Coupling Labor Supply Decisions: An Experiment in India

Matt Lowe* Madeline McKelway†

February 2022

Abstract

We study frictions in household decision-making about 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. Our motivating model and most experts predict information and discussion should raise enrollment. In reality, information had no effect, and discussion reduced enrollment by 40-50%. Negative effects were driven by couples in which the non-ticketed spouse was less supportive of female weavers, consistent with a model in which involving both spouses gives each a veto. ***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, 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 received IRB approval from UBC (#H21-00189) and Stanford (#60413).

1 Introduction

Strategic motivations of individuals may keep household decisions from being made jointly. For example, an individual may withhold information from their spouse to prevent a household decision that goes against their personal interest. This can explain why some husbands in the Philippines hide money from their wives (Ashraf 2009), and why some wives in Zambia hide contraceptive use from their husbands (Ashraf et al. 2014). Bargaining costs are another example of the conflict between individual incentives and joint decision-making (Coase 1960; Riedl 1995; Anderlini and Felli 2001; Coase 2005). A spouse may avoid joint decision-making to avoid these costs, particularly if joint decision-making would lead to an outcome that the individual personally opposes. Both information withholding and bargaining costs can lead to decisions that are inefficient from the household's perspective.

We study these frictions in the context of female labor supply in India. Many women in India cannot make unilateral decisions to work – they must have the approval of their households to supply labor (McKelway 2021a,b). By precluding joint decision-making, information withholding and bargaining costs could constrain women's employment. Further, these frictions may be especially common in decisions about female labor supply in India. Due to constraints on women's physical mobility in the country (Jayachandran 2015), husbands may be the ones who learn about employment opportunities for women. However, husbands are less supportive of women working than women themselves are (Bernhardt et al. 2018; Field et al. 2021). Husbands may therefore withhold information about employment opportunities to keep their wives from working. Regarding bargaining costs, marriages in India are often arranged, which means spouses in young couples may not know each other well. This could make bargaining especially costly in such couples.

We study the decisions of 495 young couples about whether the wife would take a job opportunity in rural Uttar Pradesh. The opportunity was offered by our partner firm Obeetee, one of India's largest carpet manufacturers. The opportunity entailed four months of paid training in carpet weaving followed by employment. To accommodate social norms, the weaving took place in all-women centers located in the villages where participants lived.

We motivate our experiment with four facts. First, women's employment is low. 82% of husbands had worked for income in the three months prior to our baseline survey, but only 13% of their wives had. Second, preferences for women's employment differ systematically between husbands and wives. Wives are significantly more supportive of women in their households working full-time outside the home, both as weavers and as workers in other occupations. Third, spousal preferences are misaligned. Over 50% of couples disagree with each other when answering a question on the perceived appropriateness of women working as weavers. Fourth, the preferences of husbands are significantly more predictive of program

enrollment than those of wives, consistent with husbands enjoying high bargaining power.

We randomized recruitment strategies that shut down the roles of information withholding and bargaining costs. We first printed enrollment tickets for a randomly-chosen 90% of the couples. Women could not enroll in the program without their unique job ticket. We used a two-by-three design among the 495 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 a ticket. As a result, ticketed spouses in $NoInfo$ had plausible deniability – they could hide the existence of the job ticket from their spouse without getting found out.

Women enrolled in the program by going to their local women's weaving center on enrollment day. They were required to present their enrollment ticket and to come with their husbands. The requirement that both spouses attend ensured that enrollment, like typical labor supply decisions in India, was a joint household decision.

We write a simple model of household decision-making about enrollment to describe how our interventions might affect job take-up. Our model adds asymmetric information and a bargaining cost to the efficient, collective model (Chiappori 1992; Browning and Chiappori 1998). In particular, 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 (Ashraf 2009; Ashraf et al. 2014). 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. This 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. Even with full information, a second friction may prevent bargaining and thus enrollment: bargaining costs (Coase 1960; Riedl 1995; Anderlini and Felli 2001; Coase 2005). 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. This model has two key predictions. First, $Info$ should raise enrollment relative to $NoInfo$ since it makes information symmetric. 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.[1]

We solicited predictions from intra-household researchers, focusing largely on individuals who had done intra-household research in developing settings, and almost all 70 experts we surveyed agree with the model's two core predictions. 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. Next, 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. Before we ran the experiment, our priors were similar to the average expert's.

We report two main findings, each contrary to our model, the predictions of experts, and our priors. First, we find no evidence that information withholding keeps enrollment low. The effect of $Info$ relative to $NoInfo$ is negative three percentage points and statistically insignificant, and we reject the mean expert prediction of a 5.5 percentage point increase ($p = 0.02$). The effects of information are similar whether the husband or the wife received the ticket, suggesting that neither husbands nor wives withheld information to prevent enrollment. We also do not see positive effects when we focus on the subgroup of couples where theory predicts the largest positive effects – those couples in which the ticket recipient was less supportive of women weaving at baseline than the non-ticketed spouse.

Consistent with these results, we find direct evidence of information diffusion. In the $NoInfo$ condition, over 70% of non-ticketed spouses reported knowledge of the job ticket on an endline survey done in the week following enrollment, suggesting that over 70% of those who received tickets shared the information with their spouses. The $Info$ treatment increased ticket awareness by around ten percentage points ($p = 0.04$), but given the null effects on enrollment, this increase likely reflects a set of couples that would not have enrolled even if the information had been shared. Spouses in these couples likely did not share the information because there was no need to.

Our second finding is that discussion lowered enrollment – by up to nine percentage points, or 50%, relative to $NoInfo$ ($p < 0.01$) and by up to six percentage points, or 40%, relative to $Info$ ($p < 0.1$). We reject the mean expert prediction of *Discuss* relative to $NoInfo$ ($p < 0.001$) and relative to $Info$ ($p < 0.001$). Focusing on the heterogeneity in expert opinion, we can reject every single expert's *Discuss*

[1]The model makes an ambiguous prediction about the effect of whether the wife or the husband gets the job ticket in $NoInfo$. Two forces push in opposite directions: the fact that husbands are less supportive of the job makes them more likely to withhold information, but the fact that husbands have more bargaining power over the ultimate decision (once information is revealed) makes them less likely to withhold information. Nevertheless, experts predict enrollment to be 3.2 percentage points higher when the wife receives the job ticket in $NoInfo$. We cannot reject this prediction at conventional levels, though we also cannot reject the null of no effect.

versus $NoInfo$ prediction at the 5% level, and 83% of the predictions relative to $Info$. Experts are not just wrong on average – they are wrong with near-unanimity. The theory is also wrong, and effects are still negative when we focus on a subgroup for which the model predicts the effects should be the most positive – the couples that disagreed about the appropriateness of women weaving at baseline. Further, we find negative effects of discussion in all but one of the six villages where we ran the experiment sequentially. Though the within-village effects are generally not statistically significant given the small samples, this pattern suggests that the surprising negative effect of discussion is not just a "fluke" result and may replicate.

Why are the theory, the experts, and our priors wrong? We find support for an alternative model that departs even further from the efficient model than our initial model. The alternative model rests on two ideas: (i) spouses do not bargain and only enroll if neither vetoes enrollment, and (ii) informing only one spouse, as in $NoInfo$, weakens the other's veto, while interventions that nudge households toward decision-making, like our $Info$ and *Discuss* treatments, give both spouses full veto power. This model rationalizes the lack of information withholding and predicts negative effects of discussion. It also implies a pattern of heterogeneity that is the near opposite of what our original model predicts: interventions toward joint decision-making should reduce enrollment among the couples in which the non-ticketed spouse is less supportive of enrollment than the ticketed one. The data confirm this prediction: having any joint decision-making intervention ($Info$ or *Discuss*) reduced enrollment by up to 16 percentage points among these couples, and had no effect for the rest. Though it is more common for a husband to be less supportive than for a wife to be, it is disagreement in spouses' preferences rather than the preferences themselves that mediate the effects – we see no heterogeneity by husbands' or wives' support individually.

