the model in Appendix E and describe it in words here. In the model, a husband and wife decide whether the wife should take the job opportunity. She will only take the opportunity if the two bargain and decide jointly for her to do so. The bargaining depends on the net utility gain to each spouse from the wife enrolling – these utility gains could include utility from extra income, shifts in bargaining power, stigma costs, or disutility of the wife's effort – and spouses have perfect knowledge of their own and their spouse's net utilities (though the assumption of perfect knowledge is not necessary for the model predictions we emphasize).

We assume two possible frictions to efficient bargaining: incomplete information and bargaining costs. Incomplete information constitutes a friction whenever only one spouse is aware of the job opportunity. In this case, the knowledgeable spouse will withhold information when two conditions are met: (i) enrollment is costly to them, and (ii) enrollment would happen if the information was revealed. Given these conditions, the spouse that is more supportive of the job will never withhold information. Bargaining costs are at play even if knowledge is symmetric. Here we assume that there is a cost of bargaining whenever the two spouses disagree. If neither spouse pays the bargaining cost (which we can think of as

(a) the ease of participating differed across treatment groups, and (b) participants could have inferred their treatment status in advance based on how their appointments were set. To address this, we modified the procedure when we were about one third of the way through the sample, at which point the experiment was ongoing in the third center's village. Under the new procedure, spouses were contacted individually to complete the baseline questionnaire in the same manner in which individuals in all other treatment groups were contacted. After the questionnaire was complete, the participant was told that surveyors had limited time to complete all surveys and therefore wanted to complete the second half of the survey with the participant's spouse present once the spouse had taken the individual questionnaire. All couples that completed individual baseline questionnaires were able and willing to schedule a second joint appointment. As it turns out, we do not see evidence for the selective attrition we worried about – *Discuss* did not affect the rate of attrition in the one-visit group, and we see balance on baseline characteristics among non-attritors in that group (Table A7, with the corresponding analyses for the two-visit group in Table A8). One remaining concern is that the postponing of job information meant that, on average, *Discuss* couples received job information a bit closer to the enrollment date. We show in Section 6.8.4 (and Table A17) that this does not explain our negative effects of discussion.

starting an uncomfortable discussion about the job), we assume that the couple reverts to non-cooperative decision-making, preventing enrollment.

Our experimental treatments shut down each bargaining friction. The information treatment ensures that both spouses are informed about the job, shutting down the incomplete information friction. The discussion treatment encouraged spouses to enter into a discussion about the job, which we think of as reducing the bargaining cost, by providing an excuse for bringing up the topic, for instance. As a result, the model makes two unambiguous predictions: First, enforcing common knowledge in *Info* increases enrollment, by preventing the strategic withholding of information. This effect should be driven by couples in which the ticket recipient is less supportive of women weaving than the non-ticketed spouse. Second, encouraging bargaining in *Discuss* increases enrollment, by enabling joint decision-making. Importantly, in the model the positive effects of discussion are driven entirely by the couples that disagree about the job – for the couples that agree, there is no bargaining cost, and so no friction preventing bargaining.

Otherwise, the model is ambiguous on the effect of $WifeTicket$. There is no mechanism in the model for an effect of $WifeTicket$ in the $Info$ or *Discuss* conditions, but there could be a positive or negative effect in $NoInfo$. On the one hand, husbands in $NoInfo$ are more likely than wives to want to withhold information given their preferences, which would tend to produce a positive effect from $WifeTicket$; on the other, wives who are opposed to enrollment are more likely than husbands to need to withhold information given their lower bargaining power, which would tend to produce a negative effect from $WifeTicket$. Thus the overall effect of $WifeTicket$ is ambiguous, but it should be larger in couples where the wife is more supportive of enrollment than the husband. Conceptually, the primary purpose of the ticket randomization was to generate variation in incentives to withhold information for heterogeneous tests of $Info$. It is possible $WifeTicket$ could have had an effect even in $Info$ or *Discuss* through a mechanism outside the model (e.g. ticketed spouses prevent enrollment by destroying the ticket), but this seemed less likely ex ante.

4 Expert Surveys

We summarize our two expert surveys in this section, and provide further details in Appendix F.

4.1 Researchers

Following DellaVigna et al. (2019), we elicited the predictions of 361 intra-household experts in May 2021 to allow us to characterize how our results compare with expert priors. Since the first draft of this

paper circulated in May 2017, our expert survey screened out any researchers that report having seen the paper and being able to remember at least some of the results. 70 experts (19% of those contacted) got past the screening questions and actually gave predictions. The 70 experts for which we have predictions skew somewhat more junior than the full sample of 361 experts (Table A4), though they are not inexperienced overall, with the average expert having 3,507 Google Scholar citations.

When eliciting predictions, we revealed the enrollment rate for one treatment cell, writing *"In the first treatment cell, i.e., when the husband got the ticket and the wife was given no information, 19% of women enrolled.* ***What percentage of women in the other seven treatment cells do you think enrolled?"*** The experts made seven predictions, rather than five (recall Figure 4), since we elicited predictions separately for the two sub-treatments of the $Info$ treatment group. Since we pool these two sub-treatments in the analysis, we also pool the expert predictions, by defining an expert's prediction for the pooled $Info$ treatment group as the average of their predictions for the two sub-treatments.[23] We did not incentivize the predictions given that experts might have been able to find the results from a previously circulated draft. Related, we removed the draft of the paper from each of our websites while the survey was live.

Experts agree with the two main predictions of the bargaining frictions model (Figure 5). They expect information withholding, predicting that enrollment will be 5.5 percentage points higher in $Info$ than in $NoInfo$. These predictions reflect near-consensus: 90% of experts predict that $Info$ will increase enrollment relative to $NoInfo$. Second, experts predict that *Discuss* will have 6.1 percentage points higher enrollment than $Info$. Experts are again nearly all in agreement: 86% expect *Discuss* to have higher enrollment than $Info$.

While the bargaining frictions model is ambiguous on whether husbands or wives should withhold information more often, experts expect husbands to be more secretive. In particular, while experts predict the $Info$ treatment to increase enrollment regardless of which spouse received the ticket, they expect $Info$ to increase enrollment by 2.5 percentage points more when the husband receives the job ticket.

4.2 Locals

We also solicited the predictions of local weavers to investigate whether understanding of local cultural norms might generate more accurate predictions of our treatment effects (Thomas et al. 2020; Thomas and Markus 2023). Given the limited education in our setting, we used a simplified version of the expert survey, which only sought to assess predictions for the effects of the ticket recipient and discussion.

The "local experts" are part of a subsequent experiment conducted with weavers in male and female loom centers affiliated with Obeetee (Kala and McKelway, 2025a). In this experiment's baseline survey,

[23] Nevertheless, we report the unpooled expert predictions in Figure A3.

we asked 241 weavers (49% female) to imagine that Obeetee was recruiting women for its female weaving program in a village nearby, and that a female surveyor would go door-to-door in the week before enrollment to inform married couples about the opportunity. There were three options for how she would inform them: inform the husbands alone, inform the wives alone, or inform the two together and encourage them to discuss the opportunity together during the meeting. Respondents were then asked to rank the three options from most to least effective at getting couples to enroll. An additional 454 weavers were asked to make predictions under alternative scenarios, where the surveyor would not be present for the discussion and/or the couple was making a decision about tutoring for their daughter; we consider these predictions in the mechanisms and external validity sections below.

Like the researchers and our ex ante model, the local experts predict the discussion intervention to be most effective (Figure 5). The average local gave discussion a ranking of 1.36 on a 1-3 scale, where 1 is most effective; 68% of the locals said it would rank in first place and 8% said it would rank last.[24] In contrast to the researchers, the locals expect informing the husband alone would be more effective than informing the wife alone, perhaps reflecting local gender norms of male authority.[25]

While the local experts were positive about the discussion treatment, they were less positive than the researchers. We construct the researchers' implied rankings by comparing their predicted enrollment rates in various arms, after averaging predictions across arms to resemble the three interventions the locals were asked to rank. Specifically, we compare: the average predicted enrollment in $NoInfo$ and $Info$ under $HusbTicket$, the average predicted enrollment in $NoInfo$ and $Info$ under $WifeTicket$, and the average predicted enrollment in $HusbTicket$ and $WifeTicket$ under $Discuss$. We find 76% of researchers ranked discussion best and 1% ranked it worst, in contrast to 68% and 8% among locals. Given what we ultimately find, these statistics show that locals are marginally better calibrated than researchers in terms of their predictions for the effects of discussion.

5 Empirical Specification and Sample

5.1 Empirical Specification

Our basic empirical specification takes the form:

$$y_i = \beta_1 Info_i + \beta_2 Discuss_i + \beta_3 WifeTicket_i + \gamma \mathbf{X}_i + \alpha \mathbf{MS}_i + \varepsilon_i$$

[24] Ties are assigned intermediate values – e.g. if a respondent said informing the husband and discussion were tied for first place, that respondent's ranking of each intervention is assigned as 1.5. The 68% and 8% statistics do not count tied options as best or worst.

[25] The local expert predictions do not differ significantly by respondent age or gender (columns 2 and 3, Table A5).

where i denotes a married couple, and $Info_i$, $Discuss_i$, and $WifeTicket_i$ are indicators for the information, discussion, and wife ticket treatments. We additionally estimate a fully saturated specification, by interacting $Info_i$ and $Discuss_i$ with $WifeTicket_i$, as well as heterogeneity specifications, including interactions between the treatments and baseline variables, to test theoretical predictions and explore mechanisms. Our main outcome is an indicator for the couple enrolling on enrollment day, though we consider additional outcomes for supplementary and mechanisms analyses and describe these outcomes as we present effects on them. We test how effects on enrollment compare to the null of no effect and also to the average prediction of researchers (the local experts did not predict magnitudes).

The vector $\mathbf{X}_i$ includes the 12 baseline survey variables used below for balance checks. We set any missing values in the baseline variables to the variable average to avoid dropping observations from the regressions. We report all estimates with and without $\mathbf{X}_i$. $\mathbf{MS}_i$ are modified strata controls. The randomization was stratified by village, hamlet, and OBC status. However, due to attrition, some strata do not have all six treatment groups represented. Additionally, when investigating how effects differ across subgroups, some strata do not have members from each treatment cell within subgroups. To make use of all observations in identifying treatment effects, we do not control for strata fixed effects, but rather control for village fixed effects and for OBC status, which we call modified strata.[26] The probability of treatment assignment did not vary by strata, so exactly how we control for strata does not meaningfully affect our results. We estimate robust standard errors throughout.

5.2 Sample Size and Characteristics

We identified 817 married women that were eligible for the job from the census survey. We printed tickets for 732 (90%) of them. The 817 women and their husbands formed the sample of married couples to be approached for the baseline survey and intervention, but we focus now only on couples with tickets.

Attrition. We consider couples to have attrited from the study if either spouse could not take the baseline survey. The requirement that both spouses be surveyed raised attrition rates but was needed because our treatments would be difficult to interpret if only one spouse was approached in some couples. Overall, 32% of the 732 couples attrited from the study. While 32% may seem high, we estimate that, at most, only 24% of the attriting couples could have known their treatment status.[27] This implies that attrition was not affected by knowing treatment for the majority of the 32%, and that the rate of attrition amongst couples

[26] For a similar reason, we do not include dummy variables for missingness of each baseline control variable, instead imputing missing values to the mean.

[27] Attriting couples could have known their treatment status if (a) one spouse took the baseline survey but the other did not, or (b) the couple was assigned to the discussion treatment with our initial protocol for setting appointments (see footnote 22).

that could have known their treatment status was at most 10%. Consistent with this, attrition does not differ significantly by treatment group (column 1, Table A6), and as discussed below, we observe balance on baseline variables among the couples that did not attrit. The vast majority (75%) of attrition was due to one or both members of the couple being out of town, away all day for work, or otherwise unavailable during the six- to eight-day baseline period (panel (a), Figure A5). The two other leading causes for attrition were surveyors running out of time to complete baseline surveys for one or both members of the couple (11%),[28] or one or both members of the couple not consenting to take the survey (9%). This leaves us with a sample of 495 married couples for analysis. Figure 4 provides the number of couples in the sample of 495 assigned to each treatment group.

Baseline Balance. Restricting to the analysis sample of 495 couples, we regress each of 12 baseline variables on indicators for $Info$, *Discuss*, and $WifeTicket$, along with modified strata controls (columns 2 to 13, Table A6). Four of the 36 coefficients, and two of 12 tests of $Info = Discuss$ are statistically significant at the 10% level, which is close to the amount of imbalance we would expect from random chance alone. That said, we also use randomization inference omnibus tests to test for the joint significance of $Info$, *Discuss*, $WifeTicket$, and $Info = Discuss$ (following Kerwin et al. (2024)). One of four, the test for *Discuss*, is statistically significant at the 10% level. In addition, some specific imbalances look potentially important.[29] We see imbalances in whether women had worked for income in the last three months, but importantly, this variable does not predict our main outcome: enrollment (Table A9), and our negative effects of discussion are not driven by working wives (Table A10). While we see husbands in the *Discuss* treatment group are more supportive of female weavers, we note that any resulting bias would be towards a positive effect of *Discuss* on enrollment, while we find a negative effect. Given the evidence, our best guess is that the balance issues are due to an unlucky draw. We proceed by estimating effects with and without the set of 12 controls, and our core conclusions are robust to either approach.

Endline Attrition and Balance. We attempted the endline survey with the 990 married adults (495×2) in the analysis sample. 830 (84%) completed the endline survey. 87% of endline attrition was due to the adult being unreachable over the phone or in person (panel (b), Figure A5). 10% of attrition was due to

[28] Enrollment dates were set in advance and therefore imposed a hard deadline on when baseline surveys had to be completed. This meant that we could not guarantee canvassing of all those eligible within the time allotted for a center's recruitment activities. To help maximize the size of our experimental sample, we prioritized married couples in the random survey order so that if surveyors ran out of time it was unmarried participants that were excluded. 96% of couples excluded from the study because surveyors ran out of time came from the first center's village, before we had precise estimates of how long baseline surveys would require and before we had implemented a system of appointment setting.

[29] The imbalances are not driven by our initial appointment-setting protocol for the discussion treatment, as discussed in footnote 22 above – when we split the balance checks by the two protocols, we do not see any evidence of imbalances being concentrated among the sample exposed to the first protocol (Tables A7 and A8). If anything, we see the opposite.

an adult not consenting to participate in the endline survey, and 4% was due to the participant not having a phone.[30] Attrition from baseline to endline is largely balanced across treatment groups (columns 1 and 2, Table A11), and is not selective: the sample that completed the endline look similarly balanced to the larger sample that completed the baseline, with the same handful of imbalances (columns 3 to 14, Table A11).

6 Main Results

6.1 Enrollment Rates

Before exploring our treatment effects, we provide descriptive evidence on enrollment rates. 15% of women in our sample enrolled. While 15% is low, it is substantial in the context of female employment rates in our setting; only 13% of women had done *any* work for income in the three months preceding baseline. Figure A1 presents the reasons for non-enrollment provided at endline by husbands and wives who had not enrolled. The most common reasons are that there would be no-one to do household chores, that the wife is pregnant or has children, that the wife's family members would not allow it, or that the household does not like this particular job (weaving). Thus the low take-up appears largely driven by constraints to women's employment in general as opposed to a dislike of the job the firm was offering.

Enrollment rates vary substantially – from 4% to 38% – across the six villages. Some of this reflects variation in the time of year recruitment was done; in the village with the lowest enrollment rate, recruitment coincided with an agricultural season that relies heavily on female labor. Variation may also reflect differing attitudes towards weaving, Obeetee, the local loom owner, or female employment in general. Indeed, the village with the highest enrollment rate is very close to a large Obeetee factory that employs many men in the village. Likewise, village-level enrollment is tightly positively correlated with village-level average responses to the question on appropriateness of women weaving (Figure A6), and these village-level attitudes are strongly predictive of a given wife's enrollment decision even after controlling for that wife's and her husband's own attitudes (Table A12). These patterns are consistent with a role for social norms.

[30] This was only a reason for attrition for the first two centers as starting from the third we conducted endline surveys in person for those that did not have phones.

6.2 Effects of $Info$

We first test whether spouses strategically withhold information to prevent enrollment. Including both couples where the wife was ticketed and those where the husband was, there is no evidence that providing information to the non-ticketed spouse increases enrollment (columns 1 and 2, Table 3). The point estimates suggest that the $Info$ treatment, if anything, reduced enrollment by three or four percentage points, relative to an enrollment rate of 18% among couples assigned $NoInfo$. We can reject the mean expert prediction of a 5.5 percentage point increase with 95% confidence. Contrary to the theory and the experts, there is no evidence that preventing information withholding increases enrollment, suggesting that there is no strategic withholding to begin with.

The lack of withholding might of course mask heterogeneity by gender – husbands may be more likely to withhold information than wives, as the experts predicted. However, the effects of $Info$ are similar whether we consider couples where the husband got the ticket or those where the wife did (columns 3 and 4, Table 3). In the case of ticketed husbands, informing the wife reduces enrollment insignificantly by four percentage points, and in the case of ticketed wives, informing the husband reduces enrollment insignificantly by three to four percentage points. We cannot reject that these two effects are equal ($p = 0.98$ with controls).[31] Neither wives nor husbands appear to be withholding information.

To test directly for information withholding, we use an indicator for the non-ticketed spouse knowing at endline that a ticket had been given to their spouse. We asked this question to adults in couples that did and did not receive tickets, phrasing the question in a way that would not give away the correct answer. In particular, the surveyor first asked whether the participant themself was given a ticket when surveyed a few days prior, and then asked whether the participant's spouse had received a ticket when surveyed. Consistent with $Info$ not increasing enrollment, information diffusion is high. In the $NoInfo$ treatment, 72% of non-ticketed spouses are aware their spouse received a ticket (Table A13). The $Info$ treatment increases this awareness by ten percentage points, but given the non-positive effects on enrollment, they suggest that information withholding only occurs when that information would not be decision-relevant – i.e. upon sharing, the wife would still not enroll. This effect of $Info$ does not vary by $WifeTicket$; husbands and wives disclosed information at similar rates.

[31] We also cannot reject the mean expert prediction that the effect of $Info$ will be 2.5 percentage points larger when the husband gets the ticket than when the wife does ($p = 0.76$ with controls). It follows that while the experts mispredict the overall effect of $Info$, their prediction for the differential effect by gender is better calibrated. Even so, a predicted treatment effect of 2.5 percentage points is small enough that even if the true effect is zero, our study is not well-powered to reject it.

6.3 Effects of *Discuss*

The bargaining frictions theory predicts that discussion could kickstart bargaining, thereby increasing enrollment. We first note that *Discuss* was successful in increasing the amount of intra-household discussion about the job, as proxied by the number of discussions. This is measured using a question at endline which asked the respondent how many times they discussed the job opportunity with their spouse. For those in the *Discuss* group, we asked for the number of times the two discussed the opportunity excluding any discussion when they were informed about the job, and then added one to this number for the 80% of couples whose surveyor reported they discussed the opportunity in the designated time. We find *Discuss* significantly increased the number of discussions, raising the number by about one from a total of roughly three in the other treatment arms (Table A13). This suggests that our discussion treatment was successful at kickstarting a decision process and did not crowd out discussions the couple had on their own.

However, we find significant, negative effects on enrollment (Table 3). *Discuss* lowers enrollment by eight to ten percentage points relative to $NoInfo$ ($p < 0.01$ with controls), and lowers enrollment by five to six percentage points relative to $Info$ ($p = 0.06$ with controls). Given that enrollment rates are low overall, these are large percentage changes, representing drops of 44 to 56% relative to $NoInfo$ and 33 to 43% relative to $Info$. We easily reject the mean expert prediction of *Discuss* relative to $NoInfo$ and relative to $Info$ ($p < 0.001$). Finally, we note that the effect of discussion does not differ significantly by the gender of the ticket recipient, though our ex ante model did not include an interaction effect and the interactions implied by the predictions of the average expert are close to zero.[32]

6.4 Effects of $WifeTicket$

Finally, we see no overall effect of giving the ticket to the wife versus the husband (Table 3). Pooling across the information treatments, 15% of couples in which the husband was ticketed enrolled. The rate is zero to one percentage point higher when the wife was given the ticket, and these differences are not significant ($p = 0.83$ with controls). Our ex ante model had an ambiguous prediction for this effect, and the average expert predicted an effect of just 1.7 percentage points, which we cannot reject. As discussed in the preceding sub-sections, we see no significant interaction effects between $WifeTicket$ and $Info$ or *Discuss*.

[32] The predictions are -2.1 for *Discuss* $\times$ $WifeTicket$ and 0.4 for *Discuss* $\times$ $WifeTicket$ - $Info$ $\times$ $WifeTicket$.

6.5 Main Results Visualized

Figure 6 summarizes these results, visualizing enrollment rates for each of the six treatment cells, and also testing for differences relative to the status quo group (the $NoInfo$, $HusbTicket$ group) and relative to the expert predictions.[33] This figure visualizes the monotonic decline in enrollment rates as the non-ticketed spouse is brought into the decision, regardless of who the non-ticketed spouse is (husband or wife), though we lack power to detect differences between the $NoInfo$ and $Info$ arms. It also highlights how our results are virtually the opposite of what expert researchers predicted – they expected monotonic increases as the non-ticketed spouse was involved, regardless of who was ticketed, and we can reject four of five expert-predicted treatment effects in this figure at the 10% level or lower. Note that while the *Discuss* bar in this figure is lower under $HusbTicket$ than $WifeTicket$, we saw in Table 3 that the effects of *Discuss* did not differ significantly by $WifeTicket$, and summing coefficients in that table, we see no significant effect of $WifeTicket$ among *Discuss* couples ($p = 0.42$ with controls).

6.6 Did Effects on Enrollment Translate Into Effects on Participation?

Our main outcome is enrollment in the program, a revealed preference measure of women's labor supply. But enrolling in the program is of course not the same as actually participating. We collected data on whether women in our sample were in the program at the end of the first month of training.[34] Dropout rates are substantial, at around 50% in the first month. Figure A7 presents data we gathered on reasons for dropout in the first month. As with non-enrollment, dropout appears largely driven by general constraints on female employment rather than issues with this job in particular; the most common reason for dropout was household chores, provided 30% of the time, while only 12% said they left because they disliked the work. Nevertheless, the treatment effects on participation look similar to those on enrollment: there is no effect of $Info$ or $WifeTicket$ and no interaction effects, but *Discuss* reduced participation relative to $Info$ and $NoInfo$ (Table A14). However, we have less power with the participation outcome than with the enrollment outcome given its lower mean.

[33] The same analyses for the full eight treatment cells are in Figure A8.

[34] We accounted for enrollment that happened after enrollment day as well as dropout. However, only nine women in our sample enrolled after enrollment day, meaning the differences between our enrollment outcome and being in the program at one month primarily reflect dropout from the program. By dropping out, we mean never participating in the program, or participating and later leaving.

6.7 Tests of Heterogeneity Predicted by Ex Ante Model

The null and negative effects on enrollment are largely contrary to what our ex ante model predicted. The model did, however, predict that positive effects should be driven by certain subgroups: do we at least see larger effects in these subgroups?

The answer generally appears to be no. The theory predicts information withholding when the ticketed spouse is less supportive of enrollment, but the effects of $Info$ do not appear to be larger in the subgroup of couples where the ticketed spouse rated female weavers as less appropriate; the difference in the $Info$ effect between this subgroup and the rest of the sample is -1 to 1 percentage point, and this difference is not significant ($p = 0.87$ with controls, columns 1 and 2 of Table 4).[35] The ex ante model also predicts that the positive effects of *Discuss* should come from couples who disagree about whether the wife should enroll. Yet the effects tend to be more *negative* when spouses rate the appropriateness of female weavers differently, and significantly so with controls (columns 3 and 4 of Table 4). We discuss an alternative model that is consistent with this heterogeneity in the next sub-section.

We do find some support for the theory's predicted heterogeneity in effects of $WifeTicket$. Among couples assigned $NoInfo$, the effect of $WifeTicket$ tends to be larger when the wife rates female weavers as more appropriate than the husband, and the difference in effects is significant with controls (columns 5 and 6 of Table 4). However, we do not see strong evidence for this heterogeneity being driven by the model's posited mechanism: information withholding. Columns 7 and 8 estimate the same regression, but consider the indicator for the non-ticketed spouse knowing about the ticket as the outcome. While the coefficients on the interactions here are positive, they are relatively small compared to the outcome mean and not statistically significant ($p = 0.38$). As we discuss below, veto decision-making provides one possible explanation for why the effects of $WifeTicket$ in the $NoInfo$ group tend to be larger when the wife is more supportive.

6.8 Mechanisms

We find suggestive evidence for veto power decision-making as an ex post explanation for our results. We discuss this evidence first, and then discuss alternative explanations, several of which we do not rule out. We note that our mechanisms analyses are our most speculative; future research should be done to more carefully test whether couples make decisions through vetoes, along with other potential explanations.

[35] We do not control for $WifeTicket$ in these regressions to be consistent with the regressions in Table 5 that also make use of variation in whether the ticket was given to the more or less supportive spouse. See footnotes 37 and 38 for further details.

6.8.1 Veto Power Mechanism: Discussion

We describe our veto power model in words here, and formally in Appendix E.2. In this model, we assume that (i) a couple enrolls only if neither spouse vetoes enrollment, (ii) spouses incur a private utility cost from exercising their veto, which can vary by gender and represent internal or external pressure to be agreeable towards one's spouse, and (iii) our treatments affect these costs by determining who felt they had a right to make the decision. In particular, informing only one spouse, as in $NoInfo$, could have made that spouse feel entitled to make the decision. Interventions that nudge households toward joint decision-making, like *Discuss*, and to a lesser extent, $Info$, could make a second spouse feel entitled to exercise a veto. Thus, while all spouses in our setting have *de jure* veto power, since enrollment requires both spouses to attend enrollment day, we posit that our treatments affect *de facto* veto power by varying how personally costly it is to veto what one's spouse wants. This model can rationalize the negative effect of *Discuss* (enrollment is less likely when it requires two spouses not to veto rather than only one), as well as the lack of information withholding (ticket-holders feel comfortable vetoing enrollment and thus do not lose by sharing information). Given gender norms in our setting, the cost of vetoing is likely to be higher for women than men. This, coupled with the fact that wives are more supportive of women weaving than husbands, means the veto power model has no prediction for the effects of $WifeTicket$ or its interactions with $Info$ and *Discuss*.

The veto model is substantively different from the efficient household bargaining model, even more so than our ex ante model. With bargaining, enrollment happens if it maximizes household welfare and even if one spouse is personally opposed, because other decisions the household makes as part of its sustained cooperation will make that spouse better off than their outside option. But in the veto world, enrollment will only happen if neither spouse vetoes it, with vetoing not constrained by ongoing cooperation but personal utility costs incurred from being disagreeable. This sort of decision-making would be consistent with bargaining being extremely costly or contracting very imperfect.

The veto model is also distinct from a model in which joint decision-making interventions increase the influence of the non-ticketed spouse in general. Such a model would tend not to deliver negative effects of joint interventions on enrollment, since half of the time joint intervention in disagreeing couples would increase the influence of the more supportive spouse (increasing enrollment) and half of the time it would increase the influence of the less supportive spouse (decreasing enrollment). Our veto model posits an asymmetry – joint intervention increases the power of the non-ticketed spouse to veto enrollment, but does not increase the power of the non-ticketed spouse to ensure enrollment – and this asymmetry can explain the negative effect of *Discuss*. Finally, we note that our requirement that both spouses attend enrollment day could have made veto decision-making especially likely as it gave individuals a straight-

forward way to exercise a veto: refuse to attend enrollment. This is a limitation of our design, and future research should explore the generality of veto decision-making.[36]

The veto model delivers the following auxiliary predictions:

1. *Asymmetric negative effects*. Intervention toward joint decision-making, i.e. $Info$ or *Discuss*, should reduce enrollment only when the ticketed spouse is more supportive of enrollment than the non-ticketed one. There should be no effects when the ticketed spouse is equally or less supportive, or when the spouses are in agreement.

2. $WifeTicket$ *effect in NoInfo*. Note that among disagreeing couples, it is random whether the ticket was assigned to the more supportive or to the less supportive spouse. According to the veto power story, assigning the ticket to the more supportive spouse should increase enrollment, even in the absence of information withholding – because non-ticketed spouses feel less entitled to veto enrollment than ticketed spouses. This effect should be strongest among couples in $NoInfo$, as $Info$ and *Discuss* reduce the veto cost of non-ticketed spouses.

3. *Effects on decision-making power*. Couples should report the non-ticketed spouse having greater influence over the decision in *Info* and *Discuss*.

In the next sub-section we report evidence on each of these predictions. In addition, we report supportive evidence from the survey of local experts.

6.8.2 Veto Power Mechanism: Evidence

Prediction (1): Asymmetric negative effects. For this prediction, we estimate treatment effects separately for two subsamples: the couples in which the ticketed spouse rated women weaving as more appropriate at baseline than the non-ticketed spouse did, and the remaining couples (in which the non-ticketed spouse was equally or more supportive of female weavers). Our $Info$ and *Discuss* treatments both significantly reduce enrollment when the ticketed spouse is more supportive of female weavers than the non-ticketed one (panel (b), Figure 7), and do not significantly affect enrollment for the other couples (panel (a); see associated regressions in columns 1 and 2 of Table 5). If joint intervention increased the influence of the non-ticketed spouse in general we would expect $Info$ and *Discuss* to *increase* enrollment in this second set of couples. The results are then more consistent with the veto power model.

[36] In principle, the physical nature of the ticket could have had a similar impact: a spouse could veto enrollment by destroying the ticket. We added a question to the endline survey for the final three centers that asked ticketed spouses where the ticket was currently. Only about 2% said it had been thrown away or destroyed. Thus, this does not appear to be an important mechanism through which spouses exercised vetoes.

In columns 5 to 8 of Table 5, we refine this test by splitting the latter set of couples into two groups: those that agree at baseline, and those in which the ticketed spouse was less supportive of female weavers. Of the three subgroups, the effects of *Info* and *Discuss* are the most negative for the couples in which the ticketed spouse is more supportive, as the veto model predicts. In our highest-power test in column 8, we see that being assigned to any joint decision-making intervention (*Info* or *Discuss*) has essentially null effects for the agreeing couples. Relative to effects among the agreeing couples, effects are slightly more negative when the ticketed spouse is less supportive but not significantly so, while effects are significantly more negative when the non-ticketed spouse is less supportive (with $p = 0.12$ for the test of equality of effects for the two types of disagreeing spouses).[37]

One concern with these analyses is that ticketed-more-supportive couples are also more likely to be $WifeTicket$ couples, so the heterogeneity could be driven by giving the wife the ticket rather than spousal disagreement. This does not appear to be the case, for two reasons. First, we saw in Table 3 that the overall effects of $Info$ and *Discuss* do not differ by $WifeTicket$. Second, we find that the interactions between ticketed-more-supportive and the $Info$ and *Discuss* treatments look very similar when we control for $WifeTicket$ and its interactions with the other treatments (Table A15).[38]

Otherwise, it is important to note that while this test of the veto power theory can rationalize the negative effect of *Discuss* relative to $NoInfo$, it does little to explain the effect of *Discuss* relative to $Info$. In particular, if *Discuss* reduced enrollment relative to $Info$ through the same channel, we would expect the interaction effect between *Discuss* and the indicator for the non-ticketed spouse being less supportive to be more negative than the same interaction effect for $Info$. In practice, the two interactions are similar in magnitude, and we cannot reject that they are equal (Table 5).

Prediction (2): $WifeTicket$ effect in *NoInfo*. We find support for this prediction in columns 5 and 6 of Table 5, where we test for equality of the coefficients on Ticketed More Supportive and Ticketed Less Supportive ($p = 0.27$ without controls, $p = 0.07$ with controls). In the *NoInfo* condition, assigning the ticket to the more supportive spouse increases enrollment by 8 to 12 percentage points. And consistent with the *Info* and *Discuss* treatments strengthening the veto power of the non-ticketed spouse, there is no effect of assigning the ticket to the more supportive spouse in the *Info* or *Discuss* arms (p-values from

[37] We do not control for $WifeTicket$ in Table 5, because this complicates the comparison between ticketed-more-supportive and ticketed-less-supportive couples, as detailed in Footnote 38.

[38] Note, however, that controlling for $WifeTicket$ complicates the comparison of ticketed-more-supportive and ticketed-less-supportive couples. In particular, this comparison is not random within husband-ticket or wife-ticket couples; within one of these treatment arms, variation in whether the more or less supportive spouse got the ticket it just driven by the endogenous variation in whether the wife or husband is more supportive.

0.58 to 0.99).[39]

Prediction (3): effects on decision-making power. We do not see support for this prediction – we estimate statistically insignificant effects on endline reports on which spouse most influenced the enrollment decision (Table A16). Related, we might expect joint intervention to increase reported spousal disagreement under the veto model, but we do not see evidence of this (columns 3 and 4, Table A17).

Local experts. After providing their rankings, we asked the local experts whether they could think of any reasons for why the discussion treatment might reduce enrollment. Surveyors matched responses to a set of answer options, but did not read the options aloud, meaning responses were unprompted. We group the answer options into six categories: the veto mechanism, discussion being awkward or argument-provoking, discussion wrongly interfering in the household, discussion preventing strategic behavior of the more interested spouse, the surveyor's presence reminding couples of local norms, and discussion leading to deeper deliberation.[40] About 51% of the experts given the scenario corresponding to our experiment could not think of any reason, consistent with them predicting positive effects from discussion. However, among those who did provide an explanation, explanations most often fell into the veto category; 47% of proposed explanations reflected the veto mechanism, with awkwardness or argument-provoking explanations being second-most common (36%), and wrong interference third (16%) (Figure A9). These results suggest there is support among locals for the veto explanation, though we caveat that the locals mispredicted the overall effectivness of *Discuss* and the nature of the question makes it difficult to categorize responses.

Summary. In sum, we see some suggestive support for the veto power model, but the evidence is mixed. Even if the veto power mechanism is operative, other mechanisms are certain to play some role. We now consider these alternative mechanisms.

[39] These p-values come from tests of Ticketed Less Supportive + $Info \times$ Ticketed Less Supportive = Ticketed More Supportive + $Info \times$ Ticketed More Supportive, and Ticketed Less Supportive + *Discuss* $\times$ Ticketed Less Supportive = Ticketed More Supportive + *Discuss* $\times$ Ticketed More Supportive (columns 5 and 6, Table 5).

[40] The answer options we categorize under the veto mechanism are "a spouse opposed to enrolling would feel more comfortable refusing to enroll", "spouses would become more committed to their views, so someone opposed to enrolling wouldn't change their mind", and "spouses might disagree about enrolling and only enroll if they are both supportive during the discussion." See footnotes for Figure A9 for the answer options considered under each of the other categories.

6.8.3 Alternative Mechanisms for No Effects of Information

The null effects of $Info$ could be rationalized by explanations outside of our two models. Spouses may have internalized norms of honesty, and consequently feel compelled to share relevant information, even if doing so is personally costly. Honesty norms in marriage seem particularly plausible given that even in lab experiments people lie surprisingly little when they have economic motives to do so (Abeler et al. 2019). Supporting this idea, when asked *"Do you think that your spouse keeps secrets from you?"* at baseline, 84% of husbands and 69% of wives answer *"Never"*. Only 2% of husbands and 6% of wives answer *"Often"* or *"Always"*, with the remainder answering *"Sometimes"*.

Another possibility is that respondents did not understand that they had plausible deniability. We find this possibility unlikely. Surveyors explained the nature of the plausible deniability at length, and there were comprehension checks throughout the script to ensure respondents understood the information being conveyed. For example, the surveyor would ask *"Will anyone else know we have given you a ticket?"* 99% of ticketed spouses in *NoInfo* give the correct answer, "no." The surveyor would also ask *"Will your spouse know what we have discussed?"* 93% of ticketed spouses give the correct answer, "no." In the rare cases that respondents answered incorrectly, the surveyor would then explain *"We will not tell any others we have given you a ticket. Note that our team is surveying many men and women in this round of surveying. Often no ticket will be given at all in these surveys. So others may see us speaking now but will not know there is a printed ticket for you to enroll or that you have that ticket."*

While participants are likely to have understood what we told them, a separate concern is that 10% of couples without tickets may have been too low for participants to feel they had plausible deniability. While we do not rule this out, we note that we did not tell the respondents the 10% number, and plausibly respondents may have imagined the number to be larger than 10% given that (i) our script was heavy-handed in emphasizing the importance of plausible deniability, and (ii) the fraction of all potentially-eligible women in a village not getting tickets was substantially higher than 10% due to attrition in between census and baseline, and the likelihood that we missed some households in the census. This may have been the fraction that effectively influenced the beliefs of our respondents.

6.8.4 Alternative Mechanisms for the Negative Effect of Discussion

We group these alternative mechanisms into two categories: those we do versus do not rule out. However, we caveat that, in both categories, our mechanism tests are often underpowered.

Mechanisms we do not rule out.

No time to prepare arguments. Our initial discussion protocol, used for 31% of the sample, had spouses

discuss the job together immediately after learning about the job. The revised protocol ensured a gap between the two – ranging from 0 to 6 days, and 0.79 days on average.[41] Could the lack of time to prepare arguments for enrollment (under the initial protocol) have driven negative effects?[42] We find mixed evidence. The effects of the initial discussion protocol are more negative than those of the second (both relative to *NoInfo* and relative to *Info*), but not statistically significantly so, given that we are quite underpowered (Table A18). The effect of $Info$ is also more negative (though not significantly so) in the one-visit sub-sample despite that protocol not changing, so any differences in *Discuss* effects may be attributable to heterogeneity across villages rather than the change in protocol. We also note that among those who received the revised discussion protocol, the number of days between the two visits does not predict enrollment (Table A19).

Forcing discussion prevented supportive spouses from bringing up the topic at an optimal time. Individuals in favor of enrolling may be better able to persuade their spouses if they bring up the topic at a particular time, e.g. over dinner or when their spouses are in a good mood. The *Discuss* treatment could have reduced employment by preventing this. We consider this mechanism possible, though we do not have an associated empirical test. That said, our exploration of external validity in Section 7 presents some evidence of negative effects of discussion even when there is greater scope for the strategic timing of discussion.

Surveyors' presence increased adherence to the norm that women not work. The presence of surveyors while couples discussed the job may have increased the salience of the norm that women not work outside their homes, reducing enrollment. We have two suggestive pieces of evidence against this story. First, increased norm adherence may also predict negative effects of discussion on perceived job desirability at endline, but we do not see such effects (columns 1 and 2, Table A17). Second, a subset of local experts were asked to consider a discussion treatment in which surveyors would encourage couples to discuss the opportunity after the surveyors had left. Surveyor presence does not affect local experts' ranking of the discussion treatment (column 4, Table A5). However, neither of these analyses is ideal – local experts could very well be wrong about the effects of surveyor presence, while job desirability captures many elements of preferences and was only asked in the latter three villages (where the effect of *Discuss* was a bit less negative, as seen above) – so greater norm adherence as a result of discussion remains a possible channel for negative effects.

Discussion treatment made misrepresenting information about the job harder. Informing couples

[41] As mentioned in footnote 22, we revised the protocol to avoid selective attrition, though we do not ultimately see evidence for the selective attrition we worried about.

[42] We thank an anonymous referee for raising this possibility.

together may make it harder for the ticketed spouse to make the job seem better than it actually is to the non-ticketed spouse. This could explain the difference in enrollment between *Discuss* and $NoInfo$, but could only explain the effect of *Discuss* relative to $Info$ if this effect were driven by couples within $Info$ that were assigned $TicketInfo$ (the sub-treatment in which the non-ticketed spouse was informed about the ticket only) rather than $FullInfo$ (the sub-treatment in which the non-ticketed spouse was given information about the ticket, the job, and enrollment). The information sub-treatment provides evidence on the effects of receiving the "official" information about the job from the research team. We find the sub-treatment made non-ticketed spouses view the job as more desirable at endline, while this story would predict the opposite (Table A3). However, as mentioned above, this question was only asked in the latter three centers. The sub-treatment did reduce enrollment, particularly when women were ticketed, but these differences are not significant (Table A3). In sum, we do not find strong evidence for this channel, but we cannot fully rule it out.

Mechanisms we rule out.

Discussion was awkward. Couples may have found our *Discuss* treatment awkward. They may have associated the job with the research team and decided not to enroll to avoid another uncomfortable experience. However, enrollment was not the only time individuals might have interacted with our team after the discussion – we also asked them to take an endline survey. If they were less likely to enroll because they were avoiding another awkward experience, it is plausible that they would be less likely to take our endline survey. While women in *Discuss* were significantly less likely to take the endline survey than women in $Info$, the corresponding difference for men points in the opposite direction and is not significant, and there are no differences in survey rates for either gender between *Discuss* and $NoInfo$ (columns 1 and 2, Table A11).

Forcing discussion led to argument. The *Discuss* treatment encouraged couples to have a discussion about a high-stakes decision at a particular time and in the presence of two surveyors. Forcing a discussion in this way could have produced arguments, leading couples to default to the status quo of not enrolling. However, we had surveyors record whether they perceived the three-minute discussions to be argumentative or not, and none of these discussions were classified as argumentative. Further, *Discuss* did not affect reports of spousal disagreement about the job opportunity at endline (columns 3 and 4, Table A17), though this question was only asked in the latter three centers.

Surveyors' presence made couple conform to norm of wives being quiet and husbands domineering. To conform to gender norms, wives may be quieter when discussing the job opportunity in public than when discussing in private. The *Discuss* treatment could then have reduced enrollment by preventing the

more supportive views of wives from being heard. Going against this, the effects of discussion were no more negative when wives were *more* supportive of weaving, relative to $NoInfo$ or to $Info$ (Table A20). Similarly, to conform to norms of male authority, husbands may be more firm in their beliefs when being watched. But the effects of discussion were no more negative when husbands were *less* supportive of weaving (Table A20).

Discussion led to deeper deliberation. We saw that the *Discuss* treatment prompted deeper deliberation, in the sense that it increased the number of discussions couples had about the opportunity. If taking the job were actually the wrong decision for many couples, deeper deliberation could have reduced enrollment. Three facts speak against this deliberation channel. First, if deliberation leads to better informed decisions, we would expect that couples in our sample who had more days between ticket receipt and enrollment, and thus more days to decide, would be less likely to enroll. In contrast, we if anything find the opposite (columns 7 to 10, Table A17). Second, as mentioned above, discussion did not affect the perceived job desirability question asked in the latter three centers (Table A17). Third, if decisions made by Discuss couples are better informed, we should see the *Discuss* effect attenuate for the participation outcome, as couples in the other treatment arms would have enrolled, learned it was a bad decision, and dropped out. But we saw similar effects of *Discuss* on participation one-month into program as on enrollment (Table A14).

Discussion group had fewer days to decide. Given the second surveyor visit, couples in the discussion treatment had fewer days between ticket receipt and enrollment. Did having fewer days to decide reduce enrollment? Evidence in Table A17 suggests not. *Discuss* couples did have less time, but these differences are small – .23-.24 days less than $NoInfo$, and .36-.37 days less than $Info$ – and significantly only when compared to $Info$. Likewise, each extra day predicts extra enrollment, but the effect of one extra day is just one percentage point. Finally, our effects on enrollment look similar when we control for the number of days to decide.

Discussion treatment involved two visits. While our *NoInfo* and *Info* treatments were delivered through one surveyor visit, our discussion treatment included a second visit to survey the spouses together. Perhaps the additional visit produced the negative effect.[43] Going against this, and as mentioned above: we find that our negative effects are, if anything, driven more by the one-visit version of the discussion treatment (Table A18), and that among those that received the two-visit discussion treatment, the number of days between the two visits does not predict enrollment (Table A19).

[43] We thank an anonymous referee for suggesting this confound.

Experimenter demand. Experimenter demand effects (de Quidt et al. 2018) on enrollment may have been triggered by discussing the job in the presence of two surveyors. However, the most plausible demand story would predict *higher* effects on enrollment given the reasonable assumption that surveyors wanted women to enroll. Experimenter demand is therefore unlikely to explain why the *Discuss* treatment reduced enrollment.

7 External Validity

How general is our finding that discussion between spouses reduces female employment? Among the plausible mechanisms discussed above, some are more generalizable than others. Effects driven by surveyor presence would not generalize to private discussions. Effects driven by forced discussion timing would likely be more generalizable, to the extent that external attempts to make decision-making more joint would also tend to force the timing. Effects driven by the veto mechanism may be the most generalizable, given that they would apply even in the absence of forced timing.

One insight into generalizability comes from the local experts. The local experts did not predict *Discuss* to be any more effective under two alternative scenarios: with the surveyor not present for the discussion, or if the couple was making a decision about tutoring for their daughter (columns 4 and 5, Table A5). However, local experts mispredicted the ranking of the discussion treatment when given the scenario corresponding to our experiment, so may have mispredicted rankings under these alternative scenarios as well.

To study more naturalistic discussions across India, we examine correlations between discussion and female employment in the most recent (2011/12) wave of the nationally representative Indian Human Development Survey (IHDS). This exercise has the advantage of generalizability, though it uses data that is somewhat dated, and it is much weaker on identification than our experiment, making the resultant correlations at best suggestive of causal effects.

We focus on the answers of married women to the question "Do you and your husband talk about... [things that happen at work or on the farm / about what to spend money on / about things that happen in the community such as elections or politics]?" Answer categories are Never = 0, Sometimes = 1, and Often = 2; with missing data only for 3.4% of married women. The question about work-related conversation is the closest proxy to our discussion treatment, and thus our core independent variable of interest.[44]

[44] Of course, we would ideally have a measure pertaining specifically to *women's* work, since the measure we use will likely often be measuring discussion about the husband's job. Relatedly, we are presuming that the measure of work discussion is a closer proxy for discussion about women's work than the measure of spending, which seems plausible, but may not hold since

We regress the wife's employment status (working or not) on the work-related discussion variable, while controlling for the other two discussion variables. Given that discussion is not randomly assigned, we think of these two discussion controls as potentially addressing omitted-variable bias. For example, they alleviate a form of reverse causality wherein the wife's employment gives the couple less time to have discussions in general. They also help address a concern that more gender-equal couples may both discuss more and be more likely to allow the wife to work.

We first regress the wife's employment on the discussion variables with only minimal, pre-determined, controls: the wife's age, years of education, and state fixed effects. Wives that discuss work-related topics with their husbands are less likely to be employed ($p < 0.01$, column 1, Table A21), with a rough magnitude of two percentage points when comparing those that discuss work-related topics often rather than never.[45] This negative correlation is robust to additional pre-determined controls like caste and the number of children (column 2), and to the addition of increasingly fine geographic fixed effects (columns 3 and 4). Making use of the fact that two women per household are surveyed in a subset of households, we find the correlation is robust to the inclusion of household fixed effects (column 5). Finally, the correlation becomes stronger when controlling for the husband's annual work hours, which helps allay the concern that the husband working generates work-related topics to discuss and also makes the wife's income unnecessary (biasing us towards a negative relationship).

Under the identification assumption that the extent of work-related discussion is random conditional on the amount of the other two types of discussion, the consistent negative coefficient on work-related discussion suggests a causal effect. Our conclusions are similar under an alternative identification assumption: that work-related discussion is random conditional on the average of *all three types* of discussion. This assumption is potentially weaker in that it compares couples that discuss a similar amount (scoring the same average across the three discussion questions), but discuss different topics.[46] We continue to find the negative relationship when controlling for the overall average of the discussion questions (Table A22). However, both identification assumptions are strong. Either would fail if, for example, women bring up the topic of their employment more when their husbands do not let them work, a form of reverse causation from employment to discussion.

households that discuss spending may well be households that are more prone to discussing economic choices in general. These are additional limitations.

[45] While the magnitude of the correlation is smaller than our experimental effects of discussion, it may still be consistent given that the work-related discussion variable is an imperfect proxy for discussion about female employment opportunities (and thus would have an implied first stage coefficient of less than one).

[46] For example, imagine that couples decide how much to discuss based on how much free time they have, and that this free time is driven by an unobservable that also drives female employment. But conditional on time spent discussing, what they discuss is randomly determined by what they see on the news. Under this story, the alternative specification gives us a causal effect but the primary specification does not.

Our takeaway is a correlation: women who discuss work-related topics more are less likely to work. We also note that discussion about spending and politics is generally *positively* correlated with the wife working; wives that discuss more with their husbands tend to be more likely to work, but not when those discussions are about work topics. While only tentative, this analysis suggests that negative effects of discussion on female employment may generalize beyond our particular experimental setting.

8 Conclusion

The key contribution of this paper is to provide evidence that prompting household discussion about an opportunity can lower take-up. This lower take-up was contrary to expert predictions, but was it necessarily bad? The answer is ambiguous from the perspective of the household. In the extreme case of our veto model, *Discuss* removes both spouses' veto costs so that couples only enroll if both spouses support enrollment; this will generate enrollment rates that are inefficiently low relative to a bargaining model in which couples can find a way to compensate the opposed spouse in exchange for enrollment. However, decisions are also not made efficiently in $NoInfo$ or $Info$ in the veto model, and it is not clear whether *Discuss* resulted in enrollment rates further from the efficient benchmark than the comparison arms. The calculus may be different if one extends the model to include, for instance, uncertainty about the costs and benefits of female employment. In this case, the status quo bias introduced by *Discuss* would reduce experimentation and keep couples who underestimate the net gains from female employment from learning. However, from the perspective of the policymaker wishing to raise female employment, *Discuss* did make outcomes worse. More generally, much anti-poverty policy introduces new opportunities to households – immunizations, agricultural technologies, workfare programs, to name a few – with the hope that households will take them up (World Bank Group 2015; Kremer et al. 2019). Our results suggest nudges towards joint decision-making can produce a status quo bias that depresses take-up. More intensive interventions could make joint discussions more productive and shift outcomes towards policymakers' objectives (Ashraf et al., 2020b; Björkman-Nyqvist et al., 2023; Boyer et al., 2022; Donald et al., 2022; Kala and McKelway, 2025b).

Why did researchers mis-predict the effects of our interventions? We think several explanations are likely. First, as discussed in the introduction, there was evidence of information withholding and benefits from encouraging discussion prior to our study, but our interventions differed in meaningful ways from many evaluated previously. This could have led experts to predict positive effects but have low certainty in these predictions; indeed, the average expert predicted just 2.7 of their 7 predictions would be within 5 percentage points of the truth. But even this low confidence was overconfidence, with only 1.1 predictions

actually falling into the 5 percentage point margin on average.[47] Our local expert survey allows us to explore two additional hypotheses for reseachers' mis-predictions: (i) experts are wrong because they fail to understand local cultural norms that local people would understand (Thomas et al. 2020; Thomas and Markus 2023), or (ii) experts and locals are both wrong, suggesting a more general misunderstanding of, and failure of introspection into, household decision-making processes. Our evidence is more consistent with the latter story, as locals also predicted positive effects from discussion, and with confidence – when asked if they felt certain of their predicted rankings, 92% said they strongly agreed, and the remaining 8% said they somewhat agreed. That said, there is also some evidence for the hypothesis on local cultural norms; we saw that locals were a bit less optimistic in their predicted rankings of the discussion treatment than researchers.

Our results suggest several directions for future research. We provide suggestive evidence for a model of veto decision-making which departs even more substantially from the efficient household model than our ex ante model of bargaining frictions; future research should test more directly for this veto decision-making, perhaps by randomizing whether a spouse has de jure veto power. Such work would also be important for establishing the generality of veto decision-making, given that our requirement that both spouses attend enrollment may have especially facilitated veto decision-making. Another direction would be to explore the importance of requiring husbands publicly condone female labor supply by attending enrollment. Women's employment is generally visible to others, but requiring husbands to visibly condone female employment may have reduced take-up. On the other hand, the fact that husbands have a great deal of control over their wives' labor supply means they are likely to be seen as responsible for employment decisions whether or not they publicly condone it.[48] Husbands in patriarchal societies have control over many household decisions; the question of whether husbands' control also increases the social sanctions they bear when their household deviates from norms is of interest in many domains aside from labor supply. Finally, the idea that individuals fail to accurately introspect into their own household decision-making processes would also be interesting to explore. One possibility is that a form of motivated reasoning is at play, where individuals would like to believe joint discussions with their spouse are productive. More generally, our results point to the importance of research on household decision-making processes – the success of intervention into the household hinges on having accurate

[47] Furthermore, confidence is uncorrelated with accuracy, as with the academic experts in DellaVigna and Pope (2018). The more experienced experts in our sample are not more confident, but they tend to give *less* accurate predictions (Table A23). In particular, Full Professors give significantly less accurate predictions than Assistant Professors (column 4), while those with more Google Scholar citations are less accurate than those with fewer (columns 5 and 6). Confidence is still uncorrelated with accuracy after controlling for citations, as is a proxy for effort: whether the expert spent above-median time on the survey (column 6).

[48] In data from a neighboring state in India, Bernhardt et al. (2018) find around 70% of husbands believe that when a woman works, her husband bears more of the social cost than she does.

mental models of how households make decisions.

References

Abeler, Johannes, Daniele Nosenzo, and Collin Raymond, "Preferences for truth-telling," *Econometrica*, 2019, *87* (4), 1115–1153.

Afzal, Uzma, Giovanna d'Adda, Marcel Fafchamps, and Farah Said, "Intrahousehold consumption allocation and demand for agency: a triple experimental investigation," Technical Report, National Bureau of Economic Research 2018.

Anderberg, Dan, Rachel Cassidy, Anaya Dam, Wendy Janssens, Karlijn Morsink, and Anouk van Veldhoven, "Keeping the Peace while Getting Your Way: Information, Persuasion and Intimate Partner Violence," IFS Working Paper W24/22, Institute for Fiscal Studies 2024.