Our ultimate support for one model of household decision-making over another is not merely academic. In a world where household bargaining is impeded by frictions, nudges toward joint decision-making naturally raise female employment. In contrast, in a world with spouses exercising vetoes, nudges toward joint decision-making lead to status quo bias, reducing female employment. The policy implication of the two worldviews is diametrically opposed, making the distinction between the two especially important.

We find little support for seven other explanations for the negative effect of discussion. We address the four most plausible here and detail all seven in Section 7. One possible explanation is that couples found our discussion treatment awkward, and subsequently avoided our research team and the job we suggested. But individuals in the *Discuss* group were no less likely to take our endline survey. A second possible explanation is that the *Discuss* treatment reduced enrollment by causing arguments. Yet surveyors did not report any of the three-minute discussions to be argumentative, and in the endline survey, individuals

in the *Discuss* treatment group did not report more disagreement with their spouse about the opportunity than those in $NoInfo$ or $Info$. A third possibility is that surveyors' presence during the three-minute discussions reminded couples of the norm that women not work outside of their homes. However, effects are no more negative in villages with lower female employment rates at baseline, where the norm against women's work is likely to be stronger. Fourth, though *Discuss* couples talked about the job opportunity more often, we do not find that deeper deliberation per se reduced enrollment. In particular, for couples not in *Discuss*, having more days between ticket receipt and enrollment – and thus having more time to make the enrollment decision – is not associated with lower enrollment.

In the final part of the paper, we judge the judges: while experts mispredicted effects, are they at least well-calibrated in their confidence? Experts expected that 2.7 of their 7 predicted enrollment rates would be within 5 percentage points of the truth, but even this low confidence is overconfidence – only 1.4 predictions are actually within 5 percentage points of the truth on average. Confidence itself is uncorrelated with accuracy (as with the experts in DellaVigna and Pope (2018)), and those with more Google Scholar citations give *less* accurate predictions.

While experts are inaccurate and overconfident in our setting, would such misjudgment generalize to predictions about household intervention in general? We cannot answer this question well, but we examine an effect from Dean and Jayachandran (2019) similar to our *Discuss* effect. The authors evaluate how a surveyor-guided conversation between female teachers and their families affected teachers' employment at their schools one year later. Like us, they find a negative effect, and though it is insignificant in their case, using their confidence interval we can again reject our mean expert prediction at the 5% level. Misjudgement of the effects of household intervention may then be a more general phenomenon.

Our paper provides new evidence on the (unintended) effects of intervention into household decision-making. While many papers document inefficient household decision-making in poor countries (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; Buchmann et al. 2021; Conlon et al. 2021), we know less about whether policy can intervene to correct these inefficiencies. In a seminal lab-in-the-field experiment, Ashraf (2009) used similar treatments as us to study the role of spousal observability and communication for saving and consumption decisions. In field experiments, Ashraf et al. (2014) study the effects of individual versus joint decision-making about contraceptive use, while Ashraf et al. (2020a) randomize whether husbands or wives are given information about maternal health costs to study information diffusion in the household and fertility. Ashraf et al. (2020b) evalute a more intensive intervention that aimed to increase household efficiency by training girls to negotiate. Our study is most similar to Ashraf et al. (2014) and Ashraf et al. (2020a), in that we study light-touch interventions in the field. We build on their work in several ways. First, we explore a different domain

of decision-making: female labor supply. Second, we evaluate an intervention designed to overcome bargaining costs, while their interventions are focused on information withholding and information diffusion. Third, our exploration of information diffusion involves verifiable information (the job ticket itself), shutting down the issue of credible communication that is central in Ashraf et al. (2020a). Fourth, we incorporate expert predictions.

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; McKelway 2021a,b), showing that efforts to encourage joint decision-making can lower labor supply in the presence of disagreement. Finally, our work introduces expert 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).

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 4500 ($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 also common: at the time of marriage, women leave their native villages and become a part of a family they do not know well in a new village. Married women are typically confined to their homes and responsible for all 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.[2] This norm is strongly ingrained in cultural mindsets, with 87% of men and 84% of women believing that husbands should earn more income than their wives.

[2]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% in 2016, Figure A1).

2.2 Partner Firm and Women's Job Opportunity

We partnered with Obeetee, one of India's largest carpet manufacturing and exporting firms. 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.[3]

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 (Figure A2) 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. First, a larger pool of labor allows Obeetee to more easily take advantage of periods of high export demand for carpets. Second, female labor helps to offset local labor shortages driven by increasing rural-urban migration of male weavers. Third, the relation to female empowerment links naturally to Obeetee's long history of promoting development in the region through corporate social responsibility projects.

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.[4] 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 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

[3]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.

[4]Even during the training period, the daily wage offered to women is only slightly below that which a skilled male weaver would earn. 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.

is low status.[5] Nevertheless, enrollment and retention rates are low. In our experiment, the enrollment rate is 15%, and the average dropout rate conditional on enrollment is 49%.

2.3 Gender Differences in 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,[6] 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 and theoretical framework.

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).[7] 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 A3).[8] Our staff partners at Obeetee also frequently mentioned opposition from husbands as a key constraint to 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),

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

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

[7] The divergence is smaller for teaching, at 7% of the husband mean, consistent with it being a less male-dominated profession. The husband mean is in each case notably lower, making the perceived appropriateness of women working on average lower than that for men.

[8] The most common reason provided for not enrolling is that there is no-one else to do household chores. It is likely that this reason also reflects 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.

Figure 1), with this disagreement concentrated among couples where the wife responds more positively than the husband. For example, in 18% of couples the wife considers it "completely appropriate" for women in her household to work outside the home as a weaver, while the husband considers it "inappropriate."

Third, the preferences of husbands are more predictive of enrollment than the preferences of wives. Job enrollment is higher when husbands and wives deem women weaving to be appropriate (panel (b), Figure 1), 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) husbands' preferences are more predictive of enrollment than wives' preferences. These facts motivate our experiment design and model.

3 Experiment Design

3.1 Timeline

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 2).

Census. 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.[9] 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,

[9]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.

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. After each census was completed, we randomly assigned treatment at the couple-level, stratifying by hamlet and an indicator for Other Backwards Castes (OBC).[10] Given that we randomized center-by-center, for our pooled regressions our strata fixed effects are the full interaction between center-by-hamlet dummies and OBC status.

Baseline and Intervention. Following the randomization we implemented a baseline survey containing our experimental intervention. This period lasted for six to eight days 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. 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, 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.

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.

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. Those that enrolled were also required to present unique enrollment tickets given to one member of the pair during the baseline survey. The requirement for women to attend with their pairs 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

[10]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.

the household. Accordingly, our theoretical framework below hinges on the idea that while a spouse can veto enrollment, they cannot unilaterally decide to enroll. Given our unexpected results, we consider an alternative interpretation in Section 7.2 in which the extent to which decision-making is joint depends on the treatment.

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 A4). No woman could enroll without her unique ticket. Crucially, we only printed tickets for a randomly chosen 90% of eligible women. 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 treatments (Figure 3).[11] 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

[11] 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 3.

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 without ever having to argue their case.