Anderlini, Luca and Leonardo Felli, "Costly bargaining and renegotiation," *Econometrica*, 2001, *69* (2), 377–411.

Anderson, Siwan and Jean-Marie Baland, "The Economics of Roscas and Intrahousehold Resource Allocation," *The Quarterly Journal of Economics*, 2002, *117* (3), 963–995.

Ashraf, Nava, "Spousal control and intra-household decision making: An experimental study in the Philippines," *The American Economic Review*, 2009, *99* (4), 1245–1277.

_ , "Gender: 2024 Nobel Symposium in Economic Sciences," 2024. Stellenbosch Institute for Advanced Study, panel discussion video. https://www.youtube.com/watch?v=CbB0KqIFgRM (accessed February 27, 2025).

_ **, Erica Field, and Jean Lee**, "Household bargaining and excess fertility: an experimental study in Zambia," *The American Economic Review*, 2014, *104* (7), 2210–2237.

_ **, Erica M Field, Alessandra Voena, and Roberta Ziparo**, "Maternal mortality risk and spousal differences in the demand for children," Technical Report, National Bureau of Economic Research 2020.

_ **, Natalie Bau, Corinne Low, and Kathleen McGinn**, "Negotiating a better future: How interpersonal skills facilitate intergenerational investment," *The Quarterly Journal of Economics*, 2020, *135* (2), 1095–1151.

Atkin, David, "Working for the future: Female factory work and child health in Mexico," *Unpublished Manuscript, Yale University*, 2009.

Bernhardt, Arielle, Erica Field, Rohini Pande, Natalia Rigol, Simone Schaner, and Charity Troyer Moore, "Male Social Status and Women's Work," *AEA Papers and Proceedings*, 2018, *108*, 363–67.

Björkman-Nyqvist, Martina, Seema Jayachandran, and Celine P Zipfel, "A Mother's Voice: Impacts of Spousal Communication Training on Child Health Investments," Technical Report, National Bureau of Economic Research 2023.

Boyer, Christopher, Elizabeth Levy Paluck, Jeannie Annan, Tvisha Nevatia, Jasper Cooper, Jackline Namubiru, Lori Heise, and Rachel Lehrer, "Religious leaders can motivate men to cede power and reduce intimate partner violence: Experimental evidence from Uganda," *Proceedings of the National Academy of Sciences of the United States of America*, 2022, *119* (31).

Browning, Martin and Pierre-André Chiappori, "Efficient intra-household allocations: A general characterization and empirical tests," *Econometrica*, 1998, pp. 1241–1278.

Buchmann, Nina, Pascaline Dupas, and Roberta Ziparo, "The Good Wife? Reputation Dynamics and Financial Decision-Making Inside the Household," Technical Report, National Bureau of Economic Research 2023.

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

_ , Alexander W Cappelen, Bertil Tungodden, Alessandra Voena, and David H Yanagizawa-Drott, "How Are Gender Norms Perceived?," Technical Report, National Bureau of Economic Research 2023.

Castilla, Carolina and Thomas Walker, "Is ignorance bliss? The effect of asymmetric information between spouses on intra-household allocations," *American Economic Review*, 2013, *103* (3), 263–268.

Chiappori, Pierre-André, "Collective Labor Supply and Welfare," *Journal of Political Economy*, 1992, *100* (3), 437–67.

Coase, Ronald H, "The problem of social cost," in "Classic papers in natural resource economics," Springer, 1960, pp. 87–137.

_ , "The institutional structure of production," in "Handbook of new institutional economics," Springer, 2005, pp. 31–39.

Conlon, John J, Malavika Mani, Gautam Rao, Matthew W Ridley, and Frank Schilbach, "Learning in the Household," Technical Report, National Bureau of Economic Research 2021.

de Quidt, Jonathan, Johannes Haushofer, and Christopher Roth, "Measuring and Bounding Experimenter Demand," *American Economic Review*, 2018.

Dean, Joshua T and Seema Jayachandran, "Changing family attitudes to promote female employment," *American Economic Association Papers and Proceedings*, 2019, *109*, 138–42.

DellaVigna, Stefano and Devin Pope, "Predicting experimental results: who knows what?," *Journal of Political Economy*, 2018, *126* (6), 2410–2456.

_ , _ , and Eva Vivalt, "Predict science to improve science," *Science*, 2019, *366* (6464), 428–429.

Donald, Aletheia, Markus Goldstein, and Lea Rouanet, "Two Heads Are Better Than One: Agricultural Production and Investment in Cote dâIvoire," 2022. World Bank Policy Research Working Paper 10047.

Duflo, Esther and Christopher Udry, "Intrahousehold resource allocation in Cote d'Ivoire: Social norms, separate accounts and consumption choices," Technical Report, National Bureau of Economic Research 2004.

Fiala, Nathan, "Business is tough, but family is worse: Household bargaining and investment decisions in Uganda," Working Paper 2018.

Field, Erica, Rohini Pande, Natalia Rigol, Simone Schaner, and Charity Troyer Moore, "On Her Own Account: How Strengthening Women's Financial Control Impacts Labor Supply and Gender Norms," *American Economic Review*, 2021, *111* (7), 2342–75.

Fletcher, Erin, Rohini Pande, and Charity Maria Troyer Moore, "Women and work in india: Descriptive evidence and a review of potential policies," 2017.

Gazeaud, Jules, Nausheen Khan, Eric Mvukiyehe, and Olivier Sterck, "With or without him? Experimental evidence on cash grants and gender-sensitive trainings in Tunisia," *Journal of Development Economics*, 2023, *165*, 103169.

Heath, Rachel and A. Mushfiq Mobarak, "Manufacturing growth and the lives of Bangladeshi women," *Journal of Development Economics*, 2015, *115*, 1–15.

_ **and Seema Jayachandran**, "The Causes and Consequences of Increased Female Education and Labor Force Participation in Developing Countries," *Oxford Handbook of Women and the Economy*, 2018.

_ **and Xu Tan**, "Intrahousehold bargaining, female autonomy, and labor supply: Theory and evidence from India," *Journal of the European Economic Association*, 2020, *18* (4), 1928–1968.

Hoff, Karla, Tauhidur Rahman, and Vijayendra Rao, "Epistemic Discrimination against Women: Experimental Evidence from India," *Available at SSRN 3853215*, 2017.

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

Kala, Namrata and Madeline McKelway, "The Impacts of Employee Excellence Rewards on Workers and Firms: The Role of Managerial Discretion vs. Worker Empowerment," AEA RCT Registry. February 07. https://doi.org/10.1257/rct.14861-3.0 2025.

_ **and** _ , "The Power of Persuasion: Causal Effects of Household Communication on Women's Employment," Working Paper 2025.

Kerwin, Jason, Nada Rostom, and Olivier Sterck, "Striking the Right Balance: Why Standard Balance Tests Over-Reject the Null, and How to Fix It," Technical Report, IZA Discussion Papers 2024.

Kremer, Michael, Gautam Rao, and Frank Schilbach, "Behavioral development economics," *Handbook of Behavioral Economics*, 2019, *2*.

McKelway, Madeline, "Women's Employment and Empowerment: Descriptive Evidence," *AEA Papers and Proceedings*, May 2022, *112*, 541â45.

_ , "How Does Women's Employment Affect Their Time Use? Evidence from a Randomized Encouragement Design in India," Working Paper, Hanover, New Hampshire 2025.

_ , "Women's Self-Efficacy and Economic Outcomes: Experimental Evidence from India," Working Paper, Hanover, New Hampshire 2025.

Mel, Suresh De, David McKenzie, and Christopher Woodruff, "Are women more credit constrained? Experimental evidence on gender and microenterprise returns," *American Economic Journal: Applied Economics*, 2009, *1* (3), 1–32.

Nyqvist, Martina Björkman and Seema Jayachandran, "Mothers care more, but fathers decide: Educating parents about child health in Uganda," *American Economic Review*, 2017, *107* (5), 496–500.

Riedl, Arno, "Weakening the weak may harm the strong: A bargaining model where opting-in is costly," Technical Report, Reihe Ökonomie/Economics Series 1995.

Robinson, Jonathan, "Limited insurance within the household: Evidence from a field experiment in Kenya," *American Economic Journal: Applied Economics*, 2012, *4* (4), 140–64.

Schaner, Simone, "Do opposites detract? Intrahousehold preference heterogeneity and inefficient strategic savings," *American Economic Journal: Applied Economics*, 2015, *7* (2), 135–74.

_ , "The cost of convenience? Transaction costs, bargaining power, and savings account use in Kenya," *Journal of Human Resources*, 2017, *52* (4), 919–945.

Thomas, Catherine C, Nicholas G Otis, Justin R Abraham, Hazel Rose Markus, and Gregory M Walton, "Toward a science of delivering aid with dignity: Experimental evidence and local forecasts from Kenya," *Proceedings of the National Academy of Sciences*, 2020, *117* (27), 15546–15553.

Thomas, Catherine Cole and Hazel Rose Markus, "Enculturating the science of international development: beyond the WEIRD independent paradigm," *Journal of Cross-Cultural Psychology*, 2023, *54* (2), 195–214.

Udry, Christopher, "Gender, agricultural production, and the theory of the household," *Journal of political Economy*, 1996, *104* (5), 1010–1046.

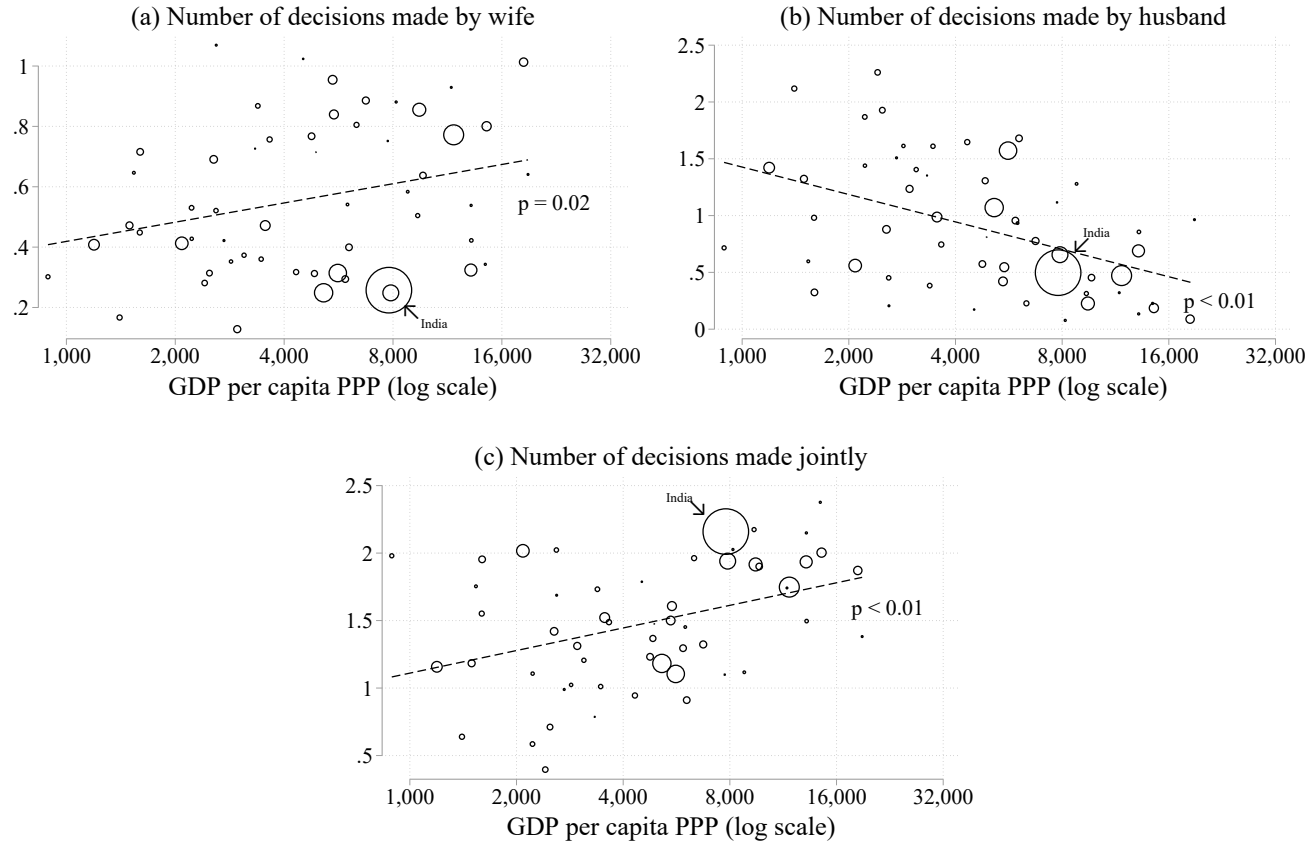
World Bank Group, *World Development Report 2015: Mind, Society, and Behavior*, Washington, DC: World Bank, 2015.

Zhang, Sally, "Asymmetric Information, Intrahousehold Cooperation, and Consumption Choices: Experimental Evidence from Kenya," Working Paper 2024.

_ , "Hidden in Plain Sight: Asymmetric Information and Hidden Income within the Household," Working Paper 2024.

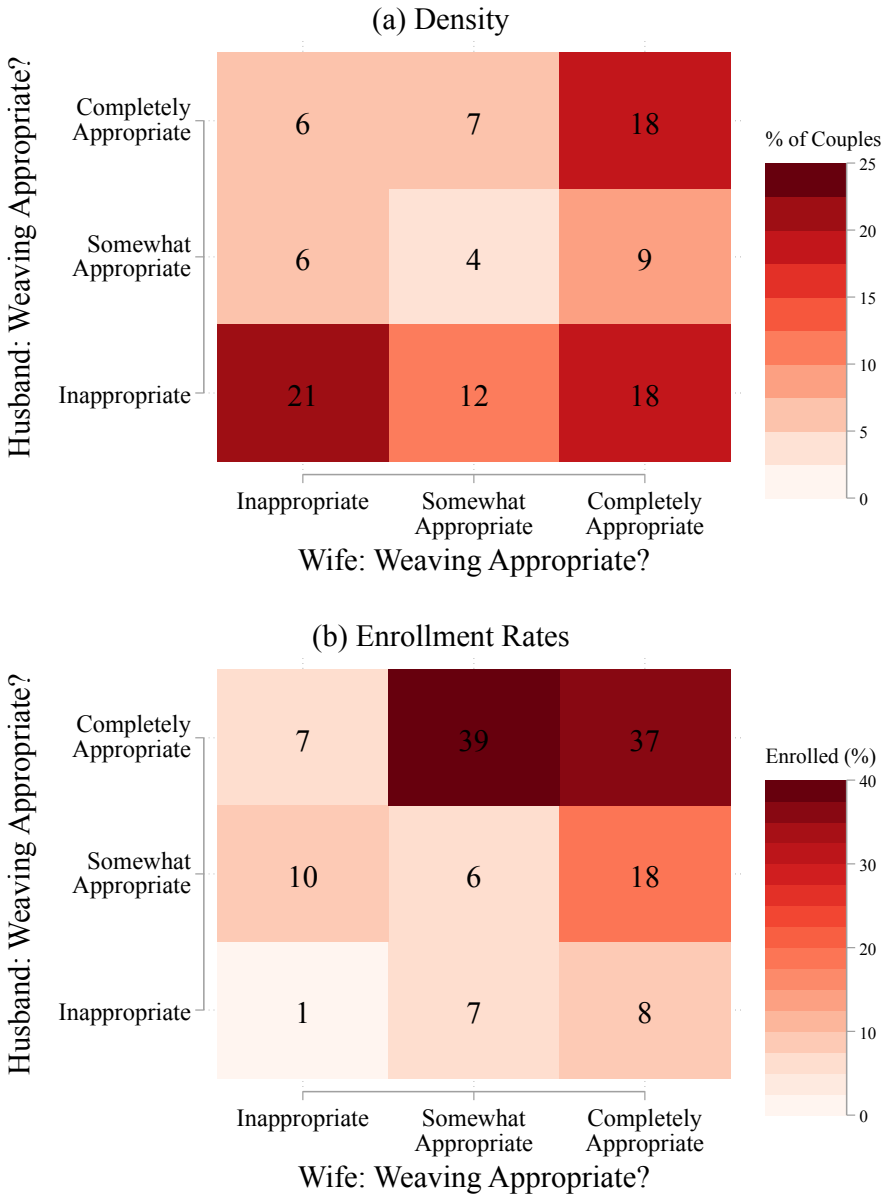
Figures

Figure 1: Couples Make More Joint Decisions in Developed Countries



Notes: Each panel correlates GDP per capita PPP with measures of household decision-making from each country's latest available Demographic and Health Survey. The decision-making question derives from the answers of married women to three questions: *Who usually makes decisions about [healthcare for yourself / making major household purchases / visits to your family or relatives]?* Respondents select one option from (i) respondent, (ii) husband, (iii) respondent and husband jointly, (iv) someone else, or (v) other. The outcome for panel (a) is the sum of questions (0 to 3) to which the respondent (the wife) answers that they make the decision; panel (b) is the equivalent for answering husband; panel (c) is the equivalent for answering respondent and husband jointly. The bubble size reflects country-level population. For further details on the data, see Appendix B. The dashed line is the linear fit from a regression of the decision-making outcome on ln(GDP per capita PPP). The p-value tests the null hypothesis that the slope is equal to zero.

Figure 2: Misaligned Preferences and Enrollment



Notes: N = 489 couples. Panel (a) shows a heatmap of the percentage of couples with each possible combination of answers given to the baseline question: *"How appropriate would it be for women in your household to hold a full-time job outside the home as a weaver?"* Panel (b) shows job enrollment rates in the experiment, pooling across all treatments, separately for each baseline preference combination.

Figure 3: Experiment Timeline

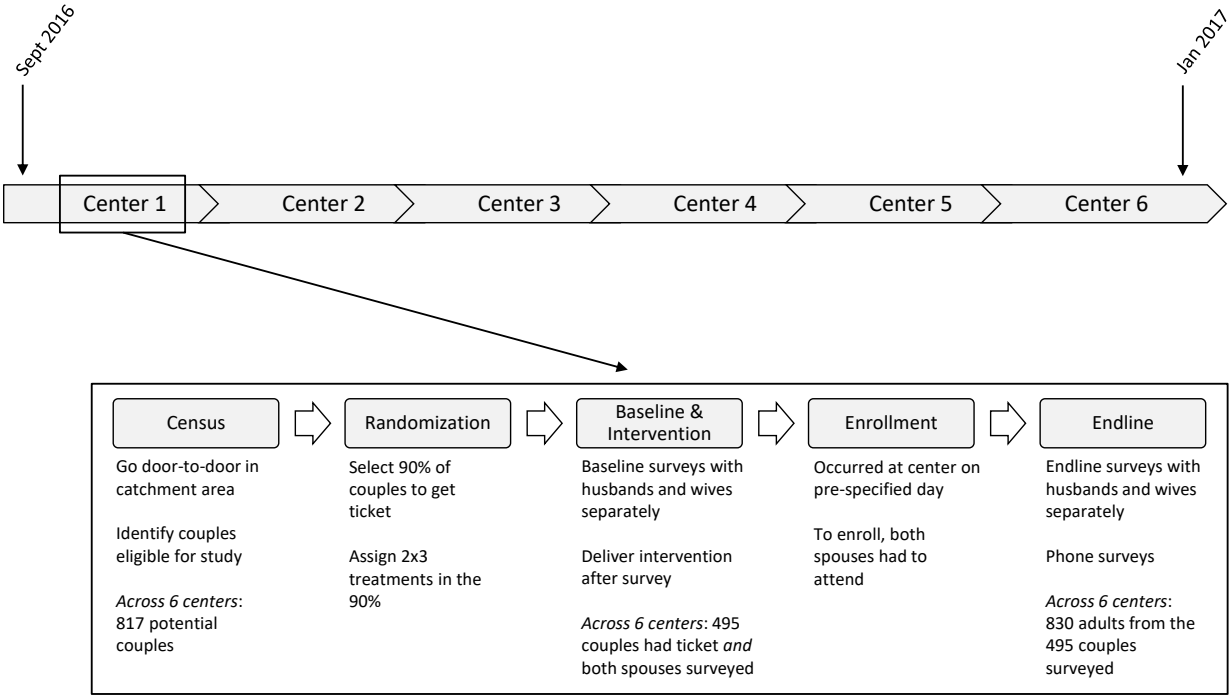
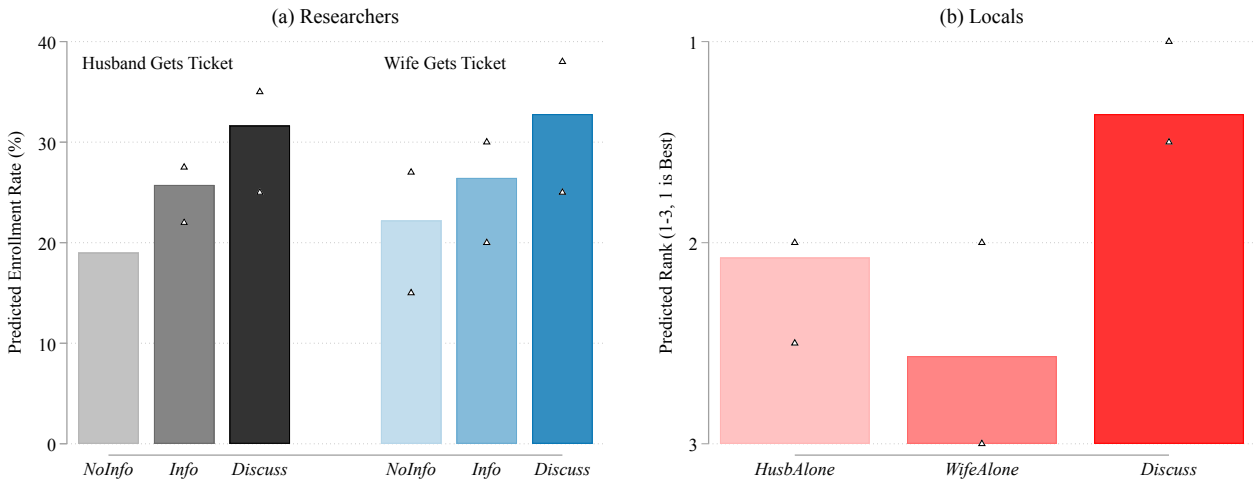


Figure 4: Experiment Design

Non-Ticketed:	*Ticketed Spouse:* **Husband**	**Wife**	
NoInfo	*77 couples*	*82 couples*	159 couples
Info	*83 couples*	*88 couples*	171 couples
Discuss	*88 couples*	*77 couples*	165 couples
	248 couples	247 couples	

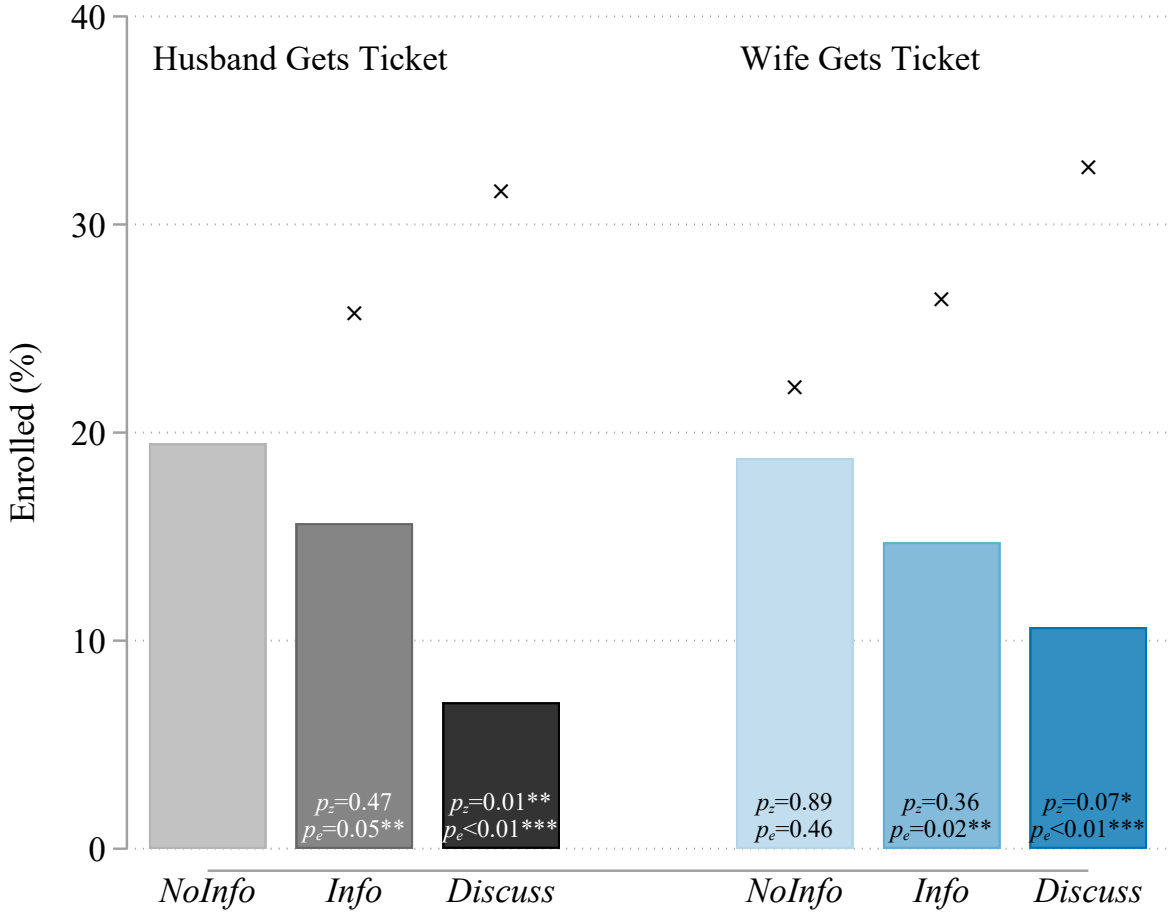
Notes: The figure visualizes our 3-by-2 experimental design, with randomization at the couple-level. The columns indicate which spouse was given the ticket required to enroll in the job. The rows indicate the information treatment for the non-ticketed spouse. In *NoInfo* the non-ticketed spouse is not aware that their spouse has a job ticket, and is aware that not all couples get job tickets, giving their spouse plausible deniability. In *Info* the non-ticketed spouse knows of the existence of the job ticket. In *Discuss* the non-ticketed spouse is present when the spouse receives the job ticket, and the couple are encouraged to discuss the opportunity together for three minutes. The numbers indicate the number of couples in our analysis sample of 495 couples who were assigned a given treatment.

Figure 5: Expert Predictions Line Up With Predictions of Bargaining Frictions Model



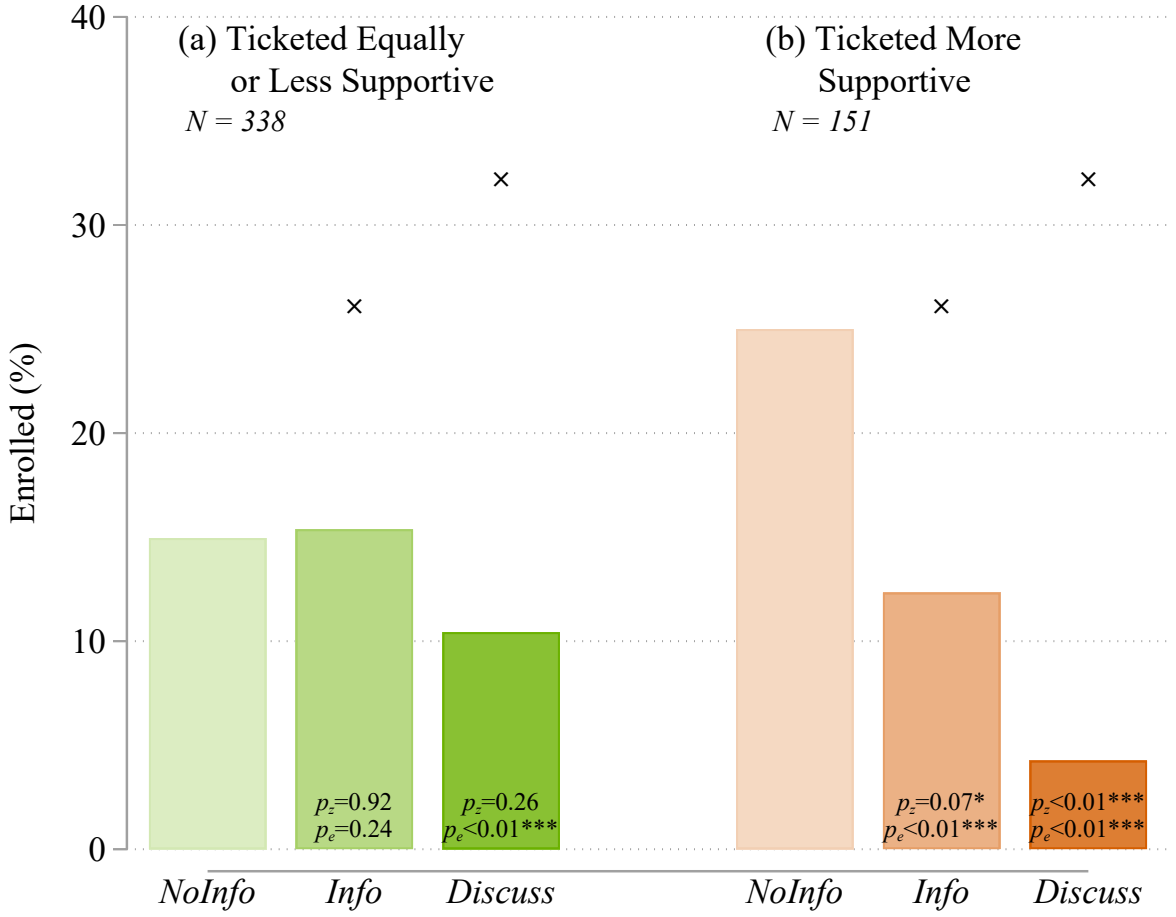
Notes: Panel (a) shows the mean predicted enrollment rate for each treatment according to our full sample of 70 experts on intra-household economics. Each expert was told the true enrollment rate (19%) of the treatment in which the husband gets the job ticket and the wife doesn't know, while the other five bars reflect predictions. Panel (b) presents responses from the 241 locals who ranked the following information campaigns in terms of their effectiveness of getting wives to enroll in Obeetee's weaving program: informing the husband alone, informing the wife alone, or informing the two together and encouraging discussion during the meeting. Each bar shows the average rank of an intervention, where lower numbered ranks mean greater predicted effectiveness. In both, triangles denote the 25th and 75th percentile response for each intervention.

Figure 6: Impacts on Enrollment Are Far Less Positive Than Experts Predict



Notes: The figure visualizes enrollment rates across the six treatments, derived from a regression ($N = 495$) of enrollment on treatment indicators, modified strata effects (i.e. village fixed effects and a dummy variable for OBC caste status), and the following controls: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for wife has worked for income in the last three months, (10) indicator for husband has worked for income in the last three months, (11) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (12) husband's answer to how appropriate for woman in household to work as weaver (zero to two), with each of these variables set to the mean when missing. Crosses denote mean predictions from the full sample of 70 experts. p_z is the p-value from a test of whether the treatment effect relative to the control group category (the husband gets the ticket, *NoInfo*) is equal to zero. p_e is the p-value from a test of whether the treatment effect is equal to the mean expert prediction.

Figure 7: Joint Decision-Making Interventions Reduce Enrollment Only When Ticketed is More Supportive



Notes: The figure visualizes enrollment rates across the three information treatments, separately for two subsamples: couples in which the ticketed spouse rated women weaving as more appropriate at baseline than the non-ticketed spouse did (panel (b)); and the rest of the couples, in which the ticketed spouse rated women weaving as equally or less appropriate (panel (a)). The enrollment rates are then derived from two regressions, each regressing enrollment on indicator variables for $Info$, *Discuss*, $WifeTicket$, modified strata effects (i.e. village fixed effects and a dummy variable for OBC caste status), and the following controls: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for wife has worked for income in the last three months, (10) indicator for husband has worked for income in the last three months, (11) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (12) husband's answer to how appropriate for woman in household to work as weaver (zero to two), with each of these variables set to the mean when missing. Crosses denote mean predictions from the full sample of 70 experts. These predictions do not differ by subsample as experts were only asked to predict for the entire sample. p_z is the p-value from a test of whether the treatment effect relative to $NoInfo$ is equal to zero. p_e is the p-value from a test of whether the treatment effect is equal to the mean expert prediction.

Tables

Table 1: Wives Are More Supportive of Women Working Outside the Home Than Husbands

	Appropriate for Man?			Appropriate for Woman?			
	Construction	Weaver	Teacher	Construction	Weaver	Teacher	Job Interest
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Wife	-0.04	0.05	-0.02	0.29***	0.31***	0.11***	0.14**
	(0.05)	(0.05)	(0.02)	(0.05)	(0.05)	(0.04)	(0.06)
Observations	976	974	975	973	973	971	500
Husband Mean	1.4	1.4	1.9	.57	.81	1.7	1.2

Notes: Robust standard errors in parentheses. The outcomes for the first three columns are the perceived appropriateness for men in the household to work full-time outside the house in construction, teaching, and weaving. The outcomes for the second three columns are the same three perceived appropriateness measures, but for women. These appropriateness outcomes take values: 0 = Inappropriate, 1 = Somewhat Appropriate, 2 = Completely Appropriate. During the baseline survey, all wives informed about the job opportunity were asked *How interested are you in this training opportunity?*, while all informed husbands were asked *How interested are you in this training opportunity for [wife's name]?* Job Interest in column (7) is the answer to this question, taking values: 0 = Not At All/Not Very Interested, 1 = Somewhat Interested, 2 = Very Interested. The final column restricts to only those couples who were both asked about their interest during baseline. *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Husbands' Preferences Are More Predictive of Enrollment Than Wives'

	Enrolled	
	(1)	(2)
Husband: Weaving Appropriate for Women in HH?	0.12***	0.11***
	(0.02)	(0.02)
Wife: Weaving Appropriate for Women in HH?	0.06***	0.05***
	(0.02)	(0.02)
Observations	489	489
p(Husband = Wife)	0.03**	0.03**
Outcome Mean	0.15	0.15
Controls	No	Yes

Notes: Robust standard errors in parentheses. Enrolled is an indicator for enrolling on enrollment day. Appropriateness outcomes take values: 0 = Inappropriate, 1 = Somewhat Appropriate, 2 = Completely Appropriate. Controls are: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for wife has worked for income in the last three months, (10) indicator for husband has worked for income in the last three months, with each of these variables set to the mean when missing. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Effects on Enrollment

	Enrolled			
	(1)	(2)	(3)	(4)
Info	-0.03	-0.04	-0.04	-0.04
	(0.04)	(0.04)	(0.06)	(0.05)
Discuss	-0.08**	-0.10***	-0.10*	-0.12**
	(0.04)	(0.03)	(0.05)	(0.05)
WifeTicket	-0.00	0.01	-0.02	-0.01
	(0.03)	(0.03)	(0.06)	(0.05)
Info × *WifeTicket*			0.01	-0.00
			(0.08)	(0.07)
Discuss × *WifeTicket*			0.05	0.04
			(0.07)	(0.07)
Observations	495	495	495	495
NoInfo Mean	0.18	0.18		
HusbTicket Mean	0.15	0.15		
NoInfo & HusbTicket Mean			0.19	0.19
p(*Info* = *Discuss*)	0.22	0.06*		
p(*Info* = Expert Mean)	0.03**	0.01**		
p(*Discuss* = Expert Mean)	<0.01***	<0.01***		
p(*Discuss* – *Info* = Expert Mean)	<0.01***	<0.01***		
p(*WifeTicket* = Expert Mean)	0.51	0.71		
p(*Info* × *WifeTicket* = *Discuss* × *WifeTicket*)			0.60	0.50
p(*Info* × *WifeTicket* = Expert Mean)			0.65	0.76
p(*Discuss* × *WifeTicket* = Expert Mean)			0.36	0.36
p(*Discuss* × *WifeTicket* – *Info* × *WifeTicket* = Expert Mean)			0.65	0.55
Modified Strata	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes

Notes: Robust standard errors in parentheses. Enrolled is an indicator for enrolling on enrollment day. Modified Strata controls are village fixed effects and an indicator for OBC status. Controls are: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for wife has worked for income in the last three months, (10) indicator for husband has worked for income in the last three months, (11) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (12) husband's answer to how appropriate for woman in household to work as weaver (zero to two), with each set to the mean when missing. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Limited Support for Heterogeneity Predicted by Ex Ante Model

	Enrolled *X = Ticketed Less Supportive*		Enrolled *X = Spouses Disagree*		Enrolled *X = Wife More Supportive*		Non-Ticketed Knows *X = Wife More Supportive*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Info	-0.03	-0.04	0.04	0.03				
	(0.05)	(0.05)	(0.06)	(0.06)				
X	-0.05	-0.06	0.03	0.05	-0.17**	-0.20***	0.12	0.14
	(0.06)	(0.05)	(0.06)	(0.06)	(0.08)	(0.07)	(0.11)	(0.13)
Info × X	-0.01	0.01	-0.12	-0.12				
	(0.08)	(0.08)	(0.08)	(0.08)				
Discuss			-0.03	-0.03				
			(0.06)	(0.05)				
Discuss × X			-0.09	-0.13*				
			(0.07)	(0.07)				
WifeTicket					-0.07	-0.09	-0.11	-0.10
					(0.08)	(0.08)	(0.11)	(0.12)
WifeTicket × X					0.12	0.20*	0.13	0.14
					(0.11)	(0.10)	(0.15)	(0.16)
Observations	489	489	489	489	489	489	412	412
Outcome Mean	0.15	0.15	0.15	0.15	0.15	0.15	0.74	0.74
Info, *Discuss*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
WifeTicket	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Info × X, *Discuss* × X	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Info × *WifeTicket*, *Discuss* × *WifeTicket*	No	No	No	No	Yes	Yes	Yes	Yes
Info × *WifeTicket* × X,								
Discuss × *WifeTicket* × X	No	No	No	No	Yes	Yes	Yes	Yes
Modified Strata	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors in parentheses. Enrolled is an indicator for enrolling on enrollment day. Non-Ticketed Knows is an indicator for the non-ticketed spouse knowing at endline that their spouse received an enrollment ticket. Ticketed Less Supportive, Spouses Disagree, and Wife More Supportive are indicators based on husbands' and wives' baseline ratings of the appropriateness of women weaving. Ticketed Less Supportive is an indicator for the ticketed spouse giving a lower rating than the non-ticketed one, Spouses Disagree is an indicator for the two spouses giving different ratings, and Wife More Supportive is an indicator for the wife giving a higher rating than the husband. Note that we only report the coefficients relevant for the associated empirical prediction; some unreported coefficients (like *Discuss* in column (1)) are still included in the regressions, as indicated in the rows towards the bottom of the table. Modified Strata controls are village fixed effects and an indicator for OBC status. Controls are: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for wife has worked for income in the last three months, (10) indicator for husband has worked for income in the last three months, (11) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (12) husband's answer to how appropriate for woman in household to work as weaver (zero to two), with each set to the mean when missing. *** p<0.01, ** p<0.05, * p<0.1.

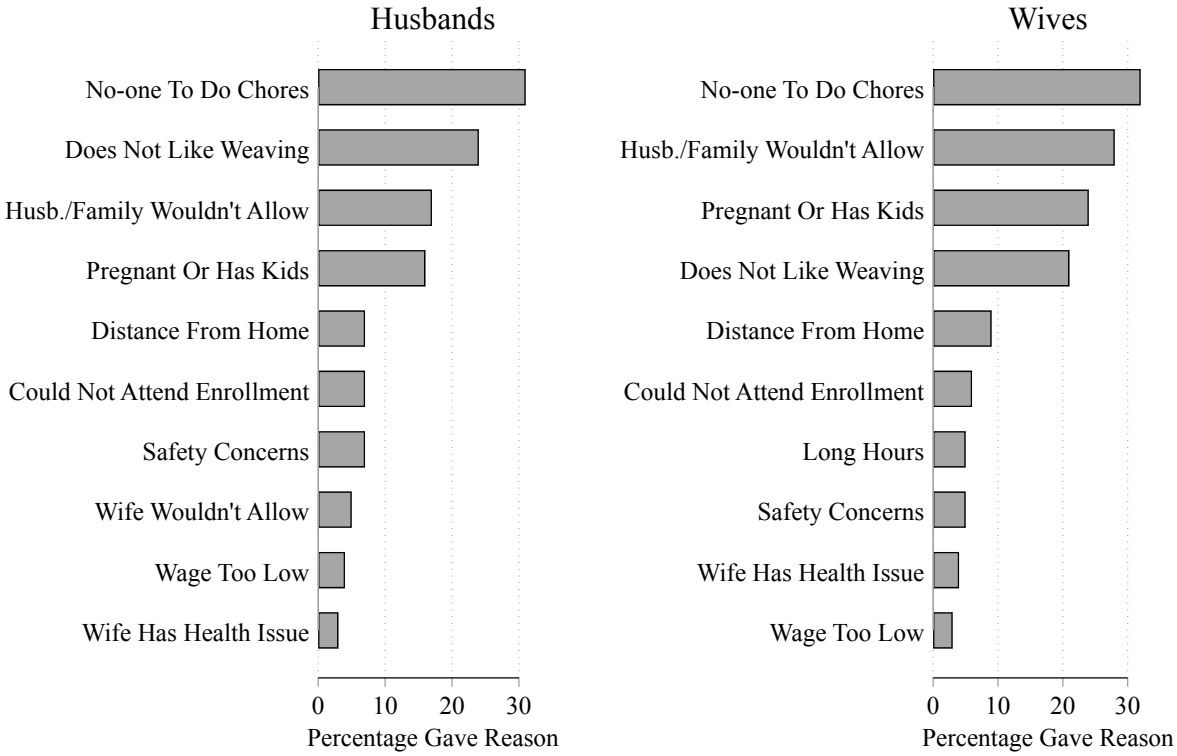
Table 5: Effects Are More Negative When Ticketed is More Supportive

	Enrolled							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Info	0.01	0.01			0.04	0.03		
	(0.05)	(0.04)			(0.06)	(0.06)		
Discuss	-0.04	-0.05			-0.03	-0.03		
	(0.04)	(0.04)			(0.06)	(0.05)		
Ticketed More Supportive	0.07	0.12*	0.07	0.12*	0.07	0.11	0.07	0.11
	(0.07)	(0.06)	(0.07)	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)
Info × Ticketed More Supportive	-0.13	-0.15*			-0.16*	-0.17*		
	(0.09)	(0.08)			(0.10)	(0.09)		
Discuss × Ticketed More Supportive	-0.12	-0.18**			-0.14	-0.20**		
	(0.08)	(0.08)			(0.09)	(0.09)		
Info or *Discuss*			-0.01	-0.02			0.01	0.00
			(0.04)	(0.04)			(0.05)	(0.05)
(*Info* or *Discuss*) × Ticketed More Supportive			-0.13*	-0.16**			-0.15*	-0.18**
			(0.08)	(0.07)			(0.08)	(0.08)
Ticketed Less Supportive					-0.01	-0.01	-0.01	-0.01
					(0.06)	(0.06)	(0.06)	(0.06)
Info × Ticketed Less Supportive					-0.07	-0.06		
					(0.09)	(0.09)		
Discuss × Ticketed Less Supportive					-0.04	-0.05		
					(0.08)	(0.08)		
(*Info* or *Discuss*) × Ticketed Less Supportive							-0.06	-0.05
							(0.08)	(0.07)
Observations	489	489	489	489	489	489	489	489
p(More Supportive = Less Supportive)					0.27	0.07*	0.27	0.08*
Modified Strata	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors in parentheses. Enrolled is an indicator for enrolling on enrollment day. Ticketed More (Less) Supportive is an indicator for the ticketed spouse rating women weaving as more (less) appropriate at baseline than the non-ticketed spouse did. Modified Strata controls include village fixed effects and an indicator for OBC status. Controls are: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for wife has worked for income in the last three months, (10) indicator for husband has worked for income in the last three months, (11) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (12) husband's answer to how appropriate for woman in household to work as weaver (zero to two), with each control variable set to the mean when missing. *** p<0.01, ** p<0.05, * p<0.1.

A Appendix [For Online Publication]

Figure A1: The Most Common Reasons for Non-Enrollment

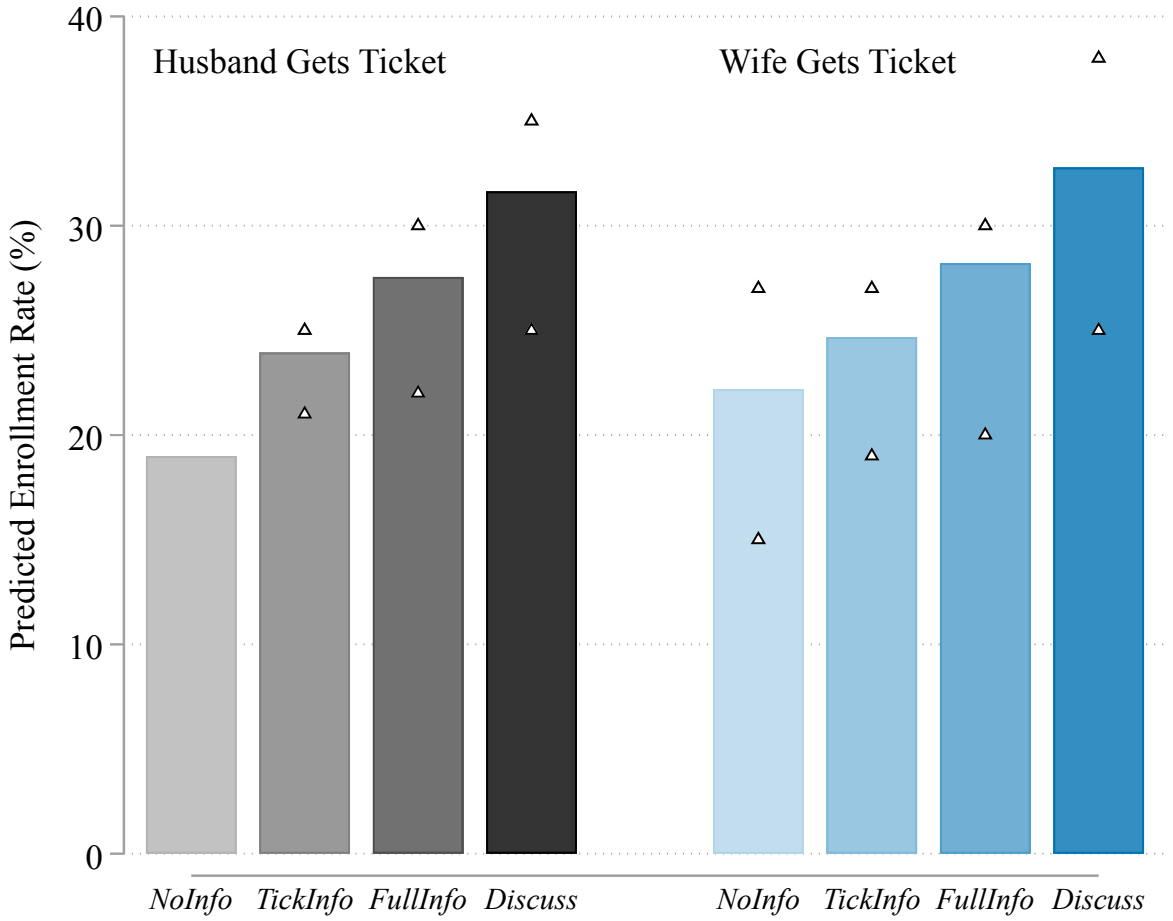


Notes: The figure shows the ten most common reasons for non-enrollment given by 334 husbands and 339 wives during the endline survey. Specifically, during the endline survey we asked wives and husbands of non-enrolling households: *What are the main reasons your household decided not to enroll in this opportunity?* Respondents could select multiple answers from among the following options: (1) fear security/safety issue, (2) no-one to take over housework, childcare, and agricultural work at home, (3) wife is pregnant or has young child, (4) hours too long, (5) wife already has employment, (6) worried what others will think, (7) husband/in-laws/family wouldn't allow, (8) wife wouldn't allow, (9) don't need the money, (10) wage too low, (11) don't like this particular job (weaving), (12) worry that a job after training wouldn't be guaranteed, (13) center too far from home, (14) woman has health issue, (15) one or both could not attend, or forgot, enrollment, (16) lost or did not have ticket, and (17) don't know. Among non-enrolling households, we did not ask this question to non-ticketed spouses in the *NoInfo* treatment group when these spouses were not aware of any job ticket.

Figure A2: An Enrollment Ticket

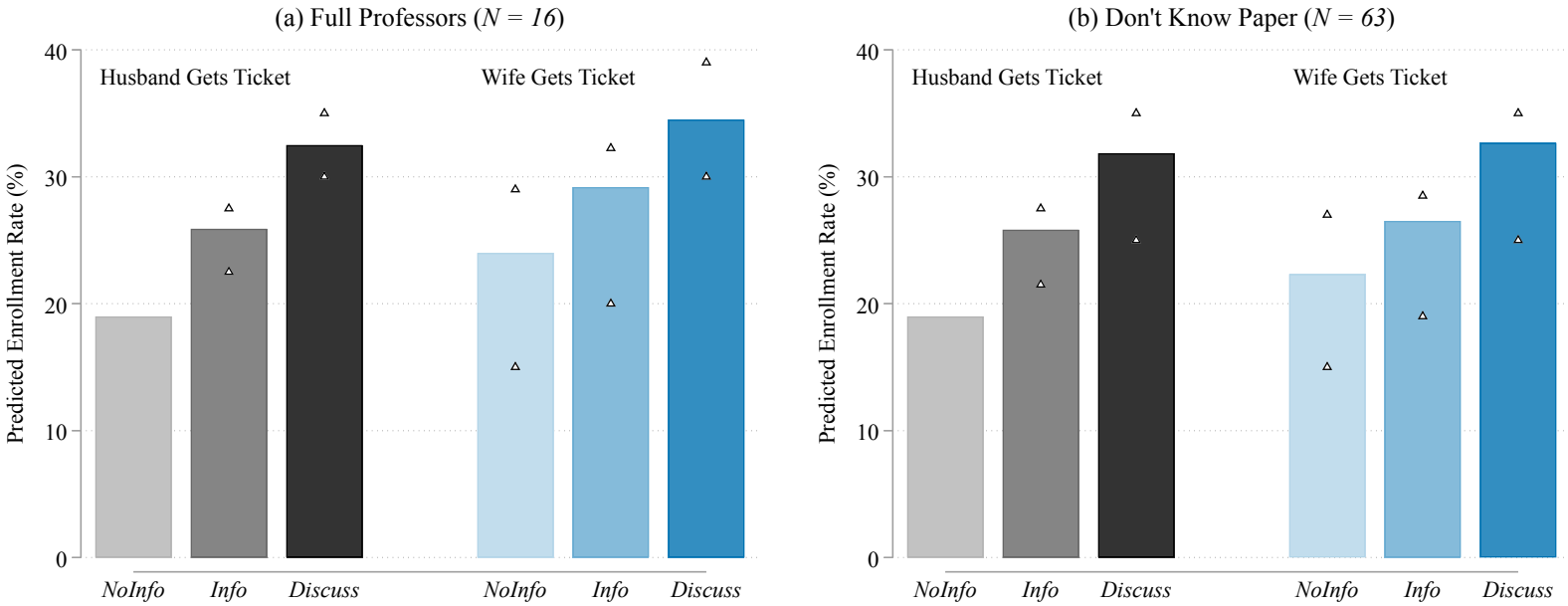
Notes: This is a photo of an enrollment ticket. The first line says "enrollment ticket." The second provides the names of the husband and wife to which this ticket corresponds. The names are blurred for confidentiality. The third line has the unique identification number we assigned to this couple. The fourth and fifth lines provide the date and time of enrollment. The last line has the location of enrollment, which was the female weaving center in the couple's village that was owned by a woman in the village and adjacent to her house. This last line has the names of the loom owner and her husband, and is blurred for confidentiality.