Information. In another one-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 ($FullInfo$). While non-ticketed spouses in $FullInfo$ have greater endline knowledge of job details than those in $TicketInfo$, the effects on job desirability and enrollment are not statistically significant and point in opposite directions (Table A1).[12] 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.[13] In contrast to $NoInfo$ couples, both spouses in $Info$ couples knew that the wife could enroll in the job and no information could be withheld. That said, spouses in $Info$ were told about the opportunity separately and the intervention did not affect how they interacted with one another.

Discussion. The final one-third of couples were assigned to the *Discuss* group. Here, the non-ticketed

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

[13] For completeness, we also report the expert predictions and core enrollment results without any pooling in the Appendix.

spouse was present while the ticket and job details were given to the ticketed spouse. Surveyors then paused the survey for a full three minutes and encouraged the couple to discuss the opportunity together. Two surveyors were present during the discussion but remained silent and did not provide any additional information about the job. At the end of the discussion, the surveyor handed the enrollment ticket to the ticketed spouse. According to surveyor evaluations, 80% of *Discuss* couples discussed the opportunity during this time, with the other couples remaining silent or discussing something else. The goal of the discussion treatment was to relax possible barriers 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.[14]

4 Theory

We consider a simple model of household decision-making which maps to our experimental treatments. The model yields two intuitive predictions: enforcing common knowledge and enforcing bargaining both increase enrollment.

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:[15] (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 a stigma cost from breaking social norms, (iv)

[14] 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, we this procedure could introduce selective attrition as surveyors would reveal treatment status when setting appointments. To address this, we modified the procedure roughly 25% of the way through. 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. Encouragingly, we do not see any selective attrition (Section 6.2). One remaining concern is that the postponing of job information meant that, on average, *Discuss* couples received job information closer to the enrollment date. We show in Section 7.2.2 (and Table A11) that enrollment is not related to proximity to enrollment day.

[15] We detail a collective model with these features in Appendix B.1, showing how the net utility parameters can be micro-founded using standard assumptions.

bargaining power shifts in the wife's favor, increasing her consumption share.[16] While the features that underlie the net utility gains are not important for our purposes, what is important is that v_h and v_w are known with certainty by both the husband and the wife.

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 $\mu_w > \mu_h$.[17] These assumptions match key features of our data on preferences: 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 $\beta > \frac{1}{2}$,

[16] 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. That said, in a subsequent experiment to ours McKelway (2021a) finds only weak evidence for effects of Obeetee's weaving job on bargaining power. Given McKelway's result, the model we describe in Appendix B.1 includes features (i) to (iii), but does not include feature (iv).

[17] As we show in Appendix B.1, the assumption of uniformly distributed net utility parameters is an implication of a collective 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.

but not a necessary one for the two empirical predictions we will emphasize.[18]

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 $v_h v_w \geq 0$,[19] 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.

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 infor-

[18] 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.

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

mation when

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

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. We map our main treatments to the decision-making types above, and derive enrollment rates implied by the model.[20] First we consider the condition closest to the status quo: the husband is approached and given the job information, and can withhold that information from his wife with plausible deniability ($Husband - NoInfo$ in Figure 3). 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 4. 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 approached and given the job information, and can withhold the information ($Wife - NoInfo$ in Figure 3):

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

The enrollment region is the shaded area in panel (b) of Figure 4. Now enrollment only occurs when the

[20] Proofs in Appendix B.2.

wife benefits on net, and is willing to pay the bargaining cost for the cases where the husband would lose out.

As we discuss below, our experts predicted that enrollment would be 3.2 percentage points higher when the wife receives the job information, with plausible deniability, than in the status quo. While this seems intuitive, the sign of $E_w^{\text{NoInfo}} - E_h^{\text{NoInfo}}$ is ambiguous in the model, 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 matter more for any subsequent collective decision.[21]

Enforcing Common Knowledge. 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 4 ($Info$ treatment in Figure 3, with no modelling here of the gender of ticketed spouse treatment). 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. As explained above, 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

[21] A third unmodelled force would also push in the same direction: if the bargaining cost c is lower for husbands than wives they would have an additional incentive for sharing information. Given the imbalance in bargaining power, this extra feature is perhaps plausible – it may be less costly for husbands to start a conversation about their spouse making a sacrifice for the household than vice versa.

as opposed to wives.

Enforcing Bargaining. 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 3). 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 4) 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 $v_h v_w < 0$. We test for heterogeneity by agreement below.

Summary. After making a set of standard and context-appropriate assumptions, the model yields two unambiguous predictions. First, enforcing common knowledge increases enrollment. This results from the fact that some spouses have incentives to strategically withhold job information in order to prevent enrollment. Second, enforcing bargaining increases enrollment. This results from the fact that even with complete information, some spouses do not pay the cost to engage in bargaining, resorting to non-cooperative decision-making, which lowers enrollment.

In contrast, the model is ambiguous on the sign of the interaction effect between $Info$ and the gender of the ticketed spouse. And while we have random variation in which spouse receives the ticket throughout, for the *Discuss* treatment we do not explicitly model a role for the ticket recipient. We nevertheless discuss the full set of treatment effects in Section 7 below, after first comparing the predictions of intra-household experts with those of the model.

5 Expert Survey

We drew up a list of academic experts that included (i) authors of papers on intra-household economics[22] 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

[22]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.

NEUDC in the last 10 years;[23] (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;[24] (ii) those without publicly available email addresses; and (iii) two academics we were certain already knew the results. We emailed the expert survey to a final list of 361 researchers in May 2021 (full questionnaire in Appendix C).

Since the first draft of this paper circulated in May 2017, our expert survey screened out any researchers that had seen and could remember at least some of the results. In particular, we asked experts *"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 (Stanford)?"* We asked this question twice – once before describing the context and details of the experiment, and once afterwards. We only asked experts for their predictions if they answered *"No"*, *"Don't Know"*, or *"Yes, but I do not remember any of the results at all"* each time.

90 experts (25% of those contacted) completed the survey, while 70 experts (19% of those contacted) got past the two 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 – for example, while 34% of the experts that gave predictions are Assistant Professors, only 20% of the experts contacted are (Table A2). Reflecting this pattern, our predictors also have fewer Google Scholar citations. That said, our sample of predictors is not "inexperienced" overall – 23% are Full Professors, and they have 3,507 Google Scholar citations on average.

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 2), 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.[25] We did not incentivize the predictions given that experts might have been able to find the results from a previously circulated draft.[26]

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

[24] We kept researchers who had shifted to the private sector very recently.

[25] Nevertheless, we report the unpooled expert predictions in Figure A5.

[26] Even without incentivizing, experts might try to find the results. To avoid this, we removed the draft of the paper from each of our websites while the survey was live. The inaccuracy of the expert predictions suggest that these efforts succeeded.

Experts agree with the two main predictions of the 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 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. The patterns 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 A6).

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?"* We document overconfidence of experts in Section 7.3.

6 Data and Sample

6.1 Empirical Specification and Outcomes

We estimate variants of the following empirical specification:

$$y_i = \alpha_s + \beta \mathbf{Treat}_i + \gamma \mathbf{X}_i + \varepsilon_i$$

where i denotes a married couple, and α_s are strata fixed effects.[27] $\mathbf{Treat}_i$ is a vector of indicators denoting treatment assignment that varies depending on the hypothesis being tested. The vector $\mathbf{X}_i$ includes the 12 baseline survey variables used below for balance checks.[28] We report core estimates with and without $\mathbf{X}_i$. y_i is an indicator for the couple enrolling on enrollment day. We estimate robust standard errors throughout.

To explore mechanisms we consider other outcomes. To test directly for information withholding,

[27] Randomization was stratified by village, hamlet, and OBC status to create 36 total strata. After allowing for baseline attrition, some strata do not have all six treatment groups represented. To make use of all observations in identifying treatment effects, we pooled strata that did not have one member of each treatment group. Whenever required, we pooled two strata in the same village and caste but different hamlets. Pooling across different hamlets rather than across castes or villages keeps members of a strata as similar to one another as possible. After pooling, we have 25 strata.

[28] We set missing observations in $\mathbf{X}_i$ to 0 and include missingness dummies for each of the 12 variables.

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.

To better understand the effects of discussion we estimate effects on retention, whether the eligible woman enrolled on or within one month of enrollment day and stayed in the job for at least one month, and on four variables from the endline survey: (i) who had the most influence over the enrollment decision, (ii) the number of discussions the couple had about the job opportunity;[29] (iii) how much spouses disagreed about whether the wife should enroll (on a zero to two scale); and (iv) how inconsiderate one's spouse was of one's own opinion (on a zero to two scale).[30]

6.2 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. 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 A7). 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%),[31] or one or both members of the couple not consenting to take the survey (9%). This leaves us

[29] We measure the number of discussions by asking for the number of times a participant discussed the job opportunity with their spouse. If the participant was in the *Discuss* group, we asked for the number of times the two discussed the opportunity excluding the discussion during the survey, and then added one to this number if the surveyor reported that the couple discussed the opportunity during their survey.

[30] The data on considerateness was only collected for the last three centers. Adults outside of the *Discuss* treatment who said they had no discussions were not asked how considerate their spouse was. Non-ticketed adults in $NoInfo$ that did not believe they had received a ticket were not asked how much the couple disagreed about the enrollment decision.

[31] 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 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, before we had precise estimates of how long baseline surveys would

with a sample of 495 married couples for analysis. Reassuringly, attrition does not differ significantly by treatment group (column 1, Table A3). Figure 3 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, the sample looks well-balanced on baseline variables. We regress each of 12 baseline variables on indicators for $Info$, *Discuss*, and Wife Gets Ticket, along with strata fixed effects (columns 2 to 13, Table A3). Four of the 36 coefficients, and two of 12 tests of joint significance of the three treatment indicators, are statistically significant at the 10% level, which is close to the amount of imbalance we would expect from random chance alone. Nevertheless, some imbalances look potentially important. 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 A4). 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. We anyway estimate the main effects with and without the set of 12 controls, and our core conclusions are robust to either approach.

Endline Attrition and Balance. Of the 990 married adults in the 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 A7). 10% of attrition was due to an adult not consenting to participate in the endline survey, and 4% was due to the participant not having a phone.[32] Attrition from baseline to endline is largely balanced across treatment groups (columns 1 and 2, Table A5), 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 chance imbalances (columns 3 to 14, Table A5).

7 Results

We find virtually the opposite of what the experts and our theory predicted. Figure 6 summarizes our results, visualizing enrollment rates for each of the six treatment cells, and also testing for differences relative to the status quo group (the $NoInfo$, Husband Gets Ticket group) and relative to the expert predictions (the same analyses for the full eight treatment cells are in Figure A8). We unpack the treatment effects in the subsections that follow, but two features of Figure 6 are worth highlighting now. First,

require and before we had implemented the system of appointment setting. Since baseline surveys were randomly ordered, attrition for this reason is not selective.

[32]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.

our treatments did not raise enrollment and, if anything, lowered it. Second, we can reject four of five expert-predicted treatment effects at at least the 10% level.

7.1 Information Withholding

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 increased enrollment (columns 1 and 2, Table 3). The point estimates suggest that the $Info$ treatment, if anything, reduced enrollment by three percentage points. We can reject the mean expert prediction of a 5.5 percentage points increase with 95% confidence, and in the specification with controls, we can reject a majority of the individual-level expert predictions at the 5% significance level. 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 two 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 ($p = 0.79$ with controls),[33] and each of these effects is again quite different to the mean expert prediction ($p = 0.07$ and $p = 0.1$ with controls). Neither wives nor husbands are withholding information.

More surprising, there is no evidence for strategic withholding even among those the theory emphasizes. The point estimates continue to be negative when we restrict to the sample of couples where the ticketed spouse rates the appropriateness of women weavers more negatively than the non-ticketed spouse (columns 5 and 6, Table 3). While in this case we lack expert predictions for comparison, we can reject positive effects of more than 2.6 percentage points at the 95% level.

Consistent with $Info$ not increasing enrollment, information diffusion is high. In the $NoInfo$ treatment, 74% of wives are aware that their spouse received a job ticket when the ticket was given to the husband. Similarly, 71% of husbands are aware when the ticket was given to the wife. The $Info$ treatment

[33] 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 ($p = 0.94$ with controls). It follows that while the experts mispredict the level of the effects of $Info$ separately for the husband-ticket and wife-ticket cases, 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.

increases this awareness by nine percentage points when the husband gets the ticket, and 13 percentage points (though not significantly different to the nine) when the wife gets the ticket (columns 7 and 8, Table 3, $p = 0.04$ for pooled $Info$ effect). These increases are small, and 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. These effects on ticket knowledge show most directly that neither husbands nor wives are strategically withholding information.

The non-positive effects of $Info$ suggest that our theory and the mental models of most experts are incomplete. When spouses are given opportunities to withhold information to advance their own interests, they do not do so. Four distinct model features could rationalize these results. First, parameter values might be such that the bargaining power of husbands is high enough that they never need to withhold information, and job preferences of wives positive enough that they never want to. But given the heterogeneity in our sample (recall Figure 1), it seems unlikely that all couples could have bargaining power and preferences within this range.

Second, spouses may have internalized norms of honesty, and consequently feel compelled to share relevant information, even if that information sharing is 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"*.

Third, respondents may not have understood or believed they had plausible deniability. We find a lack of understanding 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?" If the respondent answered incorrectly in the $NoInfo$ group, 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." If the respondent answered correctly, the surveyor would say the answer was correct before continuing. While participants are likely to have understood what we told them, we cannot fully rule out that they did not believe it. Relatedly, 10% may have been too low for the fraction of couples without tickets – individuals may not have known enough couples without tickets for not having one to seem plausible.

Fourth, decision-making may be such that information sharing is never actually costly. For example, if the household did not bargain and instead chose enrollment as long as neither spouse vetoed it, there would be no incentive to withhold information. We discuss some supportive evidence of this third

explanation in the next subsection.

7.2 The Effect of Discussion

The theory predicts that discussion ensures bargaining, maximizing enrollment. Experts agree with the prediction, expecting enrollment to be 11.6 percentage points higher in *Discuss* than in $NoInfo$. In reality, discussion lowers enrollment by seven to nine percentage points relative to $NoInfo$ (columns 1 to 2, Table 4, $p < 0.01$ with controls), and lowers enrollment by four to six percentage points relative to $Info$ ($p = 0.07$ with controls). Given that $NoInfo$ enrollment is already low at 18%, discussion reduces enrollment by 39 to 50%.

Focusing on the specification with controls, we reject the mean expert prediction of *Discuss* relative to $NoInfo$ ($p < 0.001$) and of *Discuss* relative to $Info$ ($p < 0.001$). In addition, we can reject every single expert's *Discuss* versus $NoInfo$ prediction at the 5% level, and 83% of their predictions for *Discuss* relative to $Info$. It is not just that experts mispredict the effects of discussion on average; essentially no expert even comes close.[34]

Recall our theory predicts that positive effects of discussion should be driven by couples that disagree about whether the wife should enroll. Contrary to the theory, our discussion estimates remain negative when we focus on the set of spouses that disagree about the appropriateness of women weaving (columns 3 to 4, Table 4).