Figure A3: Expert Predictions Including *Info* Sub-Treatments



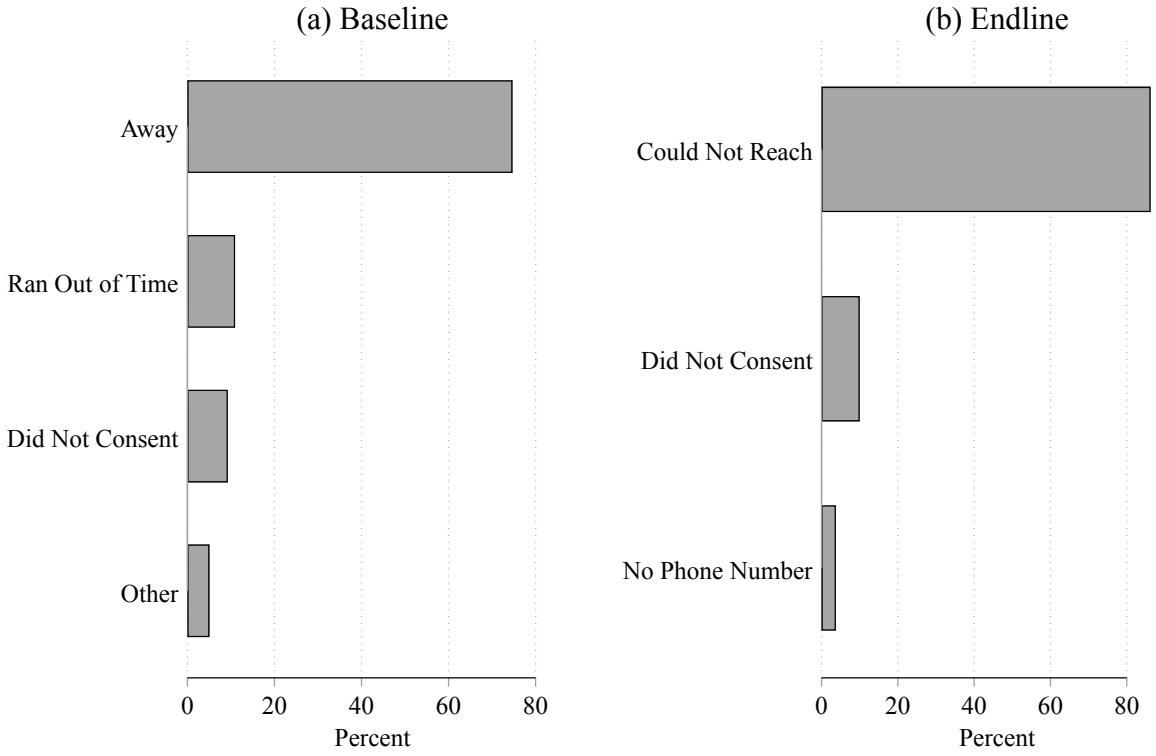
Notes: The figure shows the mean predicted enrollment rate for the full eight treatment cells according to our full sample of 70 experts on intra-household economics. Each expert was told the true enrollment rate (19%) of the treatment in which the husband gets the job ticket and the wife doesn't know (*NoInfo*), while the other seven bars reflect predictions. Triangles denote the 25th and 75th percentile prediction for each treatment. The *TickInfo* and *FullInfo* treatment cells each contain half of the sample size of the other cells, and we pool them into a combined *Info* treatment in our analysis.

Figure A4: Robustness of Expert Predictions



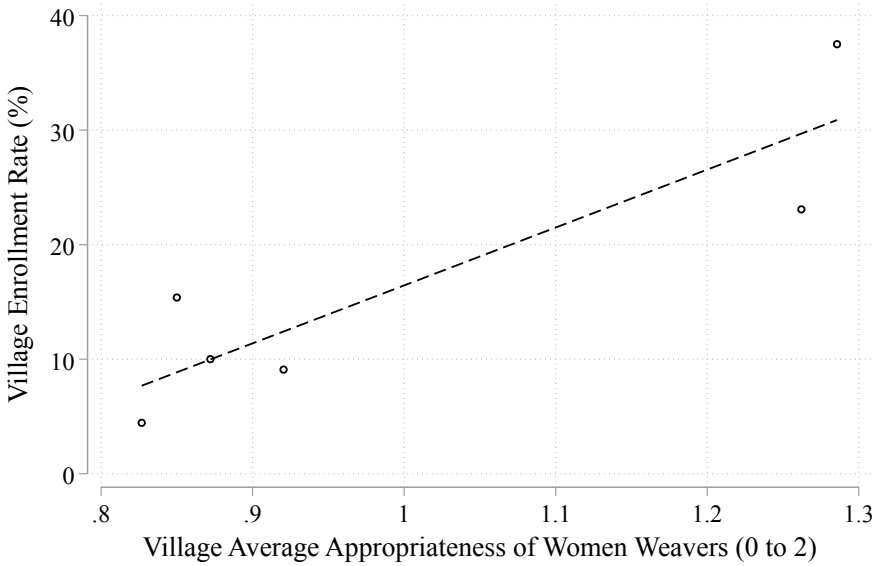
Notes: The figure shows the mean predicted enrollment rate for each treatment cell according to experts on intra-household economics. Panel (a) shows the predictions of the 16 experts who are Full Professors. Panel (b) shows the predictions of the 63 experts that answered *No* to the question *Have you seen or heard results from an experiment on household decision-making about female labor supply in India conducted by Matt Lowe (UBC) and Madeline McKelway?* after reading the details of the experiment. This sample excludes the seven experts that had seen the results of the experiment but couldn't remember them. Each expert was told the true enrollment rate (19%) of the Husband Gets Ticket-NoInfo treatment cell, while the other five bars reflect unincentivized predictions. Triangles denote the 25th and 75th percentile prediction for each treatment.

Figure A5: Reasons for Attrition



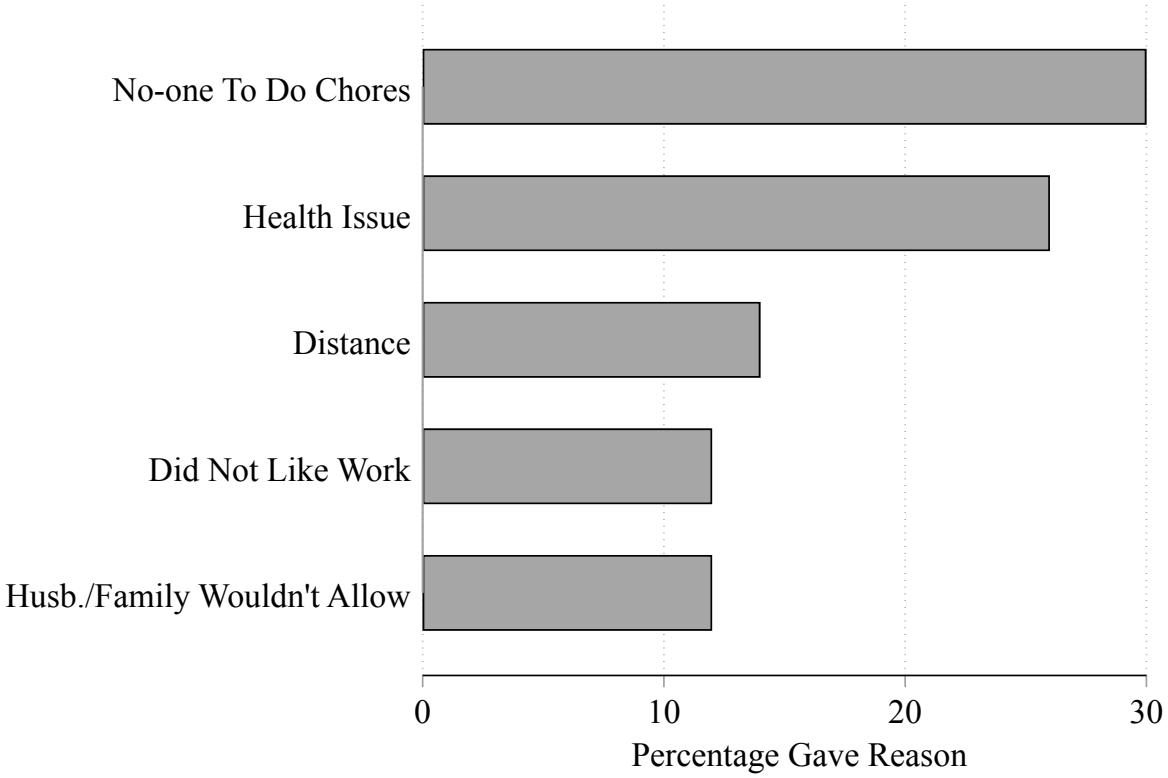
Notes: Panel (a) shows the reasons for non-completion of the Baseline survey ($N = 237$ couples). *Away* means that the husband, wife, or both, were out of town, away all day, or otherwise not available for the Baseline survey. *Ran Out of Time* means that the canvassing period prior to enrollment day ended before the couple could be surveyed. Panel (b) shows the reasons for non-completion of the Endline survey ($N = 160$ individuals, from the $N = 990$ that completed the Baseline survey).

Figure A6: Enrollment Is Higher in Villages Where Women Weavers Are Considered More Appropriate



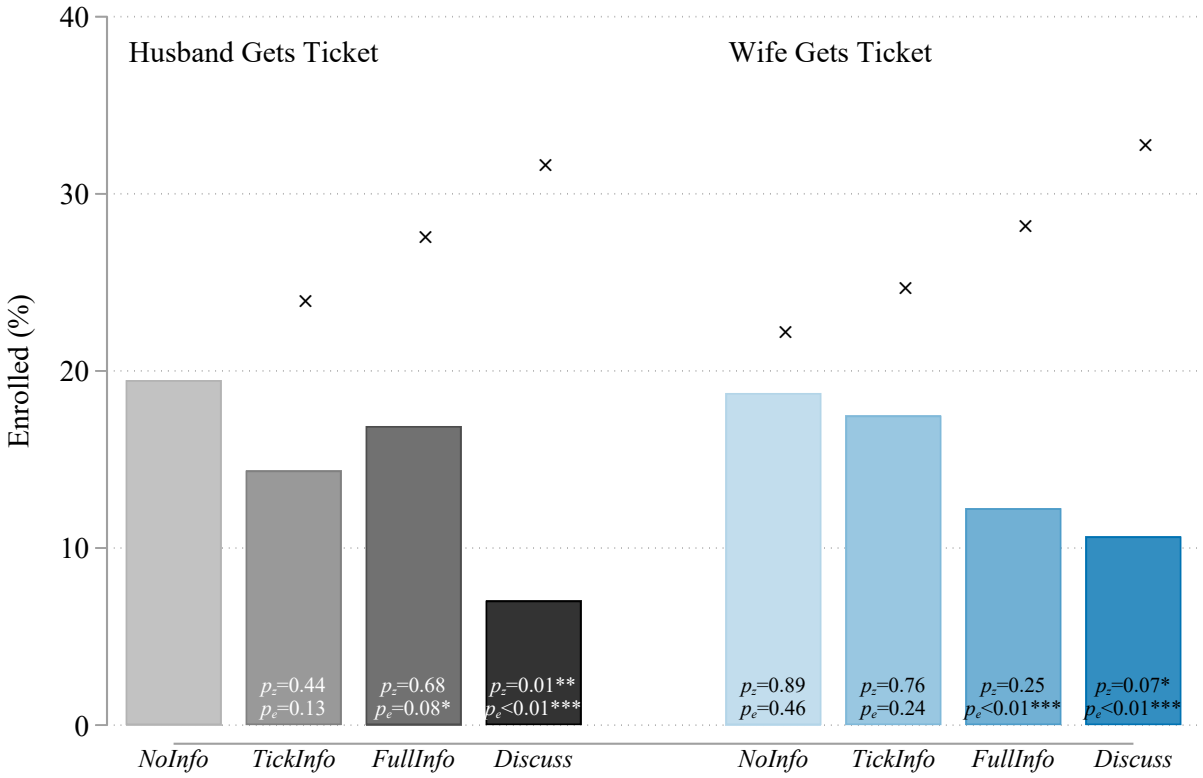
Notes: The figure plots village-level enrollment against the village-level average appropriateness of women weavers reported by couples at baseline. The dashed line is a linear fit.

Figure A7: The Most Common Reasons for Dropout



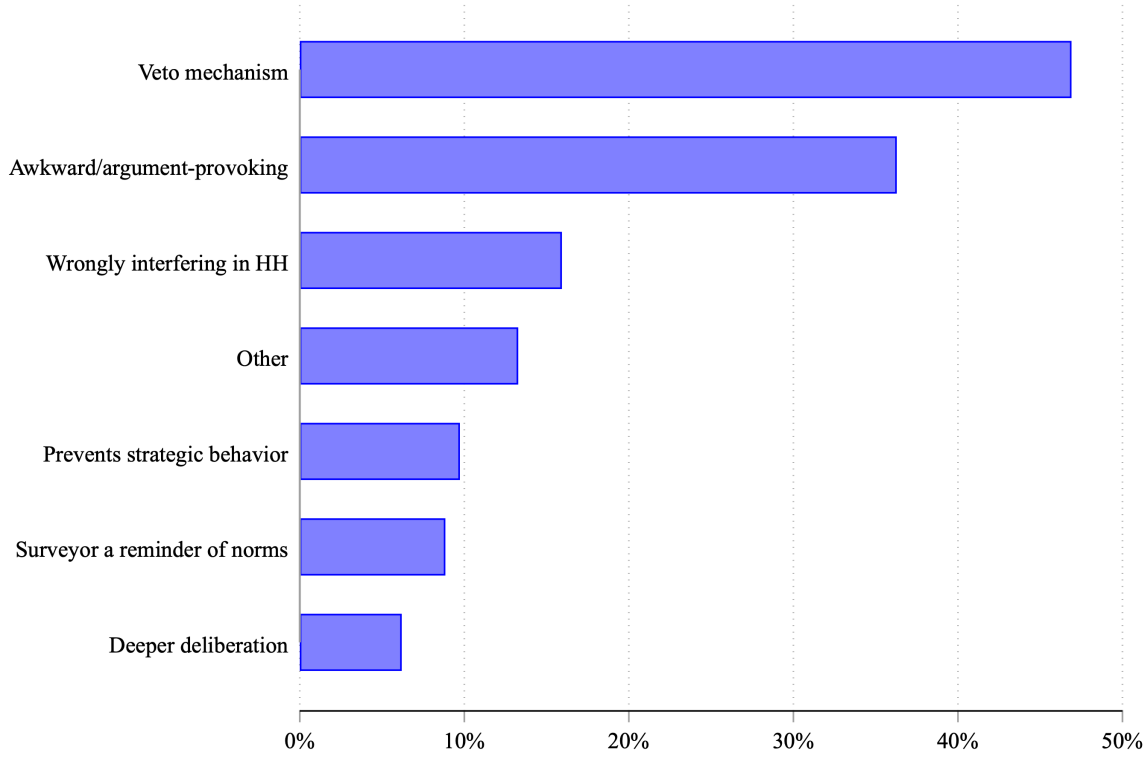
Notes: The figure shows the five most common reasons for dropping out of the firm's program, among the 43 women in our sample who dropped out of the firm's program within the first month. Data were collected by the research team through follow-up discussions with loom centers and weavers once the program had begun. Multiple reasons could be recorded for any given woman. "No-one To Do Chores" includes reasons related to tending to the home, children, sick family members, or livestock. "Health Issue" includes own illness, own pregnancy, or illness in the family. "Distance" denotes dropout due to the loom center being too far from one's home. "Husb./Family Wouldn't Allow" reflects dropout due to husband or family members not allowing participation, while "Did Not Like Work" denotes dropout due to not liking the work.

Figure A8: Results Summary Including *Info* Sub-Treatments



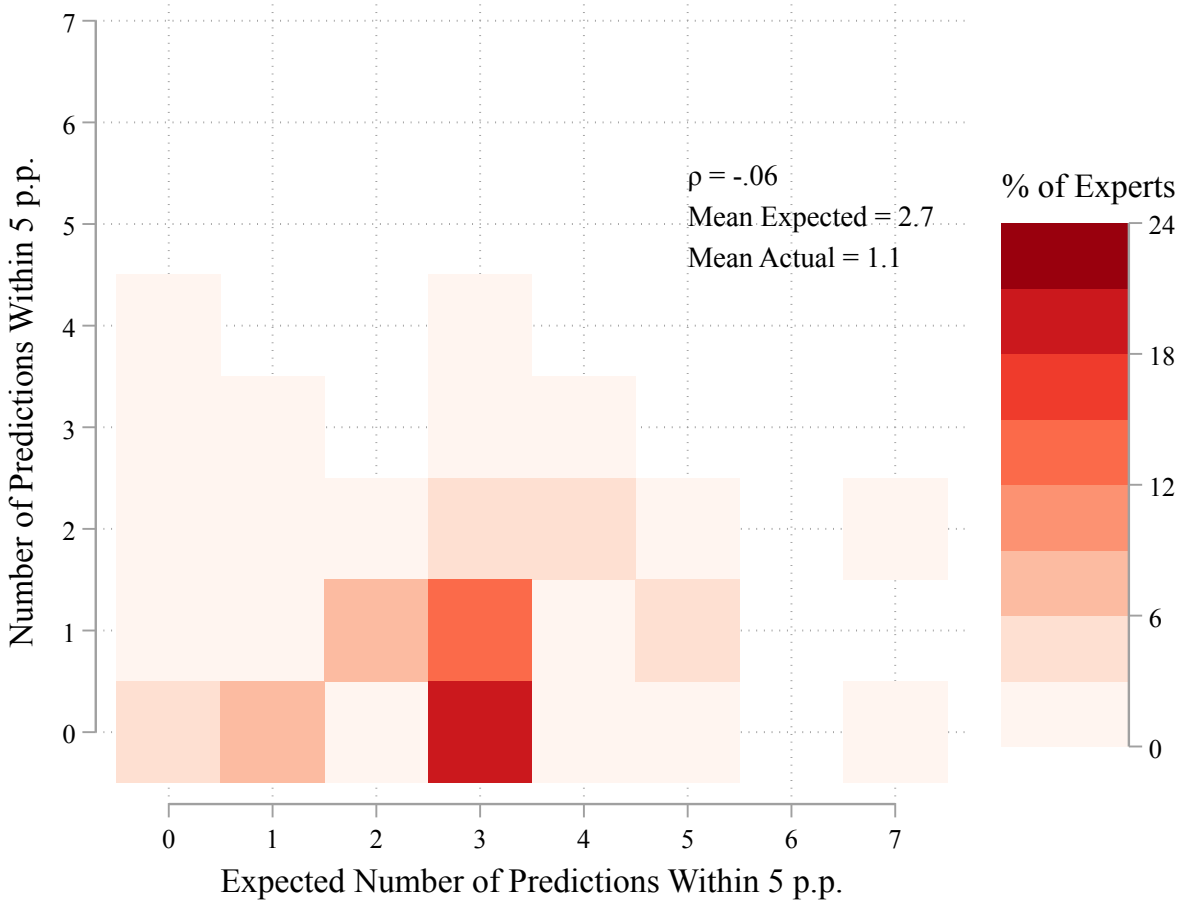
Notes: The figure visualizes enrollment rates across the eight treatments, including the two sub-treatments (with half the sample size of the other cells) of the *Info* treatment. The figure is derived from a regression ($N = 495$) of enrollment on treatment indicators, modified strata effects (i.e. village fixed effects and a dummy variable for OBC caste status), and the following controls: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for wife has worked for income in the last three months, (10) indicator for husband has worked for income in the last three months, (11) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (12) husband's answer to how appropriate for woman in household to work as weaver (zero to two), with each set to the mean when missing. Crosses denote mean predictions from the full sample of 70 experts. p_z is the p-value from a test of whether the treatment effect relative to the control group category (the husband gets the ticket, the wife doesn't know) is equal to zero. p_e is the p-value from a test of whether the treatment effect is equal to the mean expert prediction.

Figure A9: Reasons for a Negative Effect of Discussion from Local Experts



Notes: The figure shows the distribution of reasons for a negative effect of discussion provided by local experts, among those who provided a reason. Only experts given the scenario corresponding to our experiment (i.e. surveyor would be present for discussion and opportunity was Obeetee's female weaving program) are included. After ranking the intervention strategies, experts were asked if they could think of any reason discussion might have a negative effect. Surveyors matched their responses to one or more answer options, but did not read the answer options aloud. 51% of experts said they could not think of any reasons, and this figure presents responses for the 113 experts who could think of one or more reasons. The veto mechanism category includes any of the following options: "a spouse opposed to enrolling would feel more comfortable refusing to enroll"; "spouses would become more committed to their views, so someone opposed to enrolling wouldn't change their mind"; and "spouses might disagree about enrolling and only enroll if they are both supportive during the discussion." "Awkward/argument-provoking" includes the options "would be awkward/uncomfortable" and "would lead to argument." The wrongly interfering category includes the options "husband would feel the approach is wrongly interfering in the household's decision" and "wife would feel the approach is wrongly interfering in the household's decision." "Prevents strategic behavior" includes the options "would prevent spouse in favor of enrolling from bringing up the topic at a good time" and "would prevent spouse in favor of enrolling from misrepresenting information about the opportunity." The surveyor category includes the options "surveyor's presence would remind couple of norms against women's work" and "surveyor's presence would remind couple of the norm that husbands make decisions." The single deeper deliberation option is "would lead couple to think more carefully and realize the opportunity isn't good." Responses that surveyors could not fit into any of the mentioned options were categorized as "other."

Figure A10: Academic Experts Are Overconfident



Notes: The figure shows a heatmap of the percentage of academic experts with different levels of confidence in their predictions and different levels of ex-post accuracy. Confidence is measured from zero to seven on the x-axis as the answer to the question: *How many of your seven predictions do you expect are 5 percentage points or less from the actual enrollment rate in each treatment group?* Accuracy is measured from zero to seven on the y-axis as the actual number of predictions within 5 percentage points of the actual enrollment rate. Actual enrollment rates come from a regression of the enrollment indicator on seven treatment dummy variables (the omitted category being husband gets ticket, *NoInfo*), modified strata fixed effects, and controls. Suppose an expert predicts that treatment x will have an enrollment rate that is Y percentage points higher than the reference category (husband gets ticket, *NoInfo*). This prediction counts as one within 5 percentage points of the true enrollment rate if abs $\left(Y - \hat{\beta}_x\right) \leq 5$, where $\hat{\beta}_x$ is the estimated effect of x relative to the omitted category. ρ denotes Pearson's correlation coefficient between the two variables.

Table A1: Robustness of Main Results to Including Surveyor Fixed Effects

	Enrolled			
	(1)	(2)	(3)	(4)
Info	-0.03	-0.03	-0.03	-0.03
	(0.04)	(0.04)	(0.06)	(0.06)
Discuss	-0.08**	-0.10***	-0.10*	-0.12**
	(0.04)	(0.04)	(0.05)	(0.05)
WifeTicket	-0.00	0.01	-0.02	0.00
	(0.03)	(0.03)	(0.06)	(0.06)
Info × *WifeTicket*			0.00	-0.01
			(0.08)	(0.08)
Discuss × *WifeTicket*			0.04	0.04
			(0.08)	(0.07)
Observations	495	495	495	495
NoInfo Mean	0.18	0.18		
HusbTicket Mean	0.15	0.15		
NoInfo & HusbTicket Mean			0.19	0.19
p(*Info* = *Discuss*)	0.21	0.06*		
p(*Info* = Expert Mean)	0.04**	0.02**		
p(*Discuss* = Expert Mean)	<0.01***	<0.01***		
p(*Discuss* – *Info* = Expert Mean)	<0.01***	<0.01***		
p(*WifeTicket* = Expert Mean)	0.56	0.83		
p(*Info* × *WifeTicket* = *Discuss* × *WifeTicket*)			0.61	0.48
p(*Info* × *WifeTicket* = Expert Mean)			0.73	0.89
p(*Discuss* × *WifeTicket* = Expert Mean)			0.43	0.43
p(*Discuss* × *WifeTicket* – *Info* × *WifeTicket* = Expert Mean)			0.66	0.52
Modified Strata	Yes	Yes	Yes	Yes
Surveyor FE	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes

Notes: Robust standard errors in parentheses. Enrolled is an indicator for enrolling on enrollment day. Modified Strata controls are village fixed effects and an indicator for OBC status. Surveyor FE are indicators for each surveyor that take the value of 1 if the surveyor administered the intervention to a member of the couple, and 0 otherwise. In 3% of couples, we are missing information on the surveyor for one or both spouses, so we also include an indicator for either spouse's surveyor being missing (setting the other surveyor indicators to zero in such cases). Controls are: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for wife has worked for income in the last three months, (10) indicator for husband has worked for income in the last three months, (11) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (12) husband's answer to how appropriate for woman in household to work as weaver (zero to two), with each set to the mean when missing. *** p<0.01, ** p<0.05, * p<0.1.

Table A2: Effects on Enrollment, Single-Couple Households Only

	Enrolled			
	(1)	(2)	(3)	(4)
Info	-0.01	-0.04	0.01	-0.02
	(0.04)	(0.04)	(0.07)	(0.06)
Discuss	-0.06	-0.11**	-0.08	-0.13**
	(0.04)	(0.04)	(0.06)	(0.06)
WifeTicket	-0.00	0.00	-0.00	-0.01
	(0.03)	(0.03)	(0.07)	(0.06)
Info × *WifeTicket*			-0.04	-0.03
			(0.09)	(0.09)
Discuss × *WifeTicket*			0.03	0.05
			(0.09)	(0.09)
Observations	397	397	397	397
NoInfo Mean	0.19	0.19		
HusbTicket Mean	0.15	0.15		
NoInfo & HusbTicket Mean			0.19	0.19
p(*Info* = *Discuss*)	0.22	0.07*		
p(*Info* = Expert Mean)	0.13	0.03**		
p(*Discuss* = Expert Mean)	<0.01***	<0.01***		
p(*Discuss* – *Info* = Expert Mean)	<0.01***	<0.01***		
p(*WifeTicket* = Expert Mean)	0.57	0.67		
p(*Info* × *WifeTicket* = *Discuss* × *WifeTicket*)			0.39	0.31
p(*Info* × *WifeTicket* = Expert Mean)			0.91	0.98
p(*Discuss* × *WifeTicket* = Expert Mean)			0.52	0.41
p(*Discuss* × *WifeTicket* – *Info* × *WifeTicket* = Expert Mean)			0.42	0.34
Modified Strata	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes

Notes: Robust standard errors in parentheses. Sample drops couples from households in which more than one couple participated in the experiment. Enrolled is an indicator for enrolling on enrollment day. Modified Strata controls are village fixed effects and an indicator for OBC status. Controls are: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for wife has worked for income in the last three months, (10) indicator for husband has worked for income in the last three months, (11) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (12) husband's answer to how appropriate for woman in household to work as weaver (zero to two), with each set to the mean when missing. *** p<0.01, ** p<0.05, * p<0.1.