A potential concern with the surprising negative effect of discussion is that it may be a "fluke" result that would not replicate. This is of course difficult to rule out with a single experiment, but we can speak to this concern by considering effects of discussion for each village separately. As discussed above, our sample comes from six villages and the experiment was run in each village sequentially. If the discussion effect is a fluke, we might expect a small number of villages to drive it. But the *Discuss* point estimate relative to $NoInfo$ is negative in five of the six villages (Table A6).

7.2.1 Proposed Mechanism

Why did discussion reduce enrollment? We find support for an alternative theory that rests on two ideas. First, households do not bargain, perhaps because bargaining is extremely costly and couples live in non-cooperative marriages (Lundberg and Pollak 1993). Spouses will only make a decision that deviates from the status quo, like choosing to enroll in a job, if neither vetoes it. Second, informing only one

[34] Two of 70 experts predicted a negative effect of *Discuss* relative to $NoInfo$, though in column 2 of Table 4 we even reject their predictions at the 5% level. We reject one for being too negative (-17 percentage points) and the other for not being negative enough (-2 percentage points).

spouse, as in $NoInfo$, weakens the other's veto power, while interventions that nudge households toward joint decision-making, like our $Info$ and *Discuss*, give both equal veto power. Specifically, in the $Info$ and *Discuss* groups, both spouses have full and equal power to veto enrollment – enrollment only occurs when *both* spouses prefer enrollment over non-enrollment ($v_h, v_w > 0$, panel (c), Figure 7). On the other hand, informing only one spouse, like our $NoInfo$ treatment did, creates a norm that he or she make the decision. Uninformed spouses incur a utility cost from violating the norm if they veto, so they will only veto if they are very opposed to enrollment. In particular, in $NoInfo$, 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 ($v_j > v_j^*$, where $v_j^* < 0$ and $j = w$ or h) (panels (a) and (b), Figure 7). v_h^* and v_w^* need not be equal, and gender norms in our setting might suggest $v_h^* > v_w^*$ (i.e. husbands maintain more control when wives are informed than wives maintain when husbands are informed). We draw the enrollment regions in Figure 7 with $v_h^* > v_w^*$.[35]

This alternative theory rationalizes the lack of information withholding and the negative effect of discussion. In this world, the decision will never go against the ticket holder's wishes so he or she cannot lose from sharing information. Likewise, enrollment in *Discuss* requires full support from both spouses, while only one needs to be fully supportive in $NoInfo$, which could explain why enrollment was so much lower in the former than the latter.

Further, and in contrast to our initial model in Section 4, the alternative theory predicts that intervention toward joint decision-making should reduce enrollment only when the non-ticketed spouse is less supportive of enrollment than the ticketed one. While we hinted at the role of disagreement between spouses in Tables 3 and 4, we now estimate treatment effects separately for two subsamples: the couples in which the non-ticketed spouse rated women weaving as less appropriate at baseline than the 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 reduced enrollment when the non-ticketed spouse was less supportive of female weavers than the ticketed one (panel (b), Figure 8), and did not affect enrollment for the other couples (panel (a)). Table 5 tests whether the effects differ between these two subgroups. Focusing on the regressions with controls, the difference in the $Info$ effects between these two subgroups is 13 percentage points ($p = 0.12$), while the difference in the *Discuss* effects is 16 percentage points ($p = 0.04$) (column 2). Pooling the two treatments, we estimate that any joint decision-making intervention ($Info$ or *Discuss*) had a negative 16 percentage point impact among couples in which the non-ticketed spouse was less supportive, and no effect among the rest (column 4).

[35] This, coupled with the fact that more women support enrollment than men, means there is no prediction for whether enrollment will be higher when the husband or the wife receives the ticket in $NoInfo$.

One potential concern with the interpretation of these patterns of heterogeneity is that when one spouse is less supportive of female weavers than the other, it is more often husbands who are less supportive than wives (panel (a), Figure 1). Perhaps these patterns of heterogeneity by disagreement are not about disagreement per se, but rather are picking up on heterogeneity by wives' support or by husbands' opposition. But against this, we see no heterogeneity by wives' or husbands' individual reports on the appropriateness of female weavers (Table A7).

The new theory has an additional prediction for which we find suggestive support: the decision-making power of the ticket-holder should be lower in $Info$ and *Discuss* than in $NoInfo$. Ticket-holders in $Info$ and *Discuss* do in fact report less influence over the enrollment decision and more joint decision-making (Table A8), although the coefficients are not statistically significant.

We note a limitation of our explanation for the negative effect of discussion: while this theory rationalizes 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).

In sum, our initial model – the efficient, collective model with asymmetric information and a bargaining cost – predicted positive effects of information and discussion. The lack of information withholding and negative effect of discussion that we observe are instead consistent with a model with more substantial deviations from the efficient model. In particular, these results are not consistent with bargaining, but are consistent with households making decisions about enrollment through vetoes. In this world, our interventions that treated spouses symmetrically gave each veto power, while informing only one of them weakened the other's veto power. Consistent with this model, our interventions toward joint decision-making have more negative effects when non-ticketed spouses are less suppotive of female weavers than their ticketed spouses.

7.2.2 Competing Channels

Beyond the veto power explanation, we find little evidence for seven alternatives.

Discussion treatment was awkward. One possible explanation for the negative effect of discussion is that couples 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, they should also have been 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 for neither gender can we reject that individuals were surveyed at similar rates in *Discuss* and $NoInfo$ (columns 1 and 2, Table A5).

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. Two results make this argument explanation seem unlikely. First, surveyors recorded whether they perceived the three-minute discussions to be argumentative or not. None of the discussions were classified as argumentative. Second, both wives and husbands in the *Discuss* group report at endline that they were no more likely to have disagreed about the job opportunity (columns 1 and 2, Table A9), and that their spouse was not more inconsiderate of their opinion of the job (columns 3 and 4, Table A9).[36] Relative to $NoInfo$, wives and husbands actually report that their spouse was *more* considerate of their opinion.

Surveyors' presence increased adherence to the norm that women not work. Another possible explanation is that surveyors being present while couples discussed the job increased adherence to the norm that women not work outside of their homes, and this reduced enrollment. If this were true, discussion should have had a more negative effect where this norm is stronger. We use the level of female employment in a woman's village to proxy for the strength of the norm her household faces. We compare effects in the three villages with the highest female employment at baseline (17% overall) to the three villages with the lowest (8% overall), controlling for a woman's own baseline employment. We cannot reject that the effect of *Discuss* is the same across these two groups of villages (Table A10).

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 A7). Similarly, to conform to norms of male authority, husbands may be less likely to give in when being watched. But the effects of discussion were no more negative when husbands were *less* supportive of weaving (Table A7).

[36] The results discussed in this paragraph are not necessarily at odds with the joint decision-making and veto power channel described above. Under that channel, *Discuss* and $Info$ change the "rules" of the household decision, giving both spouses veto power, but that need not produce argument.

Discussion led to deeper deliberation. The *Discuss* treatment may have prompted deeper deliberation – for one thing, *Discuss* couples discussed the job opportunity more often (columns 1 to 2, Table A11). If taking the job were actually the wrong decision for many couples, deeper deliberation could have reduced enrollment. Two facts speak against this deliberation channel. First, if the deliberation in the discussion treatment group leads to better informed decisions, we should expect that for those not in the *Discuss* group, those who have more days to decide should be less likely to enroll. In contrast, we if anything find the opposite (columns 3 to 4, Table A11). Second, if decisions made by *Discuss* couples are better informed, we should see higher retention among enrolling *Discuss* couples than among other enrolling couples, as these other enrolling couples learn on the job that the job is low quality. This in turn implies that the negative effects of *Discuss* on enrollment should attenuate over time. In contrast, *Discuss* has similar effects on whether the couple enrolled and remained in the job for at least one month (columns 5 to 6, Table A11) – a 43 to 53% drop in enrollment relative to $NoInfo$, and a statistically significant drop in enrollment relative to $Info$ at at least the 5% level.

Discussion treatment made misrepresenting information about the job harder. Another possibility is that informing couples together made it harder for the ticketed spouse to make the job seem better than it was 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. The sub-treatment improved knowledge about the job, but the effect on job desirability is not statistically significant and the point estimate is actually positive (columns 1 to 5, Table A1). The sub-treatment did reduce enrollment (column 6, Table A1), particularly when women were ticketed (Figure A8), but neither the overall effect nor the effect when women were ticketed is statistically significant. In sum, we do not find strong evidence for this channel, though we lack power to completely rule it out.

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.

7.3 Judging the Judges

Experts predicted that household intervention, relative to the status quo, would increase enrollment. These predictions were inaccurate, with our interventions reducing enrollment. Of our five interventions, we can reject four of five expert-predicted treatment effects at at least the 10% level (summarized in Figure 6, and shown for the full eight treatment cells in Figure A8). Nevertheless, while inaccurate overall, are experts well-calibrated in their confidence? And are some types of experts predictably more accurate than others?

Experts have relatively low confidence in their predictions. On average, experts expected that 2.7 of their 7 predicted enrollment rates would be within 5 percentage points of the actual enrollment rate. Nevertheless, experts are overconfident – only 1.4 predictions are actually within 5 percentage points of the truth (Figure 9). Furthermore, confidence is uncorrelated with accuracy, as with the academic experts in DellaVigna and Pope (2018). In both settings, confidence is not a helpful heuristic for good judgment.

While experience does not predict confidence (columns 1 to 3, Table A12), the more experienced experts tend to give *less* accurate predictions (columns 4 to 6). 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). Going from the lowest to highest-ranked citations is associated with an increase in average absolute forecast error of 54% of the mean. 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).

8 Conclusion

While inefficiency in household decision-making suggests a role for intervention, the success of intervention hinges on accurate mental models of decision-making processes. In this paper we show that theory, experts, and we mispredicted the effects of intervention into household decision-making in India. Our initial model – the efficient, collective model with asymmetric information and a bargaining cost – predicted that our information and discussion interventions would raise enrollment in a job for women. Instead, we find no effects of information and negative effects of discussion. We argue that these unexpected results are consistent with a model further from the efficient one. In our alternative model, spouses do not bargain and only choose to enroll if neither vetoes enrollment. Informing only one spouse about the job opportunity weakened the other's veto, while nudging households toward joint decision-making granted each full veto power.

In our setting, experts tended to be both inaccurate and overconfident in their predictions about joint

decision-making. An important question remains: are experts wrong about intervening in the household in general, or were our study results somehow unusually difficult to predict? While we cannot answer this question well without a meta-analysis of intra-household expert surveys,[37] we note the results of a study with important parallels to ours. Dean and Jayachandran (2019) ran a field experiment, concurrent with ours, on the retention of female teachers in South India. Their *Conversation* treatment has parallels with our *Discuss* treatment: a surveyor guided a conversation between the female teacher and her family members about the pros and cons of her working. While Dean and Jayachandran (2019) cannot reject the null hypothesis of no effect (given a smaller sample of $N = 171$ teachers), the estimated effect of *Conversation* on retention is negative, and similar to our own estimated effects of *Discuss* at –6 percentage points. Furthermore, if experts predicted the same treatment effect for *Conversation* as they did for *Discuss* (relative to $NoInfo$, i.e. an 11.6 percentage point effect), this prediction would be rejected by Dean and Jayachandran (2019) at the 5% level. While suggestive, these results suggest that experts may more generally mispredict the effects of interventions into household decision-making in India.

[37] Such a meta-analysis may eventually be possible given the increasing popularity of the Social Science Prediction Platform (socialscienceprediction.org, DellaVigna et al. 2019).

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.

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

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

_ , 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.

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, "Investment Decisions with Endogenous Budget Share Allocations Inside the Household," 2021.

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.

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.

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.

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.

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.

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

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.

Lundberg, Shelly and Robert A Pollak, "Separate spheres bargaining and the marriage market," *Journal of political Economy*, 1993, *101* (6), 988–1010.

McKelway, Madeline, "The Empowerment Effects of Women's Employment: Experimental Evidence," Working Paper, Hanover, New Hampshire 2021.

_ , "Women's Employment in India: Intra-Household and Intra-Personal Constraints," Working Paper, Hanover, New Hampshire 2021.

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.

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.

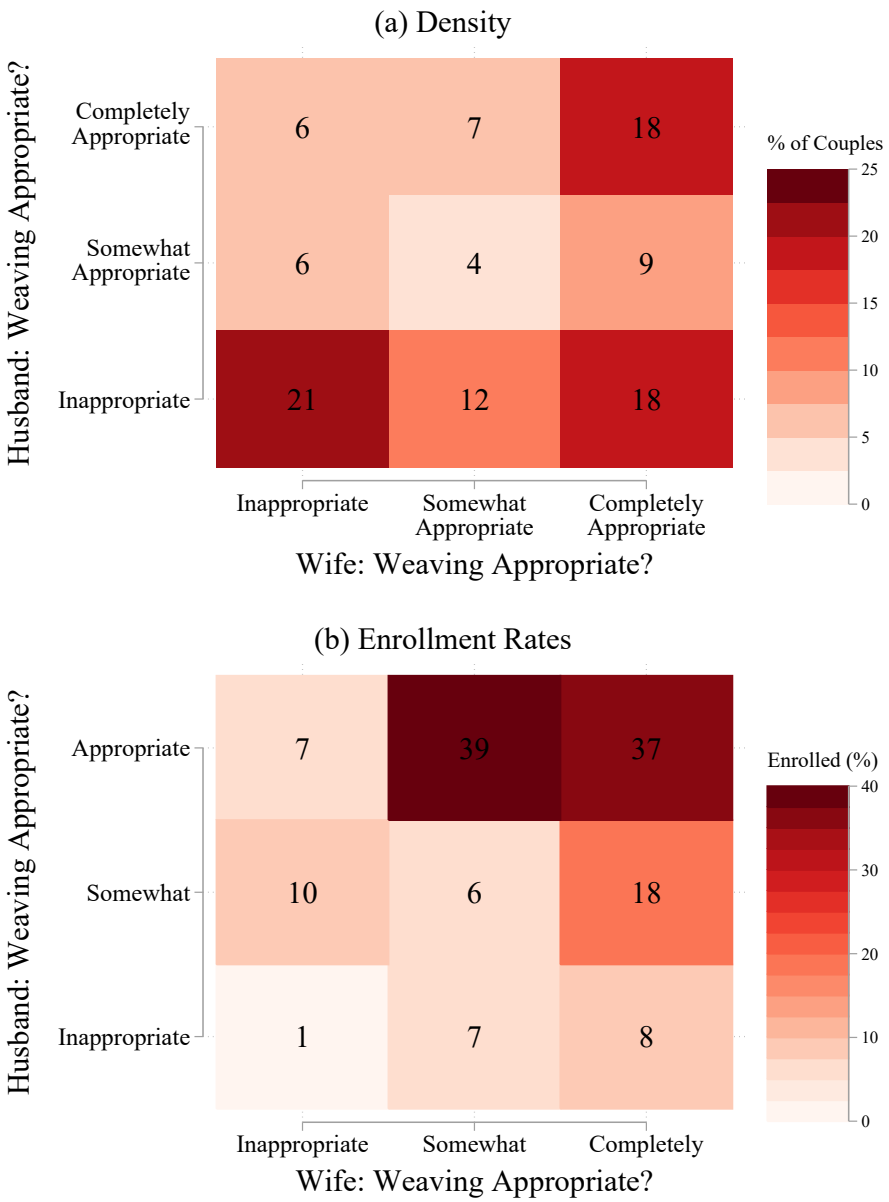
Schaner, Simone, "Do opposites detract? Intrahousehold preference heterogeneity and inefficient strategic savings," *American Economic Journal: Applied Economics*, 2015, *7* (2), 135–74.