Table A3: No Detectable Enrollment Effect of *Full Info* Versus *Ticket Info*

	Non-Ticketed Knows Pay				Non-Ticketed Job Desirability				Enrolled			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
FullInfo	0.33***	0.33***	0.37***	0.37***	0.48**	0.40**	0.60**	0.53**	-0.02	-0.02	0.01	0.03
	(0.08)	(0.08)	(0.11)	(0.11)	(0.19)	(0.19)	(0.26)	(0.24)	(0.05)	(0.05)	(0.08)	(0.07)
WifeTicket	0.06	0.06	0.09	0.10	-0.02	0.04	0.13	0.19	-0.00	0.01	0.01	0.03
	(0.04)	(0.04)	(0.12)	(0.12)	(0.10)	(0.11)	(0.30)	(0.33)	(0.03)	(0.03)	(0.08)	(0.08)
FullInfo × *WifeTicket*			-0.08	-0.08			-0.27	-0.29			-0.05	-0.08
			(0.15)	(0.15)			(0.38)	(0.39)			(0.11)	(0.11)
Observations	414	414	414	414	234	234	234	234	495	495	495	495
TicketInfo Mean	0.51	0.51	0.51	0.51	0.85	0.85	0.85	0.85	1.00	1.00	1.00	1.00
NoInfo	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Discuss	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NoInfo × *WifeTicket*	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Discuss × *WifeTicket*	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Modified Strata	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors in parentheses. The outcomes in columns (1) to (8) come from the endline survey. Only responses of non-ticketed spouses, i.e. those whose information set was affected by the sub-treatment within *Info*, are included in these columns. Knows pay is a dummy for knowing the initial pay for the women's weaving program. Job Desirability equals zero if the respondent considers the job to be completely undesirable, one for somewhat desirable or somewhat undesirable, and two for completely desirable. The Job Desirability question was asked only for the final three centers. Enrolled is an indicator for enrolling on enrollment day. Modified Strata controls include village fixed effects and an indicator for OBC status. Controls are: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for wife has worked for income in the last three months, (10) indicator for husband has worked for income in the last three months, (11) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (12) husband's answer to how appropriate for woman in household to work as weaver (zero to two), with each set to the mean when missing. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Representativeness of Experts Surveyed

	All Experts Contacted	Experts That Predicted
PhD Student	0.03	0.06
Assistant Professor	0.20	0.34
Associate Professor	0.18	0.17
Full Professor	0.32	0.23
Other Position	0.26	0.20
Years Since PhD	14.92	12.09
Has Google Scholar Profile	0.84	0.87
Google Scholar Citations	6104.09	3506.54
Number of Experts	361	70

Notes: The first column shows the mean characteristics for the full set of experts we invited to complete the expert survey. The second column shows the mean characteristics for only those experts that completed the survey and gave predictions. Years Since PhD is calculated only for those with known PhD completion years and for PhD candidates, for which we set Years Since PhD to zero. Years Since PhD is then non-missing for 350 of 361 experts in the first column, and for 69 of 70 experts in the second column. Google Scholar Citations is the total citations as of September 2021, only for those with a Google Scholar profile.

Table A5: Additional Analyses of Local Expert Survey

	Predicted Rank (1-3, 1 is Best)				
	(1)	(2)	(3)	(4)	(5)
WifeAlone	0.49***	0.49***	0.51*	0.49***	0.49***
	(0.07)	(0.08)	(0.26)	(0.07)	(0.09)
Discuss	-0.71***	-0.76***	-0.79***	-0.71***	-0.76***
	(0.07)	(0.10)	(0.30)	(0.07)	(0.11)
Female		-0.03			
		(0.08)			
WifeAlone × Female		0.00			
		(0.13)			
Discuss × Female		0.09			
		(0.14)			
Age			-0.00		
			(0.00)		
WifeAlone × Age			-0.00		
			(0.01)		
Discuss × Age			0.00		
			(0.01)		
No Surveyor				-0.03	
				(0.06)	
WifeAlone × No Surveyor				-0.04	
				(0.10)	
Discuss × No Surveyor				0.13	
				(0.11)	
Education Topic					0.06
					(0.08)
WifeAlone × Education Topic					-0.27**
					(0.13)
Discuss × Education Topic					0.10
					(0.16)
Observations	723	723	720	1371	714
Omitted Group Mean	2.07	2.09	2.07	2.07	2.09
Strata	N/A	N/A	N/A	Yes	Yes

Notes: Data are at the respondent × intervention level. Standard errors clustered by respondent in parentheses. Predicted Rank is the respondent's predicted rank of the intervention, where lower ranks mean greater effectiveness. Female is an indicator for the respondent being female, while age is the respondent's age. No Surveyor is an indicator for the discussion intervention described in the hypothetical scenario not having a surveyor present for the couple's discussion. Education Topic is an indicator for the topic of the decision being an opportunity for girls' tutoring rather than Obeetee's female weaving program. No Surveyor was randomly assigned across all respondents, while Education Topic was randomly assigned among male respondents (all female respondents were given the decision about Obeetee's program). The omitted group mean is the mean prediction for the *HusbAlone* intervention in columns (1) and (3), the mean prediction for *HusbAlone* from male respondents in column (2), the mean prediction for *HusbAlone* from respondents assigned to the surveyor condition in column (4), and the mean prediction for *HusbAlone* from respondents assigned to the Obeetee topic in column (5). Strata are loom center fixed effects. Columns (1)-(3) are restricted to respondents given the scenario corresponding to our experiment (i.e. surveyor present and Obeetee topic). Column (4) is restricted to respondents given the Obeetee topic. Column (5) is restricted to male respondents (Obeetee vs. Education topic did not vary among female respondents) and respondents given a scenario where the surveyor was present for discussion. *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Baseline Characteristics and Balance

	Attrit:	Age of		No Education			Years Married			Employed		Women Weaver Preference	
	Baseline	Wife	Husb.	Wife	Husb.	Fertility	10+	5-10	0-5	Wife	Husb.	Wife	Husb.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Info	-0.06	0.52	0.12	0.07	0.07*	0.10	0.03	-0.01	-0.02	0.01	0.02	-0.12	-0.00
	(0.04)	(0.57)	(0.60)	(0.05)	(0.04)	(0.18)	(0.05)	(0.05)	(0.05)	(0.03)	(0.04)	(0.09)	(0.09)
Discuss	-0.03	0.83	0.46	0.03	0.03	0.19	0.00	0.04	-0.04	0.09**	0.03	-0.12	0.24**
	(0.04)	(0.58)	(0.62)	(0.05)	(0.04)	(0.17)	(0.06)	(0.05)	(0.05)	(0.04)	(0.04)	(0.09)	(0.10)
WifeTicket	0.02	0.28	0.53	0.03	-0.04	-0.07	-0.04	-0.01	0.04	0.05*	0.01	0.07	-0.04
	(0.03)	(0.48)	(0.51)	(0.04)	(0.03)	(0.14)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.08)	(0.08)
Observations	732	490	493	490	493	490	485	485	485	490	493	490	493
p(*Info*=*Discuss*)	0.41	0.61	0.59	0.42	0.31	0.62	0.61	0.41	0.76	0.03**	0.74	0.97	0.01**
Joint RI p-value: *Info*			0.72										
Joint RI p-value: *Discuss*			0.07*										
Joint RI p-value: *Info*=*Discuss*			0.22										
Joint RI p-value: *WifeTicket*			0.30										
Outcome Mean	.32	26	29	.39	.13	2.3	.4	.32	.28	.13	.82	1.1	.81
Modified Strata	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. Joint RI p-values are randomization inference tests (500 draws) for the joint significance of *Info*, *Discuss*, and *WifeTicket*, and the equality of *Info* with *Discuss*, across columns 2 to 13. Modified Strata controls include village fixed effects and an indicator for OBC status. Outcome variables are: (1) indicator for attrited (at least one spouse did not complete the Baseline), (2) age of wife, (3) age of husband, (4) indicator for wife has no education, (5) indicator for husband has no education, (6) number of children, (7) indicator for married ten or more years, (8) indicator for married five to ten years, (9) indicator for married zero to five years, (10) indicator for wife has worked for income in the last three months, (11) indicator for husband has worked for income in the last three months, (12) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (13) husband's answer to how appropriate for woman in household to work as weaver (zero to two). *** p<0.01, ** p<0.05, * p<0.1.

Table A7: Baseline Characteristics and Balance: One-Visit Discussion Treatment

	Attrit:	Age of		No Education			Years Married			Employed		Women Weaver Preference	
	Baseline	Wife	Husb.	Wife	Husb.	Fertility	10+	5-10	0-5	Wife	Husb.	Wife	Husb.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Info	-0.06	1.00	1.42	0.10	0.04	0.10	0.11	-0.04	-0.07	0.02	0.04	-0.12	0.02
	(0.07)	(1.06)	(1.20)	(0.10)	(0.08)	(0.33)	(0.10)	(0.09)	(0.08)	(0.05)	(0.08)	(0.15)	(0.17)
Discuss	0.01	-0.44	-0.82	-0.02	-0.02	-0.03	-0.07	-0.03	0.10	0.11*	-0.01	-0.25	0.15
	(0.08)	(1.05)	(1.13)	(0.10)	(0.08)	(0.30)	(0.10)	(0.10)	(0.10)	(0.07)	(0.09)	(0.18)	(0.18)
WifeTicket	0.03	0.44	1.54	0.09	-0.03	0.13	0.03	-0.01	-0.02	0.07	-0.05	0.06	0.02
	(0.06)	(0.89)	(0.97)	(0.08)	(0.06)	(0.26)	(0.08)	(0.08)	(0.07)	(0.05)	(0.07)	(0.13)	(0.14)
Observations	238	151	152	151	152	151	149	149	149	151	152	151	152
p(*Info=Discuss*)	0.38	0.20	0.05**	0.22	0.47	0.67	0.07*	0.89	0.07*	0.18	0.60	0.42	0.45
Joint RI p-value: *Info*			0.58										
Joint RI p-value: *Discuss*			0.47										
Joint RI p-value: *Info=Discuss*			0.34										
Joint RI p-value: *WifeTicket*			0.52										
Outcome Mean	.36	26	30	.4	.18	2.2	.43	.28	.3	.11	.77	1.3	.95
Modified Strata	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. Sample includes only those surveyed during the period of the first *Discuss* treatment protocol. Joint RI p-values are randomization inference tests (500 draws) for the joint significance of *Info*, *Discuss*, and *WifeTicket*, and the equality of *Info* with *Discuss*, across columns 2 to 13. Modified Strata controls include village fixed effects and an indicator for OBC status. Outcome variables are: (1) indicator for attrited (at least one spouse did not complete the Baseline), (2) age of wife, (3) age of husband, (4) indicator for wife has no education, (5) indicator for husband has no education, (6) number of children, (7) indicator for married ten or more years, (8) indicator for married five to ten years, (9) indicator for married zero to five years, (10) indicator for wife has worked for income in the last three months, (11) indicator for husband has worked for income in the last three months, (12) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (13) husband's answer to how appropriate for woman in household to work as weaver (zero to two). *** p<0.01, ** p<0.05, * p<0.1.

Table A8: Baseline Characteristics and Balance: Two-Visit Discussion Treatment

	Attrit:	Age of		No Education			Years Married			Employed		Women Weaver Preference	
	Baseline	Wife	Husb.	Wife	Husb.	Fertility	10+	5-10	0-5	Wife	Husb.	Wife	Husb.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Info	-0.01	0.17	-0.80	0.04	0.07*	0.07	-0.02	0.01	0.01	0.01	0.02	-0.12	-0.02
	(0.05)	(0.67)	(0.66)	(0.06)	(0.04)	(0.21)	(0.07)	(0.06)	(0.06)	(0.04)	(0.05)	(0.12)	(0.11)
Discuss	-0.01	1.24*	0.86	0.04	0.04	0.25	0.03	0.06	-0.09	0.08*	0.05	-0.07	0.26**
	(0.05)	(0.68)	(0.73)	(0.06)	(0.04)	(0.20)	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.11)	(0.12)
WifeTicket	0.02	0.09	-0.13	-0.01	-0.05	-0.17	-0.07	-0.00	0.08	0.04	0.03	0.07	-0.07
	(0.04)	(0.56)	(0.58)	(0.05)	(0.03)	(0.17)	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.09)	(0.09)
Observations	433	339	341	339	341	339	336	336	336	339	341	339	341
p(*Info=Discuss*)	0.94	0.13	0.02**	0.93	0.50	0.38	0.49	0.42	0.10	0.10	0.42	0.64	0.01**
Joint RI p-value: *Info*			0.53										
Joint RI p-value: *Discuss*			0.17										
Joint RI p-value: *Info=Discuss*			0.19										
Joint RI p-value: *WifeTicket*			0.45										
Outcome Mean	.21	26	29	.39	.11	2.4	.38	.34	.28	.14	.84	1	.74
Modified Strata	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. Sample includes only those surveyed during the period of the second *Discuss* treatment protocol. Joint RI p-values are randomization inference tests (500 draws) for the joint significance of *Info*, *Discuss*, and *WifeTicket*, and the equality of *Info* with *Discuss*, across columns 2 to 13. Modified Strata controls include village fixed effects and an indicator for OBC status. Outcome variables are: (1) indicator for attrited (at least one spouse did not complete the Baseline), (2) age of wife, (3) age of husband, (4) indicator for wife has no education, (5) indicator for husband has no education, (6) number of children, (7) indicator for married ten or more years, (8) indicator for married five to ten years, (9) indicator for married zero to five years, (10) indicator for wife has worked for income in the last three months, (11) indicator for husband has worked for income in the last three months, (12) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (13) husband's answer to how appropriate for woman in household to work as weaver (zero to two). *** p<0.01, ** p<0.05, * p<0.1.

Table A9: Wife's Baseline Employment Does Not Predict Enrollment

	Enrolled			
	(1)	(2)	(3)	(4)
Wife Employed	-0.005	0.013	-0.009	0.005
	(0.047)	(0.047)	(0.044)	(0.045)
Observations	490	490	490	490
Outcome Mean	0.15	0.15	0.15	0.15
Modified Strata	No	Yes	Yes	Yes
Controls	No	No	Yes	Yes
Treatments	No	No	No	Yes

Notes: Robust standard errors in parentheses. Enrolled is an indicator for enrolling on enrollment day. Wife Employed is an indicator from the baseline survey for whether the wife had worked for income in the last three months. Modified Strata controls include village fixed effects and an indicator for OBC status. Controls are: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for husband has worked for income in the last three months, (10) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (11) husband's answer to how appropriate for woman in household to work as weaver (zero to two), with each set to the mean when missing. Treatments are *Info*, *Discuss*, and *WifeTicket* indicators. *** p<0.01, ** p<0.05, * p<0.1.

Table A10: Negative Effects Are Not Driven by Working-At-Baseline Women

	Enrolled			
	(1)	(2)	(3)	(4)
Info	-0.03	-0.04	-0.05	-0.05
	(0.04)	(0.04)	(0.04)	(0.04)
Discuss	-0.08**	-0.10***	-0.09**	-0.12***
	(0.04)	(0.03)	(0.04)	(0.04)
WifeTicket	-0.00	0.01	-0.01	0.00
	(0.03)	(0.03)	(0.03)	(0.03)
Info × Wife Employed			0.14	0.07
			(0.15)	(0.13)
Discuss × Wife Employed			0.11	0.11
			(0.11)	(0.11)
WifeTicket × Wife Employed			0.04	0.02
			(0.11)	(0.10)
Observations	495	495	490	490
p(*Info* = *Discuss*)	0.22	0.06*	0.20	0.04**
Modified Strata	Yes	Yes	Yes	Yes
Wife Employed Control	No	No	Yes	Yes
Controls	No	Yes	No	Yes

Notes: Robust standard errors in parentheses. Enrolled is an indicator for enrolling on enrollment day. Wife Employed is an indicator for the wife having worked for income in the last three months (reported at baseline). Modified Strata controls are village fixed effects and an indicator for OBC status. Controls are: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for wife has worked for income in the last three months, (10) indicator for husband has worked for income in the last three months, (11) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (12) husband's answer to how appropriate for woman in household to work as weaver (zero to two), with each set to the mean when missing. *** p<0.01, ** p<0.05, * p<0.1.

Table A11: Baseline Characteristics and Balance in Endline Sample

	Attrit:		Age of		No Education			Years Married			Employed		Women Weaver Preference	
	Endl., Wife	Endl., Husb.	Wife	Husb.	Wife	Husb.	Fertility	10+	5-10	0-5	Wife	Husb.	Wife	Husb.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Info	-0.05	0.06	0.30	0.02	0.03	0.07*	0.07	0.01	-0.02	0.00	0.02	0.01	-0.14	-0.04
	(0.04)	(0.04)	(0.59)	(0.63)	(0.06)	(0.04)	(0.19)	(0.06)	(0.06)	(0.05)	(0.04)	(0.05)	(0.10)	(0.10)
Discuss	0.05	0.02	0.96	0.39	-0.02	0.03	0.18	-0.02	0.06	-0.03	0.10**	0.06	-0.06	0.23**
	(0.04)	(0.04)	(0.64)	(0.68)	(0.06)	(0.04)	(0.19)	(0.06)	(0.06)	(0.06)	(0.04)	(0.05)	(0.10)	(0.11)
WifeTicket	0.01	0.01	0.17	0.24	0.04	-0.05	-0.07	-0.03	-0.01	0.05	0.07**	-0.00	0.10	-0.05
	(0.03)	(0.03)	(0.51)	(0.55)	(0.05)	(0.03)	(0.16)	(0.05)	(0.05)	(0.05)	(0.03)	(0.04)	(0.08)	(0.08)
Observations	495	495	411	411	411	411	411	409	409	409	411	411	411	411
p(*Info*=*Discuss*)	0.02**	0.30	0.30	0.60	0.47	0.34	0.58	0.54	0.20	0.51	0.06*	0.29	0.42	0.01***
Joint RI p-value: *Info*			0.76											
Joint RI p-value: *Discuss*			0.14											
Joint RI p-value: *Info*=*Discuss*			0.14											
Joint RI p-value: *WifeTicket*			0.30											
Outcome Mean	.17	.16	26	29	.4	.13	2.3	.4	.32	.28	.12	.82	1.2	.83
Modified Strata	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. Joint RI p-values are randomization inference tests (500 draws) for the joint significance of *Info*, *Discuss*, and *WifeTicket*, and the equality of *Info* with *Discuss*, across columns 3 to 14. Outcome variables are: (1) indicator for wife attrited (completed Baseline but not Endline), (2) indicator for husband attrited (completed Baseline but not Endline), (3) age of wife, (4) age of husband, (5) indicator for wife has no education, (6) indicator for husband has no education, (7) number of children, (8) indicator for married ten or more years, (9) indicator for married five to ten years, (10) indicator for married zero to five years, (11) indicator for wife has worked for income in the last three months, (12) indicator for husband has worked for income in the last three months, (13) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (14) husband's answer to how appropriate for woman in household to work as weaver (zero to two). *** p<0.01, ** p<0.05, * p<0.1.

Table A12: Village-Level Appropriateness of Women Weavers Predicts Enrollment Above and Beyond Own Perceived Appropriateness

	Enrolled	
	(1)	(2)
Wife: Weaving Appropriate for Women in HH?	0.04**	0.03**
	(0.02)	(0.01)
Husband: Weaving Appropriate for Women in HH?	0.11***	0.11***
	(0.02)	(0.02)
Village-level Average: Weaving Appropriate for Women in HH?		0.32*
		(0.13)
Observations	495	495
Wild Bootstrap p-value		0.03
OBC Dummy	Yes	Yes
Controls	Yes	Yes
Info, *Discuss*, *WifeTicket*	Yes	Yes

Notes: Column 1 uses robust standard errors, column 2 uses village-clustered standard errors. Since there are only six villages we also report the Wild bootstrap p-value to test the coefficient on the village-level appropriateness of women weavers. Enrolled is an indicator for enrolling on enrollment day. The coefficients shown are for the wife and husband's answer to how appropriate it would be for a woman in their household to work as a weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), and the village-level average of these answers (average over both wives and husbands). Controls are: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for wife has worked for income in the last three months, (10) indicator for husband has worked for income in the last three months. *** p<0.01, ** p<0.05, * p<0.1.

Table A13: "First Stage" Effects on Information and Discussion

	Non-Ticketed Knows				N Discussions	
	(1)	(2)	(3)	(4)	(5)	(6)
Info	0.10**	0.10*	0.08	0.08	-0.00	0.09
	(0.05)	(0.05)	(0.07)	(0.07)	(0.35)	(0.35)
WifeTicket	0.05	0.05	-0.03	-0.02	-0.12	-0.07
	(0.04)	(0.04)	(0.08)	(0.08)	(0.31)	(0.31)
Info × *WifeTicket*			0.05	0.03		
			(0.10)	(0.11)		
Discuss					1.00**	0.99**
					(0.42)	(0.44)
Observations	414	414	414	414	829	829
NoInfo Mean	0.72	0.72			3.04	3.04
NoInfo & HusbTicket Mean			0.74	0.74		
p(*Info* = *Discuss*)					0.01***	0.02**
Modified Strata	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes
Discuss	Yes	Yes	Yes	Yes	Yes	Yes
Discuss × *WifeTicket*	No	No	Yes	Yes	No	No

Notes: Data are at the couple-level in columns (1)-(4) and individual-level in columns (5)-(6). Standard errors are robust in columns (1)-(4) and clustered by couple in columns (5)-(6). Non-Ticketed Knows is an indicator for the non-ticketed spouse knowing at endline that their spouse received an enrollment ticket. N Discussions is the number of discussions the couple had about the job opportunity, constructed using responses from the endline survey and surveyor reports on whether couples in the *Discuss* treatment discussed the opportunity in the designated time for discussion. Modified Strata controls are village fixed effects and an indicator for OBC status. Controls are: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for wife has worked for income in the last three months, (10) indicator for husband has worked for income in the last three months, (11) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (12) husband's answer to how appropriate for woman in household to work as weaver (zero to two), and (13) the gender of the respondent (in the final two columns only), with each set to the mean when missing. *** p<0.01, ** p<0.05, * p<0.1.

Table A14: Treatment Effects on Participation in the First Month

	In Program at End of First Month			
	(1)	(2)	(3)	(4)
Info	0.01	0.02	0.02	0.01
	(0.03)	(0.03)	(0.05)	(0.05)
Discuss	-0.05	-0.05*	-0.05	-0.06
	(0.03)	(0.03)	(0.04)	(0.04)
WifeTicket	-0.04	-0.03	-0.04	-0.04
	(0.02)	(0.02)	(0.05)	(0.04)
Info × *WifeTicket*			-0.00	0.00
			(0.07)	(0.06)
Discuss × *WifeTicket*			0.02	0.02
			(0.06)	(0.06)
Observations	495	495	495	495
Omitted Group Mean	0.12	0.12	0.12	0.12
p(*Info* = *Discuss*)	0.04**	0.02**		
Modified Strata	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes

Notes: Robust standard errors in parentheses. The outcome is an indicator for the wife having enrolled in the weaving program within the first month, and not having dropped out in that month. Modified Strata controls are village fixed effects and an indicator for OBC status. Controls are: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for wife has worked for income in the last three months, (10) indicator for husband has worked for income in the last three months, (11) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (12) husband's answer to how appropriate for woman in household to work as weaver (zero to two), with each set to the mean when missing. *** p<0.01, ** p<0.05, * p<0.1.

Table A15: Ticketed-More-Supportive Heterogeneity Not Driven by $WifeTicket$

	Enrolled							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Info × Ticketed More Supportive	-0.15	-0.16*			-0.17*	-0.18*		
	(0.09)	(0.09)			(0.10)	(0.10)		
Discuss × Ticketed More Supportive	-0.14	-0.20**			-0.15	-0.21**		
	(0.09)	(0.08)			(0.10)	(0.09)		
(*Info* or *Discuss*) × Ticketed More Supportive			-0.15*	-0.18**			-0.16*	-0.19**
			(0.08)	(0.08)			(0.09)	(0.08)
Info × Ticketed Less Supportive					-0.07	-0.05		
					(0.10)	(0.09)		
Discuss × Ticketed Less Supportive					-0.02	-0.03		
					(0.09)	(0.08)		
(*Info* or *Discuss*) × Ticketed Less Supportive							-0.04	-0.04
							(0.08)	(0.07)
Observations	489	489	489	489	489	489	489	489
Info	Yes	Yes	No	No	Yes	Yes	No	No
Discuss	Yes	Yes	No	No	Yes	Yes	No	No
Ticketed More Supportive	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Info or *Discuss*	No	No	Yes	Yes	No	No	Yes	Yes
Ticketed Less Supportive	No	No	No	No	Yes	Yes	Yes	Yes
WifeTicket	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
WifeTicket × *Info*	Yes	Yes	No	No	Yes	Yes	No	No
WifeTicket × *Discuss*	Yes	Yes	No	No	Yes	Yes	No	No
WifeTicket × (*Info* or *Discuss*)	No	No	Yes	Yes	No	No	Yes	Yes
Modified Strata	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Notes: This table presents the ticketed-more-supportive heterogeneity analyses, controlling for *WifeTicket* and its interaction with the other treatment variables. Robust standard errors in parentheses. Enrolled is an indicator for enrolling on enrollment day. Ticketed More (Less) Supportive is an indicator for the ticketed spouse rating women weaving as more (less) appropriate at baseline than the non-ticketed spouse did. Modified Strata controls include village fixed effects and an indicator for OBC status. Controls are: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for wife has worked for income in the last three months, (10) indicator for husband has worked for income in the last three months, (11) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (12) husband's answer to how appropriate for woman in household to work as weaver (zero to two), with each control variable set to the mean when missing. *** p<0.01, ** p<0.05, * p<0.1.

Table A16: Who Had Most Influence over the Enrollment Decision?

	Ticketed Spouses Only					
	Me	Me	Equal	Equal	Spouse	Spouse
	(1)	(2)	(3)	(4)	(5)	(6)
Info	-0.01	-0.02	0.04	0.04	-0.03	-0.02
	(0.06)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)
Discuss	-0.03	-0.05	0.05	0.05	-0.02	0.00
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Observations	407	407	407	407	407	407
Omitted Group Mean	.33	.33	.31	.31	.36	.36
Modified Strata	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes
WifeTicket	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. The three outcomes come from the endline question: *Out of you and your spouse, who would you say had the most influence over the decision of whether to enroll or not?* Me is an indicator for the respondent saying *Me*, Equal is an indicator for the respondent saying *Me and my spouse had same influence*, and Spouse is an indicator for the respondent saying *My spouse*. The sample includes only ticketed spouses, as non-ticketed spouses in *NoInfo* who did not know about the ticket (a selected sample) were not asked the question. Modified Strata controls include village fixed effects and an indicator for OBC status. Controls are: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for wife has worked for income in the last three months, (10) indicator for husband has worked for income in the last three months, (11) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (12) husband's answer to how appropriate for woman in household to work as weaver (zero to two), with each set to the mean when missing. *** p<0.01, ** p<0.05, * p<0.1.

Table A17: Other Mechanisms Analyses: Effects on Job Desirability, Spousal Disagreement, and Time Before Enrollment

	Job Desirability		Spousal Disagreement, Ticketed Spouses Only		Days Before Enrollment		Enrolled			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Info	0.18*	0.20**	0.03	-0.00	0.12	0.14			-0.03	-0.04
	(0.09)	(0.09)	(0.12)	(0.11)	(0.19)	(0.19)			(0.04)	(0.04)
Discuss	0.11	0.08	-0.04	-0.01	-0.24	-0.23			-0.07*	-0.10***
	(0.10)	(0.09)	(0.13)	(0.12)	(0.19)	(0.19)			(0.04)	(0.03)
Days Before Enrollment							0.01	0.01*	0.01	0.01*
							(0.01)	(0.01)	(0.01)	(0.01)
Observations	478	478	245	245	494	494	494	494	494	494
Omitted Group Mean	0.99	0.99	1.19	1.19	4.99	4.99			0.18	0.18
p(*Info* = *Discuss*)	0.43	0.18	0.53	0.95	0.04**	0.04**			0.29	0.09*
Modified Strata	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
WifeTicket	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes

Notes: Data are at the individual-level in columns (1) and (2), and at the couple-level elsewhere. Standard errors are clustered by couple in columns (1) and (2), and robust elsewhere. Job Desirability equals zero if the respondent considers the job to be completely undesirable, one for somewhat desirable or somewhat undesirable, and two for completely desirable. The Job Desirability question was asked only for the final three centers. Spousal Disagreement is the extent to which the respondent reported disagreeing with their spouse about whether the woman should enroll, measured on a 0-2 scale. The Spousal Disagreement question was asked only for the final three centers. The question was not asked of non-ticketed spouses in *NoInfo* who were unaware of the ticket; to avoid selection, we therefore estimate effects on this outcome among ticketed spouses only. Days Before Enrollment is the number of days between enrollment day and when the couple's ticket was given to the ticketed spouse. Enrolled is an indicator for enrolling on enrollment day. Modified Strata controls are village fixed effects and an indicator for OBC status. Controls are: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for wife has worked for income in the last three months, (10) indicator for husband has worked for income in the last three months, (11) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (12) husband's answer to how appropriate for woman in household to work as weaver (zero to two), and (13) the gender of the respondent (in the second column only), with each set to the mean when missing. *** p<0.01, ** p<0.05, * p<0.1.