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

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

Figures

Figure 1: Misaligned Preferences and Enrollment



Notes: Panel (a) shows a heatmap of the percentage of couples with each possible combination of answer 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 by each baseline preference combination.

Figure 2: Experiment Timeline

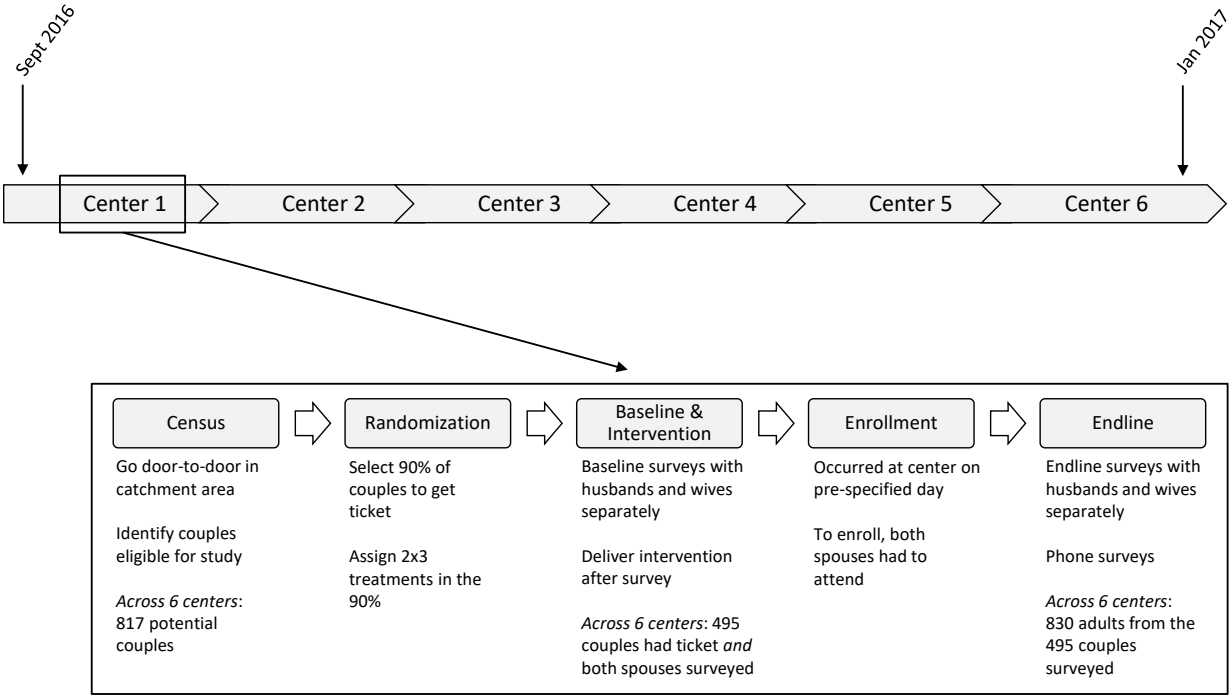
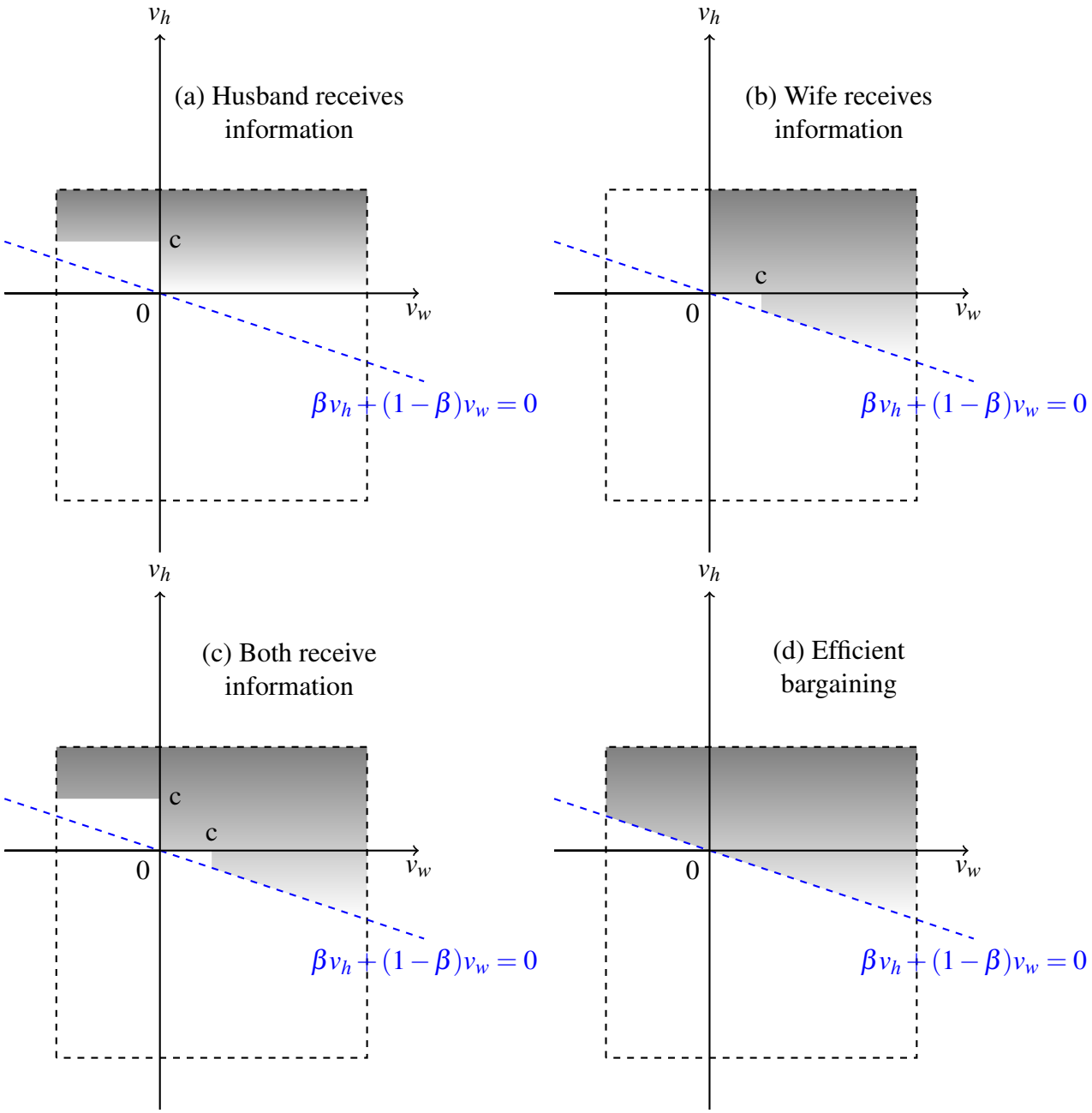