Table A18: Effects by Version of Discussion Treatment Protocol

	Enrolled	
	(1)	(2)
Info	-0.07	-0.08
	(0.09)	(0.08)
Discuss	-0.18**	-0.19**
	(0.09)	(0.08)
Version 2	-0.04	-0.08
	(0.10)	(0.10)
Info × Version 2	0.06	0.06
	(0.10)	(0.09)
Discuss × Version 2	0.15	0.12
	(0.09)	(0.09)
Observations	495	495
Omitted Group Mean	0.36	0.36
p(*Info* × Version 2 = *Discuss* × Version 2)	0.28	0.44
Modified Strata	Yes	Yes
Controls	No	Yes
WifeTicket	Yes	Yes

Notes: Robust standard errors in parentheses. Enrolled is an indicator for enrolling on enrollment day. Version 2 is an indicator for being treated after the change to the protocol for the discussion treatment. This change was made midway through implementation in the third village. Modified Strata controls are village fixed effects and an indicator for OBC status. Controls are: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for wife has worked for income in the last three months, (10) indicator for husband has worked for income in the last three months, (11) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (12) husband's answer to how appropriate for woman in household to work as weaver (zero to two), with each set to the mean when missing. *** p<0.01, ** p<0.05, * p<0.1.

Table A19: Days Between Visits Does Not Predict Enrollment

	Enrolled	
	(1)	(2)
Number of Days Between Visits	-0.014	-0.009
	(0.018)	(0.017)
Observations	118	118
Outcome Mean	0.08	0.08
Modified Strata	Yes	Yes
Controls	No	Yes
WifeTicket	Yes	Yes

Notes: Robust standard errors in parentheses. Enrolled is an indicator for enrolling on enrollment day. Sample includes only those assigned to the two-visit version of *Discuss* (our revised protocol). Modified Strata controls are village fixed effects and an indicator for OBC status. Controls are: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for wife has worked for income in the last three months, (10) indicator for husband has worked for income in the last three months, (11) wife's answer to how appropriate for woman in household to work as weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate), (12) husband's answer to how appropriate for woman in household to work as weaver (zero to two), with each set to the mean when missing. *** p<0.01, ** p<0.05, * p<0.1.

Table A20: No Heterogeneity by Individual Preferences

	Enrolled	
	(1)	(2)
Info	-0.06	-0.06
	(0.05)	(0.05)
Discuss	-0.11***	-0.13***
	(0.04)	(0.04)
Wife's Weaving Preference (0 to 2)	0.01	-0.01
	(0.03)	(0.03)
Husband's Weaving Preference (0 to 2)	0.13***	0.13***
	(0.04)	(0.04)
Info × Wife's Weaving Preference (0 to 2)	0.05	0.06
	(0.04)	(0.04)
Discuss × Wife's Weaving Preference (0 to 2)	0.03	0.05
	(0.04)	(0.04)
Info × Husband's Weaving Preference (0 to 2)	-0.03	-0.05
	(0.05)	(0.05)
Discuss × Husband's Weaving Preference (0 to 2)	-0.03	-0.04
	(0.05)	(0.05)
Observations	489	489
p(*Info* + *Info* × Wife's Pref. = *Discuss* + *Discuss* × Wife's Pref.)	0.05**	0.02**
p(*Info* + *Info* × Husband's Pref. = *Discuss* + *Discuss* × Husband's Pref.)	0.20	0.26
Modified Strata	Yes	Yes
Controls	No	Yes
WifeTicket	Yes	Yes

Notes: Robust standard errors in parentheses. Enrolled is an indicator for enrolling on enrollment day. Wife's (Husband's) Weaving Preference is the wife's (husband's) answer to how appropriate it is for a woman in her household to work as a weaver (0 = inappropriate, 1 = somewhat appropriate, 2 = completely appropriate). Modified Strata controls include village fixed effects and an indicator for OBC status. Controls are: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband has no education, (5) number of children, (6) indicator for married ten or more years, (7) indicator for married five to ten years, (8) indicator for married zero to five years, (9) indicator for wife has worked for income in the last three months, (10) indicator for husband has worked for income in the last three months, with each set to the mean when missing. *** p<0.01, ** p<0.05, * p<0.1.

Table A21: Wives That Discuss Work With Their Husbands Are Less Likely to Work

	Wife is currently working for pay (=0/1)					
	(1)	(2)	(3)	(4)	(5)	(6)
Couple discusses work (0-2)	-0.011***	-0.015***	-0.011***	-0.008*	-0.022*	-0.031**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.013)	(0.015)
Couple discusses spending (0-2)	0.020***	0.018***	0.007*	0.008	0.013	0.027*
	(0.004)	(0.004)	(0.004)	(0.005)	(0.013)	(0.014)
Couple discusses politics (0-2)	0.006	0.006*	0.007*	0.005	0.003	0.008
	(0.004)	(0.004)	(0.004)	(0.004)	(0.012)	(0.013)
Observations	36,673	36,572	36,572	36,567	7,246	6,152
Outcome Mean	0.24	0.24	0.24	0.24	0.17	0.18
Age and Education	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	No	No	No	No
Religion and Caste FE	No	Yes	Yes	Yes	Yes	Yes
Children, In-Laws, Marriage Age	No	Yes	Yes	Yes	Yes	Yes
District FE	No	No	Yes	No	No	No
Primary Sampling Unit FE	No	No	No	Yes	No	No
Household FE	No	No	No	No	Yes	Yes
Husband Work Hours	No	No	No	No	No	Yes

Notes: Robust standard errors in parentheses. The table uses data from the most recent wave (2011/12) of the Indian Human Development Survey. The outcome variable is a dummy variable equal to one if the wife answers that she is currently working for pay. The three displayed independent variables are from the question: Do you and your husband talk about... [things that happen at work or on the farm / about what to spend money on / about things that happen in the community such as elections or politics] with answers Never = 0, Sometimes = 1, and Often = 2. Other control variables are (1) wife's age, (2) wife's number of years of education, (3) number of living children, (4) dummy variables for whether mother-in-law and father-in-law are alive, (5) wife's age at marriage, and (6) husband's total yearly hours worked. *** p<0.01, ** p<0.05, * p<0.1.

Table A22: Wives That Discuss Work With Their Husbands Are Less Likely to Work: Alternative Identification Assumption

	Wife is currently working for pay (=0/1)					
	(1)	(2)	(3)	(4)	(5)	(6)
Couple discusses work (0-2)	-0.021***	-0.025***	-0.018***	-0.014**	-0.028	-0.045**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.019)	(0.021)
Mean of discussion variables (0-2)	0.035***	0.034***	0.022***	0.019**	0.024	0.050*
	(0.007)	(0.007)	(0.008)	(0.008)	(0.025)	(0.027)
Observations	36,673	36,572	36,572	36,567	7,246	6,152
Outcome Mean	0.24	0.24	0.24	0.24	0.17	0.18
Age and Education	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	No	No	No	No
Religion and Caste FE	No	Yes	Yes	Yes	Yes	Yes
Children, In-Laws, Marriage Age	No	Yes	Yes	Yes	Yes	Yes
District FE	No	No	Yes	No	No	No
Primary Sampling Unit FE	No	No	No	Yes	No	No
Household FE	No	No	No	No	Yes	Yes
Husband Work Hours	No	No	No	No	No	Yes

Notes: Robust standard errors in parentheses. The table uses data from the most recent wave (2011/12) of the Indian Human Development Survey. The outcome variable is a dummy variable equal to one if the wife answers that she is currently working for pay. Couple discusses work is from the question: Do you and your husband talk about things that happen at work or on the farm? with answers Never = 0, Sometimes = 1, and Often = 2. Mean of discussion variables is the mean of Couple discusses work (0-2), Couple discusses spending (0-2), and Couple discusses politics (0-2). Other control variables are (1) wife's age, (2) wife's number of years of education, (3) number of living children, (4) dummy variables for whether mother-in-law and father-in-law are alive, (5) wife's age at marriage, and (6) husband's total yearly hours worked. *** p<0.01, ** p<0.05, * p<0.1.

Table A23: The Predictions of Experienced Researchers Are Less Accurate

	Confidence (0 to 7)			Mean Prediction Accuracy		
	(1)	(2)	(3)	(4)	(5)	(6)
Associate Professor	.625			-.84		
	(.636)			(1.85)		
Full Professor	-.458			-3.94**		
	(.445)			(1.77)		
Other Position	.0694			-6.44*		
	(.542)			(3.29)		
Has Google Scholar Profile		-.712	-.606		-8.97	-11.1
		(1.55)	(1.61)		(8.46)	(8.15)
Google Scholar Citation Ranking (1 to 61)		.00014	-.00103		.0962*	.124**
		(.0111)	(.0117)		(.0527)	(.0612)
Above-Median Time on Survey			.143			-3.47
			(.414)			(2.63)
Confidence						.651
						(.514)
Observations			70			
Outcome Mean		2.73			-14.62	

Notes: Robust standard errors in parentheses. Confidence is the number of predictions (out of 7) that the expert expected to be within 5 percentage points of the actual enrollment rate. Mean Prediction Accuracy is the negative of the mean absolute forecast error across the 7 predictions. The omitted academic rank category is Assistant Professor, while Other Position includes, among others, PhD students and postdoctoral fellows. Has Google Scholar Profile is a dummy variable equal to one if the expert has a Google Scholar profile. Google Scholar Citation Ranking is the rank among the 61 experts with a Scholar profile of total Google Scholar citations as of September 2021 (lower means more citations). Above-Median Time on Survey is a dummy variable equal to one if the expert spent above-median time completing the survey (a proxy for effort). *** p<0.01, ** p<0.05, * p<0.1.

B Secondary Data Description

Development and Decision-Making. Figure 1 shows cross-country correlations of household decision-making with per capita income. Each panel includes the 57 countries with a Demographic and Health Survey (DHS) containing household decision-making variables. For each country we plot data from the most recent DHS survey.

Household Decision-Making. The DHS contains five questions that measure which household members make decisions. We exclude two questions with selected coverage: one on contraceptive use, only asked to those that use contraceptives, and one on how the husband's income is spent, only asked if the husband is working. This leaves us with the three questions on healthcare, major household purchases, and visits to family and relatives, detailed in the note to Figure 1.

We collapse these three decision-making variables to the country-level following these steps: first, we find the latest DHS survey of women for each country (the start year ranges from 2003 to 2023). Second, we keep only currently married women with non-missing data for all three of the decision-making variables. Third, we sum across each woman's answers to the three questions, giving us three measures, all from 0 to 3: the number of decisions made by the wife alone, the number made by the husband alone, and the number made jointly by both the wife and the husband. Fourth, we collapse each of these three measures to the country-level, applying the DHS-provided sampling weights.

Income and Population. Our measure of income is GDP per capita, PPP in constant 2021 international $. We take this measure, and our measure of total population, from the World Bank World Development Indicators (series NY.GDP.PCAP.PP.KD and SP.POP.TOTL). For both, we use the year(s) of the relevant DHS survey. For example, the latest available DHS survey for India is 2019/21. We then use India's average GDP per capita and population across the three years 2019 to 2021.

Joint Interventions in Pre-Registered Experiments. To measure the prevalence of joint household decision-making interventions among experiments, we start by downloading the full set of 10,055 pre-registered experiments (as of March 3, 2025) from the AEA RCT Registry. Next, we restrict to the 402 pre-registrations that mention any of "intra-household", "wife", "husband", or "married", including spelling variants, in any part of the registration.

We then use the OpenAI API to have GPT-4 code each of the pre-registrations as whether it involves joint household interventions or not. We provide the following definition as part of the prompt:

An example of a joint intervention is getting the wife and husband to discuss a possible health tool together. Whereas an intervention that just approaches the wife with a health tool would not be categorised as joint.

To make each decision, we provide GPT-4 with the title, abstract, and intervention details fields of each pre-registration. For entries with missing intervention details, or intervention details with 40 characters or less, we provide the experimental design field instead of the intervention details field.

To verify that GPT-4 classifies experiments accurately, we manually classified a random sample of 10% (40) of the 402 pre-registrations, blind to the classification of GPT-4. Encouragingly, the manual classification matches the GPT-4 classification in 39 of 40 cases, and in the one case of mismatch, we noted that the study was an "edge case" when manually classifying.

C Deviations from Pre-Registration

The pre-registration for the experiment can be found at https://www.socialscienceregistry.org/trials/1678. The one important deviation concerns our initial plan to study the effects of employment by randomizing which women get positions among those that enroll. Unfortunately, this was not feasible given low take-up of the program. There was oversubscription for the program in just one of our six weaving centers, giving us a sample of only 30 women for a second-stage experiment. Further, many women initially not assigned a job were eventually given one due to dropout.

Otherwise, one minor deviation is that we aimed to reach a sample size of 800 married couples, but in the end we were only able to reach 495 married couples for the six weaving centers to be opened.

D Discussion Script

We instructed enumerators to remain silent throughout the three minutes allocated to discussion. One of the two surveyors present (the one with the enrollment ticket) read out the following script to start the discussion:

> I will now give you two 3 minutes to talk. The intention is for you two to discuss the job opportunity but you can discuss something else or nothing at all if you prefer. I will remain here during your discussion but you should feel free to have an open and honest discussion. Everything you say will remain strictly confidential. Please only talk to one another at this time and do not ask me questions about the job.

The other surveyor was given the following instructions:

> pause again while other surveyor introduces 3 minute discussion period. Resume to give envelope with gift when other surveyor does so.

In addition, we coded the SurveyCTO form to require a minimum of 200 seconds (3 minutes + a 20-second buffer to allow the surveyor to read their part) to pass before the survey continues.

E Theory

E.1 Bargaining Frictions Model

We first present the ex ante model of household decision-making with bargaining frictions which maps to our experimental treatments. We outline the alternative model with veto rights following this.

Agents and Utility. A household consists of a husband, h, and a wife, w. If the wife does not take the job opportunity with Obeetee, the husband and wife receive utility U_h and U_w. If instead she does take it, the husband and wife receive utilities $U_h' = U_h + v_h$ and $U_w' = U_w + v_w$. The net utility gains to the husband and wife from enrolling, v_h and v_w, can be considered reduced form outcomes from a more complete model with any combination of the following features associated with the wife working:[49] (i) the wife earns extra income and this income is shared as per a consumption sharing rule, (ii) the wife has some disutility of effort, (iii) the husband and wife face, potentially different, stigma costs from breaking social norms, (iv) bargaining power shifts in the wife's favor, increasing her consumption share.[50] We assume v_h and v_w are known with certainty by both the husband and the wife, though this assumption is not essential for our model predictions.[51]

We assume a simple form of heterogeneity for the net utility parameters. They are independently and uniformly distributed: $v_h \sim U\left[\mu_h - \frac{\varphi}{2}, \mu_h + \frac{\varphi}{2}\right], v_w \sim U\left[\mu_w - \frac{\varphi}{2}, \mu_w + \frac{\varphi}{2}\right]$, with the same variance, but different means, such that $\mu_w > \mu_h$.[52] These assumptions match key features of our data on preferences:

[49] We detail a collective model with these features in Section E.3 below, showing how the net utility parameters can be microfounded using standard assumptions.

[50] For example, Atkin (2009) finds that Mexican women induced to work in manufacturing jobs have taller children and report stronger bargaining power, while Jensen (2012) and Heath and Mobarak (2015) find that labor market opportunities for women lead to delayed marriage and childbirth in Indian and Bangladesh respectively. However, the model we describe in Appendix E.3 includes features (i) to (iii), but does not include feature (iv).

[51] There are ways in which we could introduce imperfect information and still generate the same predictions. Consider the following model sketch: each individual knows their own preferences (v_i), but only knows the distribution from which v_j is drawn, i.e. $v_j \sim U\left[\mu_j - \frac{\varphi}{2}, \mu_j + \frac{\varphi}{2}\right]$. If an individual receives the job ticket and information and their spouse does not, they now withhold the information if (1) they are personally opposed ($v_i < 0$), and (2) revealing *could* lead to enrollment, given v_i, β, and the *range* of values v_j might take. Assume preferences become common knowledge after information is shared, e.g. through how the spouse reacts to the information, and bargaining proceeds as in our current model. This alternative version of the model yields the same key empirical predictions: $Info$ and *Discuss* both raise enrollment, but $WifeTicket$ has an ambiguous effect given the offsetting effects of husbands being less supportive of enrollment but also having higher bargaining power.

[52] As we show in Section E.3, the assumption of uniformly distributed net utility parameters is an implication of a collective

wives are on average more positive about women weaving than husbands, but there exist couples in which the husband is more positive than the wife. Building on this, we also assume that pro-job and anti-job types exist among both wives and husbands, i.e. $\mu_w - \frac{\varphi}{2} < 0 < \mu_h + \frac{\varphi}{2}$.

Household Decision-Making. We now consider three types of decision-making, from the most to the least efficient, and interpret our experimental treatments as shifting households between these different types.

Collective Model. Households may decide as in the collective model (Chiappori 1992; Browning and Chiappori 1998) – they bargain efficiently, and the wife enrolls whenever $\beta U_h' + (1-\beta) U_w' > \beta U_h + (1-\beta) U_w$, which is whenever

$$\beta v_h + (1-\beta) v_w > 0 \tag{1}$$

where β is the bargaining weight of the husband. We assume that $\beta > \frac{1}{2}$ given our specific evidence on the predictiveness of husbands' preferences of enrollment, as well as wide-ranging evidence in general for the high bargaining power of men in India. To simplify subsequent derivations, we additionally assume that:

$$\begin{aligned} \beta \left(\mu_h - \frac{\varphi}{2}\right) + (1-\beta) \left(\mu_w + \frac{\varphi}{2}\right) < 0 \\ \beta \left(\mu_h + \frac{\varphi}{2}\right) + (1-\beta) \left(\mu_w - \frac{\varphi}{2}\right) > 0 \end{aligned} \tag{2}$$

which is to say that the most anti-job husband would never reach a collective decision to enroll, while wives of the most pro-job husband would always enroll. This is a natural assumption, given that $\beta > \frac{1}{2}$, but not a necessary one for the empirical predictions we will emphasize.[53]

Bargaining Costs. Alternatively, households may fail to bargain efficiently if there is some cost to starting the negotiation process. Here we assume that there is a cost $c > 0$ of bargaining efficiently whenever there is disagreement, with $v_h v_w < 0$. Intuitively, this cost reflects the difficulty in striking up an uncomfortable discussion. After either spouse pays this cost, a collective decision, as above, will follow. Instead, when $v_h v_w \geq 0$,[54] the cost of bargaining is zero. In this case, the husband and wife want the same thing, making a conversation about the job a simple one.

model with uniformly distributed, additive, costs to each spouse of the wife working. These costs can be understood as stigma costs as in Field et al. (2021), but provided that they are additive, the particular interpretation is not important.

[53] These assumptions restrict the number of cases (i.e. parameter regions) we have to solve for, such that the expressions for enrollment probabilities do not depend on parameter regions.

[54] Recall the assumption that each spouse knows the net utility gain of the other with certainty.

If neither spouse pays the cost, we assume that the couple reverts to non-cooperative decision-making. In our case, spouses could not unilaterally enroll – both had to go to the weaving center on enrollment day if the wife was to enroll. Instead, with non-cooperative decision-making, either spouse can veto enrollment – by refusing to attend on enrollment day, for example. It follows that whenever either of v_h or v_w is negative, and neither spouse pays the bargaining cost, the woman will not enroll.

Summarizing, with bargaining costs, couples will enroll without needing to pay the cost whenever both spouses have positive net utility gains from enrollment. Couples will enroll after paying the cost whenever $v_w > c$ or $v_h > c$, and equation 1 is satisfied. Finally, we make the assumption that $-\frac{1-\beta}{\beta}\left(\mu_w - \frac{\varphi}{2}\right) \leq c < \mu_h + \frac{\varphi}{2}$. This ensures that the cost is low enough that at least some husbands and some wives pay it, and high enough to simplify subsequent derivations, without affecting the key empirical predictions.

Incomplete Information. If only one spouse is aware of the job opportunity, the household decision operates in two stages. In the first stage, the knowledgeable spouse decides whether to pass on the information to the other spouse or to withhold it. We assume that there is no direct cost of passing on the information. If the information is not passed on, there is no enrollment, because in our setting enrollment requires both spouses to go to the weaving center on enrollment day. If the information is passed on, we reach the second stage, in which spouses decide whether to pay the bargaining cost as above. If one spouse pays the cost, the couple bargain efficiently, as in the first decision-making type. If neither pay the cost, the couple bargains non-cooperatively.

Spouse i will withhold the information whenever (i) enrollment is net costly to them ($v_i < 0$), and (ii) enrollment would happen if information were revealed ($v_h, v_w > 0$, or $\beta v_h + (1-\beta) v_w > 0$ and $\max\{v_h, v_w\} > c$). Combining these conditions, husbands and wives respectively withhold job information when

$$\begin{aligned} -\frac{1-\beta}{\beta} v_w < v_h < 0, v_w > c \\ -\frac{\beta}{1-\beta} v_h < v_w < 0, v_h > c \end{aligned} \tag{3}$$

It follows that, ceteris paribus, a spouse is more likely to strategically withhold information when they have low bargaining power – this is precisely when the second stage decision is more likely to go against their interests.

Intuitively, this stylized model aims to capture the idea of an opportunity arising for one spouse that is considered in the household's best interest but requires the other spouse to make a sacrifice. In our context, the opportunity is a job for the wife, and a possible sacrifice would be the cost for the husband

from breaking a local social norm. More familiar to academics, one might instead imagine a promotion for one spouse that would require relocating to a new city. The other spouse may prefer to stay, but the household nevertheless reaches a joint decision to move. In each of these cases, a spouse has an incentive to withhold information.

Enrollment Rates and Treatment Effects. We map our main treatments to the decision-making types above, and derive effects on enrollment implied by the model.[55]

Effect of $WifeTicket$ There is no mechanism in our model for $WifeTicket$ to affect enrollment under $Info$ or *Discuss*, but it could have an effect under $NoInfo$. We begin by deriving the enrollment rate when the husband is given the ticket in $NoInfo$ ($Husband-NoInfo$ in Figure 4, the condition closest to the status quo). In this case, the enrollment rate is

$$E_h^{\text{NoInfo}} = \frac{\varphi\left(\mu_h + \frac{\varphi}{2}\right) + c\left(\mu_w - \frac{\varphi}{2}\right)}{\varphi^2}$$

The enrollment region is shown graphically as the shaded area in panel (a) of Figure B1. Intuitively, wives only enroll when the husband benefits on net from enrollment, and when the husband is willing to pay the bargaining cost for the cases where the wife would lose out from enrolling. Enrollment is increasing locally in both the average net utility gain for wives and husbands, and falling in the bargaining cost.

Next consider the parallel case: the wife is given the ticket in $NoInfo$ ($Wife-NoInfo$ in Figure 4):

$$E_w^{\text{NoInfo}} = \frac{\left(\mu_w + \frac{\varphi}{2}\right)\left(2\beta\left(\mu_h + \frac{\varphi}{2}\right) + (1-\beta)\left(\mu_w + \frac{\varphi}{2}\right)\right) - (1-\beta)c^2}{2\beta\varphi^2}$$

The enrollment region is the shaded area in panel (b) of Figure B1. Now enrollment only occurs when the wife benefits on net, and is willing to pay the bargaining cost for the cases where the husband would lose out.

The sign of $E_w^{\text{NoInfo}} - E_h^{\text{NoInfo}}$ is ambiguous, due to two offsetting forces. The more anti-job preferences of husbands push husbands toward withholding information more often than wives. In contrast, their high bargaining power ($\beta > \frac{1}{2}$) pushes them to share information more often, since their preferences are weighted more in any subsequent collective decision.

Effect of $Info$. In the experiment, we shut down the possibility of strategic withholding by removing the plausible deniability of ticket-receiving spouses. Through the lens of the model, this increases enrollment, which is now reflected by the shaded area in panel (c) of Figure B1 corresponding to $Info$ treatment in

[55] Proofs in Section E.4.

Figure 4.[56] Now couples have full information, and enrollment occurs whenever both spouses benefit ($v_h, v_w > 0$) or the couple collectively benefits ($\beta v_h + (1-\beta) v_w > 0$) and one spouse is willing to pay the bargaining cost. The impact on enrollment relative to when the wife gets the ticket is

$$\theta_w = \frac{\left(\mu_h + \frac{\varphi}{2} - c\right)\left(\frac{\varphi}{2} - \mu_w\right)}{\varphi^2} > 0$$

which is increasing in μ_h and decreasing in μ_w. In this sense, enforcing common knowledge after the wife gets the job ticket will not matter that much if women tend to be very pro-job and men tend to be very anti-job. This is because women rarely have the incentive to strategically withhold information. The impact on enrollment relative to when the husband gets the ticket is

$$\theta_h = \frac{1-\beta}{\beta} \frac{\left(\mu_w + \frac{\varphi}{2} - c\right)\left(c + \frac{1}{2}\left(\mu_w + \frac{\varphi}{2} - c\right)\right)}{\varphi^2} > 0$$

which is increasing in μ_w and falling in β. The more bargaining power husbands have, the less they need to strategically withhold information to prevent household decision-making going against their interests. The sign of $\theta_w - \theta_h$ is ambiguous given the two offsetting forces of job preferences and bargaining power. As a result, the model predicts that enforcing common knowledge will increase enrollment, but not necessarily that this effect will be larger when forcing husbands to share information as opposed to wives.

Effect of Discuss. In the experiment, we try to kickstart bargaining by having couples receive the information together and giving them several minutes to discuss it (*Discuss* treatment in Figure 4). Through the lens of the model, we consider this treatment as one that eliminates the bargaining cost, by giving individuals an excuse to bring up the job, for example. This would shift couples to making decisions as in the collective model (moving from panel (c) to panel (d) of Figure B1) and again increase enrollment, with couples now enrolling whenever $\beta v_h + (1-\beta) v_w > 0$. The increase in enrollment is

$$\theta^{\text{Discuss}} = \frac{\frac{1-\beta}{2\beta}\left(c^2 - \left(\frac{\varphi}{2} - \mu_w\right)^2\right) + c\left(\frac{\varphi}{2} - \mu_w\right)}{\varphi^2} > 0$$

The discussion effect is driven entirely by the couples that disagree – i.e. those with $v_h v_w < 0$, a dimension of heterogeneity we test for.

[56] As mentioned above, there is no mechanism in the model for enrollment rates to differ in $Info$ based on who got the ticket (though of course the *effect* of $Info$ on enrollment can vary by $WifeTicket$).

E.2 Veto Power Model

In the veto power model, a couple enrolls only if neither spouse vetoes it, but spouses incur a private utility cost from exercising a veto. Our treatments shift the costs by changing who feels entitled to veto; ticketed spouses always feel entitled to veto, while *Info* and *Discuss* make non-ticketed spouses feel more comfortable exercising a veto. Suppose, for simplicity, that being assigned the ticket or being in *Info* or *Discuss* eliminates the cost entirely (though of course it is plausible that *Discuss* would have reduced costs by a larger amount than *Info*). Then in the *Info* and *Discuss* groups, enrollment only occurs when *both* spouses prefer enrollment over non-enrollment ($v_h, v_w > 0$, panel (c), Figure B2). In *NoInfo* on the other hand, the couple enrolls whenever the ticketed spouse supports enrollment ($v_i > 0$, for $i = h$ or w) and the non-ticketed spouse is not too opposed to it ($v_j > -c_j$, where $c_j > 0$ is the cost of vetoing and $j = w$ or h) (panels (a) and (b), Figure B2). c_h and c_w need not be equal, and gender norms in our setting might suggest that $c_w > c_h$ (i.e. women incur a greater private utility cost from vetoing their husbands' preferred outcome than husbands do from vetoing their wives' preferred outcome). We draw the enrollment regions in Figure B2 with $c_w > c_h$. This, coupled with the fact that wives are more supportive of women weaving than husbands, means the veto power model has no prediction for whether enrollment will be higher when the husband or the wife receives the ticket in $NoInfo$, or for the interactions between $WifeTicket$ and $Info$/*Discuss*.

E.3 Micro-founding the net utility parameters

Adapting Field et al. (2021), we can write the maximization problem of a collective household as:

$$\begin{aligned}\max_{(h_i,c_i)_{i\in\{h,w\}}} \; & \beta\left[u_h\left(1-h_h,c_h\right)-\gamma_h 1\left(h_w>0\right)\right]+\\ & (1-\beta)\left[u_w\left(1-h_w,c_w\right)-\gamma_w 1\left(h_w>0\right)\right]\\ & \text{subject to } c_w+c_h \leq \Sigma_{i=w,h}\left[y_i+w_i h_i\right]\\ & 0 \leq h_i \leq 1 \text{ for } i \in \{h,w\}\end{aligned}$$

To solve for the net utility parameters v_h and v_w, we solve this maximization problem twice: once for the case in which the wife doesn't work ($h_w = 0$) and again for the case in which she does ($h_w > 0$). The optimum in the former case yields utility U_h for the husband and U_w for the wife, while in the latter case the resulting utility is U_h' and U_w'. The difference in utility from having the spouse work is then $U_h' - U_h = v_h$ for the husband and $U_w' - U_w = v_w$ for the wife. To make this concrete, we now solve for U_i and U_i' for the case where $u_i\left(1-h_i,c_i\right) = \ln\left(c_i\right) + \phi \ln\left(1-h_i\right)$.

Case I: Wife Does Not Work. Here, we take the wife's labor supply to be constrained at $h_w = 0$. For simplicity, call combined unearned income $y = y_h + y_w$. The household's problem now collapses to:

$$\begin{aligned} \max_{h_h, c_h, c_w} \beta\left(\ln c_h + \phi \ln\left(1 - h_h\right)\right) + \left(1 - \beta\right)\left(\ln c_w\right) \\ \text{subject to } c_w + c_h \leq y + w_h h_h \\ 0 \leq h_h \leq 1 \end{aligned}$$

The Lagrangian is:

$$\mathscr{L} = \beta \ln\left(c_h\right) + \beta\phi \ln\left(1 - h_h\right) + \left(1 - \beta\right)\ln\left(c_w\right) + \lambda\left(w_h h_h + y - c_h - c_w\right)$$

The first order conditions are:

$$\frac{\beta}{\lambda} = c_h,\ \frac{1-\beta}{\lambda} = c_w,\ \frac{\beta\phi}{\lambda w_h} = 1 - h_h$$

Using the budget constraint to solve for optimal consumption and the husband's labor supply, we get:

$$c_w^* = \frac{1-\beta}{1+\beta\phi}\left(w_h + y\right),\ c_h^* = \frac{\beta}{1+\beta\phi}\left(w_h + y\right),\ h_h^* = \frac{w_h - y\beta\phi}{w_h\left(1+\beta\phi\right)}$$

Given that $w_h, y, \beta, \phi > 0$, labor supply is at an interior solution provided that unearned income is sufficiently small relative to wages ($y\beta\phi \leq w_h$). This is clearly the relevant parameter region to consider given that 82% of husbands in our experimental sample had worked for income in the three months prior to the baseline survey. Equilibrium utils are then:

$$U_h = \left(1+\phi\right)\ln\left(\frac{\beta}{1+\beta\phi}\left(w^h + y\right)\right) + \phi\ln\left(\frac{\phi}{w^h}\right),\ U_w = \ln\left(\frac{1-\beta}{1+\beta\phi}\left(w^h + y\right)\right)$$

Case II: Wife Works. Assuming that the wife works, the household's problem becomes:

$$\begin{aligned} \max_{\left(h_i, c_i\right)_{i\in\{h,w\}}} \beta\left[\ln c_h + \phi\ln\left(1 - h_h\right) - \gamma_h\right] + \\ \left(1-\beta\right)\left[\ln c_w + \phi\ln\left(1 - h_w\right) - \gamma_w\right] \\ \text{subject to } c_w + c_h \leq y + \Sigma_{i=w,h} w_i h_i \\ 0 \leq h_i \leq 1 \text{ for } i \in \{h, w\} \end{aligned}$$

Since the γs do not affect the optimization problem, the Lagrangian becomes:

$$\mathscr{L} = \beta \ln(c_h) + \beta\phi \ln(1-h_h) + (1-\beta)\ln(c_w) + (1-\beta)\phi \ln(1-h_w) + \lambda\left(w_h h_h + w_w h_w + y - c_h - c_w\right)$$

The first order conditions are now:

$$\frac{\beta}{\lambda} = c_h,\ \frac{1-\beta}{\lambda} = c_w$$

$$\frac{\beta\phi}{\lambda w_h} = 1-h_h,\ \frac{(1-\beta)\phi}{\lambda w_w} = 1-h_w$$

Using the budget constraint, and with several steps of rearranging, we can solve for optimal labor supply as:

$$h_h' = \frac{1+\phi(1-\beta)}{1+\phi} - \frac{\phi\beta}{1+\phi}\frac{w_w+y}{w_h}$$

$$h_w' = \frac{1+\phi+\beta(1-\beta)\phi^2}{(1+\phi)(1+\phi(1-\beta))} - \frac{\phi(1-\beta)}{1+\phi}\frac{w_h+y}{w_w}$$

We again assume that parameters are such that these expressions satisfy the labor supply constraints that $0 \leq h_i \leq 1$.[57] Substituting into the budget constraint and simplifying, we can solve for total earned income as:

$$y_e' = w_h h_h' + w_w h_w' = \frac{w_h + w_w - \phi y}{1+\phi}$$

It follows that $c_h' = \beta y_e'$ and $c_w' = (1-\beta) y_e'$. With optimal consumption and labor supply in hand, we solve for equilibrium utility:

$$U_h' = (1+\phi)\ln\left(\frac{\beta}{1+\phi}(w_h + w_w + y)\right) + \phi\ln\left(\frac{\phi}{w_h}\right) - \gamma_h$$

$$U_w' = (1+\phi)\ln\left(\left(\frac{1-\beta}{1+\phi}\right)(w_h + w_w + y)\right) + \phi\ln\left(\frac{\phi}{w_w}\right) - \gamma_w$$

The Collective Decision Of Whether To Work. The household collectively decides that the wife should work whenever:

$$\beta U_h' + (1-\beta)U_w' > \beta U_h + (1-\beta)U_w$$

$$\Rightarrow \beta\left(U_h' - U_h\right) + (1-\beta)\left(U_w' - U_w\right) > 0$$

$$\Rightarrow \beta v_h + (1-\beta)v_w > 0$$

[57] Specifically, $w_h, y, \beta, \phi > 0$ and $\beta < 1$ ensure that $h_h', h_w' < 1$, while for $h_h' > 0$ we require that $\frac{w_w+y}{w_h} < \frac{1+\phi(1-\beta)}{\phi\beta}$, and for $h_w' > 0$ we require that $\frac{w_h+y}{w_w} < \frac{1+\phi+\beta(1-\beta)\phi^2}{\phi(1-\beta)(1+\phi(1-\beta))}$. These last two conditions are more likely to hold when unearned income is small and gender-specific wages not too different.

Our net utility parameters are:

$$v_h = U_h' - U_h = (1+\phi)\ln\left(\frac{1+\beta\phi}{1+\phi}\frac{(w_h+w_w+y)}{(w_h+y)}\right) - \gamma_h$$

$$v_w = U_w' - U_w = (1+\phi)\ln\left(\left(\frac{1-\beta}{1+\phi}\right)(w_h+w_w+y)\right) - \ln\left(\frac{1-\beta}{1+\beta\phi}(w_h+y)\right) + \phi\ln\left(\frac{\phi}{w_w}\right) - \gamma_w$$

In our model in Appendix E.1 we assume that the net utility parameters are independently and uniformly distributed. Given the expressions above, this assumption can be microfounded by assuming (i) no couple-level heterogeneity in gender-specific wages (w_i), bargaining power (β), labor-leisure preferences (ϕ), and unearned income (y), and (ii) independently and uniformly distributed norm costs (γ_i).

E.4 Solving for enrollment rates

The enrollment rate when husbands receive the ticket and can withhold information is the fraction of couples with $v_h, v_w > 0$ plus the fraction of couples with $v_h > c, v_w < 0$:

$$\begin{aligned} P(\text{enroll} \mid \text{husband-ticket,bargaining-cost}) &= \frac{\left(\mu_w + \frac{\varphi}{2}\right)\left(\mu_h + \frac{\varphi}{2}\right)}{\varphi^2} + \frac{\left(\frac{\varphi}{2} - \mu_w\right)\left(\mu_h + \frac{\varphi}{2} - c\right)}{\varphi^2} \\ &= \frac{\varphi\left(\mu_h + \frac{\varphi}{2}\right) + c\left(\mu_w - \frac{\varphi}{2}\right)}{\varphi^2} \end{aligned}$$

The equivalent rate for when wives receive the ticket is the fraction of couples with $v_h, v_w > 0$ plus the fraction of couples with $v_w > c, \beta v_h + (1-\beta) v_w > 0$. Equivalently, the enrollment rate is the fraction of couples with $v_w > 0, \beta v_h + (1-\beta) v_w > 0$ less the fraction of couples with $v_w < c, \beta v_h + (1-\beta) v_w > 0$:

$$\begin{aligned} P(\text{enroll} \mid \text{wife-ticket,bargaining-cost}) &= \frac{\left(\mu_w + \frac{\varphi}{2}\right)\left(\mu_h + \frac{\varphi}{2}\right)}{\varphi^2} + \frac{\frac{1-\beta}{2\beta}\left(\mu_w + \frac{\varphi}{2}\right)^2}{\varphi^2} - \frac{\frac{1-\beta}{2\beta}c^2}{\varphi^2} \\ &= \frac{\left(\mu_w + \frac{\varphi}{2}\right)\left(2\beta\left(\mu_h + \frac{\varphi}{2}\right) + (1-\beta)\left(\mu_w + \frac{\varphi}{2}\right)\right) - (1-\beta)c^2}{2\beta\varphi^2} \end{aligned}$$

The effect of enforcing common knowledge on enrollment when the wife receives the ticket is equal to the fraction of couples with $v_h > c, v_w < 0$. This is clearly $\theta_w = \frac{\left(\mu_h + \frac{\varphi}{2} - c\right)\left(\frac{\varphi}{2} - \mu_w\right)}{\varphi^2}$,and is positive given the assumptions that $\mu_w - \frac{\varphi}{2} < 0$ and $c < \mu_h + \frac{\varphi}{2}$.

The effect of enforcing common knowledge on enrollment when the husband receives the ticket is equal to the fraction of couples with $v_w > c, \beta v_h + (1-\beta) v_w > 0$. This is the sum of the area of a

rectangle and a triangle (see panel (c), Figure B1) divided by φ^2:

$$\begin{aligned}\theta_h &= \frac{\frac{1-\beta}{\beta} c \left(\mu_w + \frac{\varphi}{2} - c\right)}{\varphi^2} + \frac{\frac{1-\beta}{2\beta} \left(\mu_w + \frac{\varphi}{2} - c\right)^2}{\varphi^2} \\ &= \frac{1-\beta}{\beta} \frac{\left(\mu_w + \frac{\varphi}{2} - c\right) \left(c + \frac{1}{2} \left(\mu_w + \frac{\varphi}{2} - c\right)\right)}{\varphi^2}\end{aligned}$$

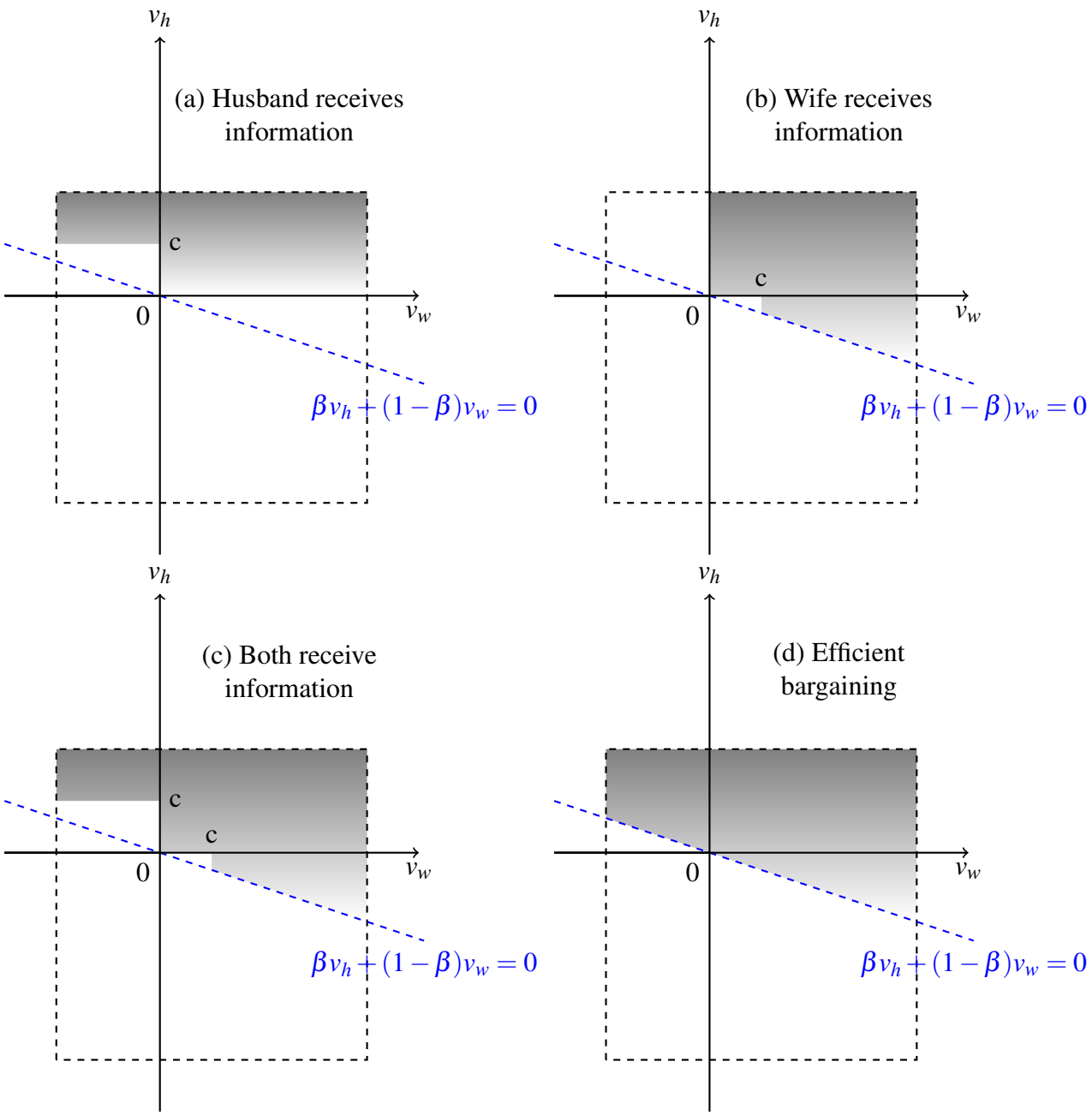
which is positive given that $c < \mu_h + \frac{\varphi}{2} \Rightarrow c < \mu_w + \frac{\varphi}{2}$, since $\mu_w > \mu_h$.

The effect of enforcing bargaining on enrollment is equal to the fraction of couples with $v_w < c, v_h < 0, \beta v_h + (1-\beta) v_w > 0$ plus the fraction of couples with $v_w < 0, v_h < c, \beta v_h + (1-\beta) v_w > 0$:

$$\begin{aligned}\theta^{\text{Discuss}} &= \frac{\frac{1-\beta}{2\beta} c^2 + \frac{1-\beta}{2\beta} \left(\frac{\varphi}{2} - \mu_w\right)^2 + \left(c - \frac{1-\beta}{\beta} \left(\frac{\varphi}{2} - \mu_w\right)\right) \left(\frac{\varphi}{2} - \mu_w\right)}{\varphi^2} \\ &= \frac{\frac{1-\beta}{2\beta} \left(c^2 - \left(\frac{\varphi}{2} - \mu_w\right)^2\right) + c \left(\frac{\varphi}{2} - \mu_w\right)}{\varphi^2}\end{aligned}$$

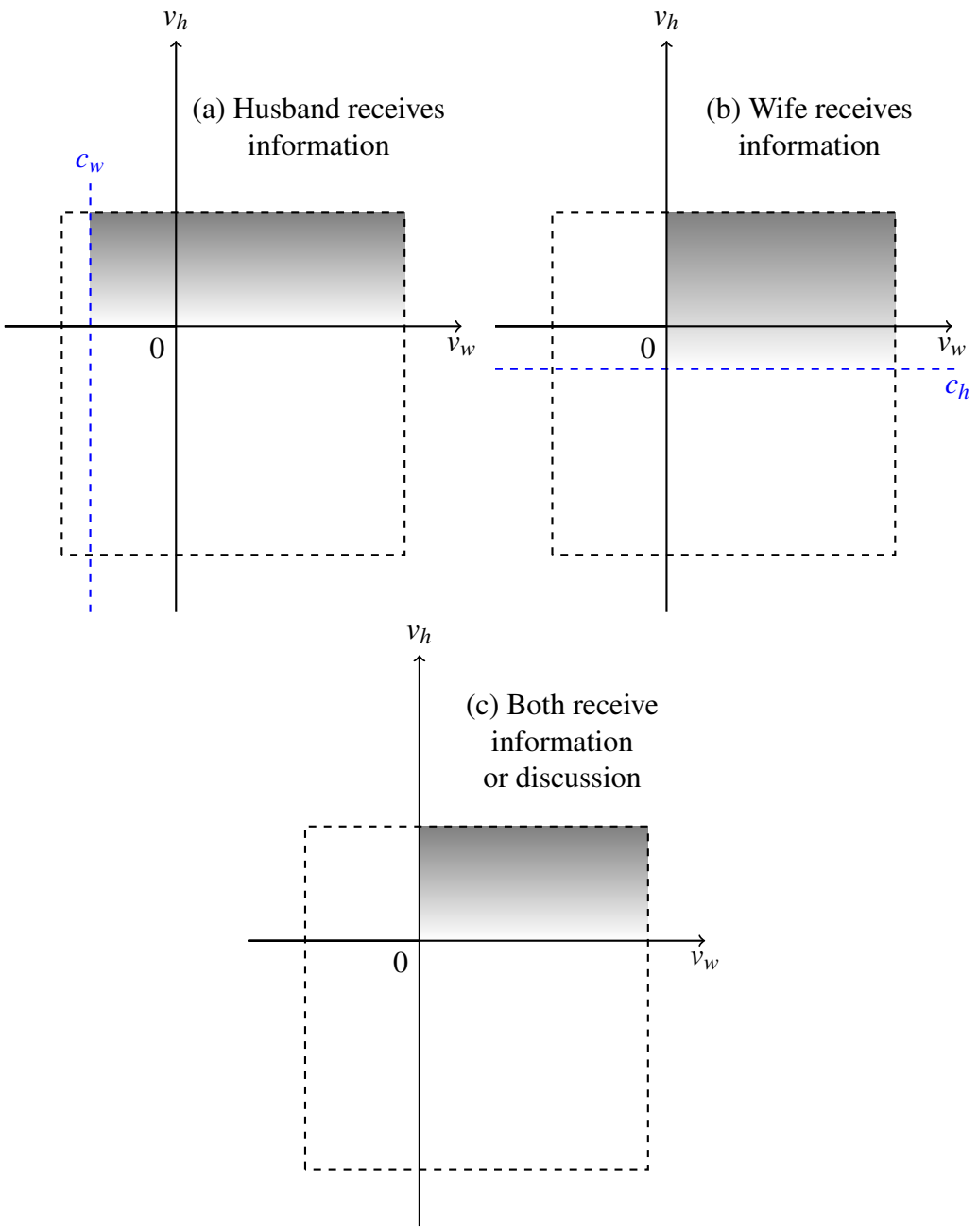
which is positive since $c \geq -\frac{1-\beta}{\beta} \left(\mu_w - \frac{\varphi}{2}\right)$.

Figure B1: Enrollment Regions Under Bargaining Frictions Model



Notes: The figure is drawn for $v_h \sim U[-4,2], v_w \sim U[-2,4], \beta = \frac{3}{4}, c = 1$. The black dashed line box shows the region in which the preferences of husbands and wives are independently and uniformly distributed. The blue dashed line is the line of indifference for collective households. The shaded regions indicate the preference regions in which wives enroll in the job under different assumptions about household decision-making: (a) husband receives information and can withhold, (b) wife receives information and can withhold, (c) both receive information but there remains a cost of bargaining, (d) both receive information and efficient bargaining.

Figure B2: Enrollment Regions Under Veto Power Model



Notes: The figure is drawn for $v_h \sim U[-4,2], v_w \sim U[-2,4], c_h = -0.75, c_w = -1.5$. The black dashed line box shows the region in which the preferences of husbands and wives are independently and uniformly distributed. The blue dashed lines represent the level of utility below which the uninformed spouse will veto enrollment. The shaded regions indicate the preference regions in which wives enroll in the job under different assumptions about household decision-making: (a) husband receives information, (b) wife receives information, and (c) both receive information with or without discussion. Unlike Figure B1, we now assume couples make decisions through vetoes, there is no cost to either spouse from vetoing in (c), there is a cost from vetoing for the uninformed spouse in (a) and (b), and there is full information diffusion.

F Expert Surveys: Additional Details

F.1 Researchers

Sampling. We drew up a list of experts that included (i) authors of papers on intra-household economics[58] published in the last 10 years in any of the top-5 economics journals, *American Economic Journal: Applied Economics,* or *Journal of Development Economics*; (ii) presenters of intra-household research at NEUDC in the last 10 years;[59] (iii) authors of intra-household papers we cited in our October 2019 draft; and (iv) authors of intra-household papers in the syllabi for the PhD development economics courses at MIT and Stanford. From this list, we dropped (i) inactive researchers, including the retired, emeritus faculty, and those who had shifted to industry at least a few years ago; (ii) those without publicly available email addresses; and (iii) two academics we were certain already knew the results. This resulted in a list of 361 researchers.

Confidence. Following DellaVigna and Pope (2018), as a proxy for an expert's confidence we also asked: *"How many of your seven predictions do you expect are 5 percentage points or less from the actual enrollment rate in each treatment group?"*

Robustness. The prediction patterns we report in the main text are similar if we include only full professors, or if we exclude the seven experts that had seen the paper but couldn't remember the results (Figure A4).

F.2 Locals

Due to constraints on the length of the baseline surveys and the smaller number of female weaving centers, we assigned our prediction questions to all female weavers and a randomly selected half of male weavers in the Kala and McKelway (2025a) baseline survey. Half of the respondents were randomly assigned to rank the three intervention options described in the text, where the surveyor would be present for the discussion. The interventions were described as follows:

[58] Papers with (a) "intra-household," "intrahousehold," "within the household," or something similar in the title or abstract, or (b) at least one "O" category JEL code and at least one of D13, J12, and J16.

[59] Some of the older NEUDC websites in the 10-year window no longer exist, so we did not include experts from those conferences.

As you may know, Obeetee has arranged a program which provides women paid training and employment in carpet weaving. Imagine Obeetee were recruiting for this program in a village nearby. Women and their husbands would have to go together on a certain day to enroll. In the week before, a female surveyor would go door-to-door to inform married couples about the opportunity.

There are three options for how she informs them:

Option 1: inform husbands alone

Option 2: inform wives alone

Option 3: inform the two together and encourage them to discuss the opportunity together while she's meeting them

For the other half of respondents, the third intervention option instead involved the surveyor informing the two together and encouraging them to discuss the opportunity together once she had left, reading: *Option 3: inform the two together and encourage them to discuss the opportunity together once she's left*. We additionally randomized the decision topic among male weavers; half got the scenario about the Obeetee job, while the other half were asked to rank the recruitment strategies in the context of getting parents to enroll their daughters in a tutoring program for girls. In this case, the first paragraph of the prompt read: *Imagine there were an NGO named "girls run the world" which offered a tutoring program for girls of primary school age. The tutoring would be subsidised but not free. Imagine the NGO were recruiting for this program in a village nearby. Mothers and fathers would have to go together on a certain day to enroll. In the week before, a female surveyor would go door-to-door to inform parents about the opportunity.* The rest of the prompt was identical, except the options referred to fathers and mothers instead of husbands and wives. This randomization was only done among male weavers, with all females getting the Obeetee scenario, because of the smaller number of female weavers.

All randomizations were stratified by loom center. The questions had to be cut altogether about 75% of the way through the baseline surveys due to issues with survey length. In all, we solicited predictions from 695 local experts, 33% of whom are female weavers and the rest male weavers. 241 locals made predictions for the scenario corresponding to our experiment (Obeetee job, surveyor present), 49% of whom are female.