Figure 3: Experiment Design

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	

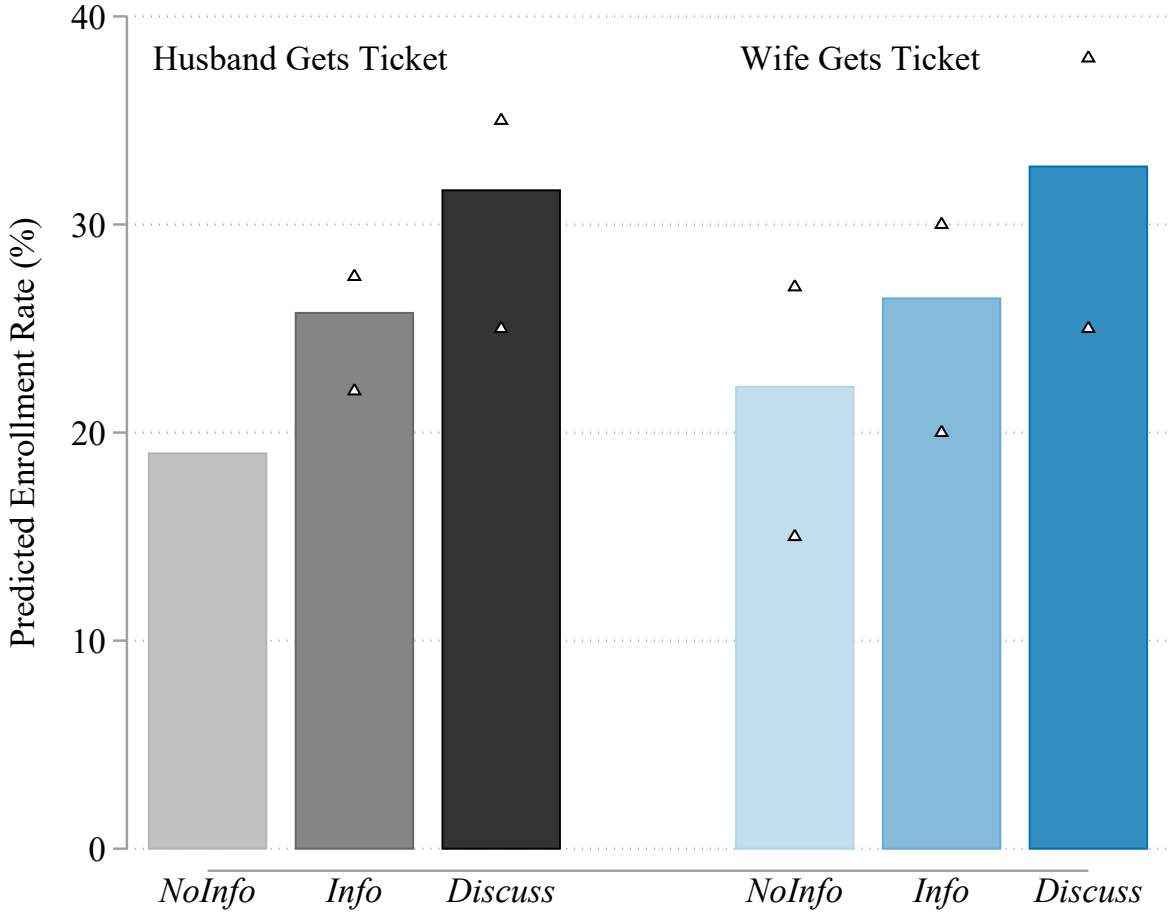
Notes: The figure visualizes our 3-by-2 experimental design, with randomization at the couple-level. The columns indicate which spouse was given the ticket required to enroll in the job. The rows indicate the information treatment for the non-ticketed spouse. In *NoInfo* the non-ticketed spouse is not 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 4: Enrollment Regions Under Initial Model



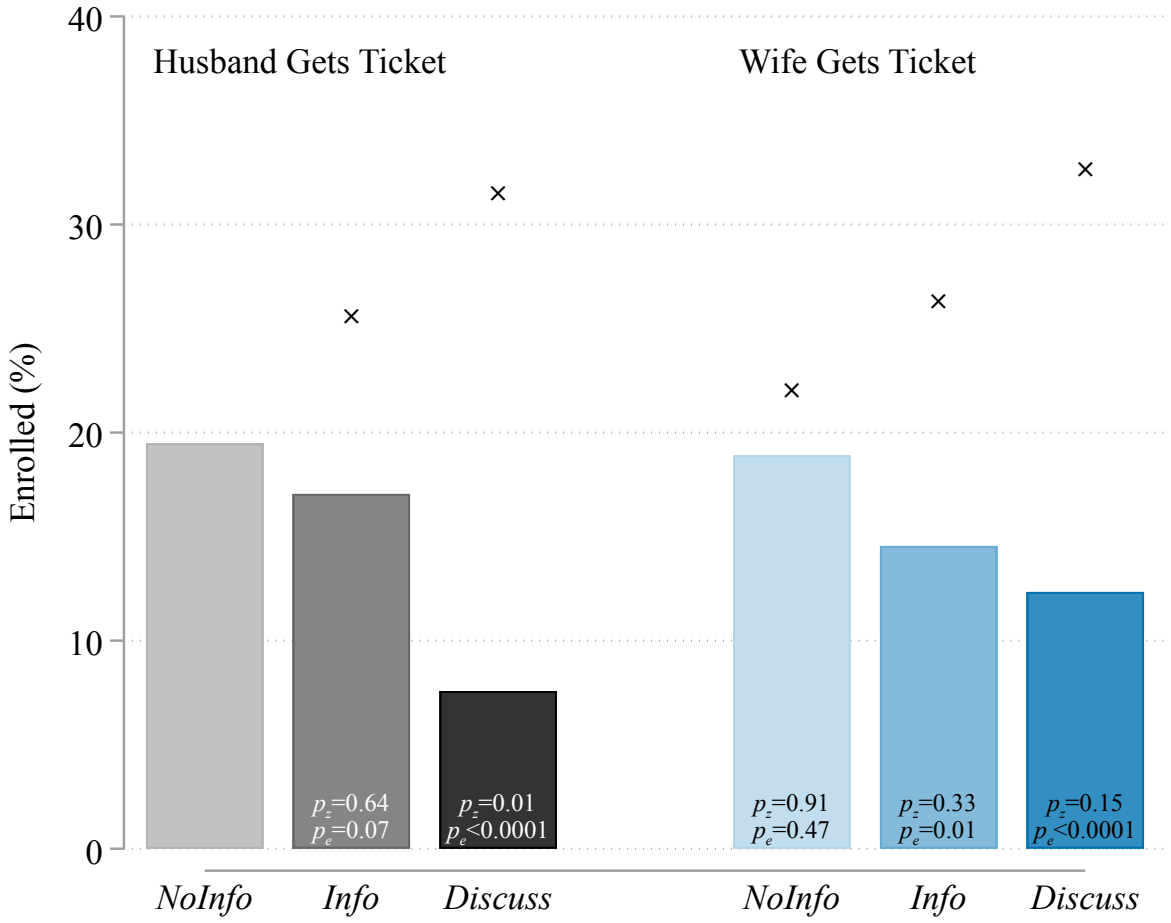
Notes: The figure is drawn for $v_h \sim U[-4,2], v_w \sim U[-2,4], \beta = \frac{3}{4}, c = 1$. The black dashed line box shows the region in which the preferences of husbands and wives are independently and uniformly distributed. The blue dashed line is the line of indifference for collective households. 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 5: Expert Predictions Line Up With Model Predictions



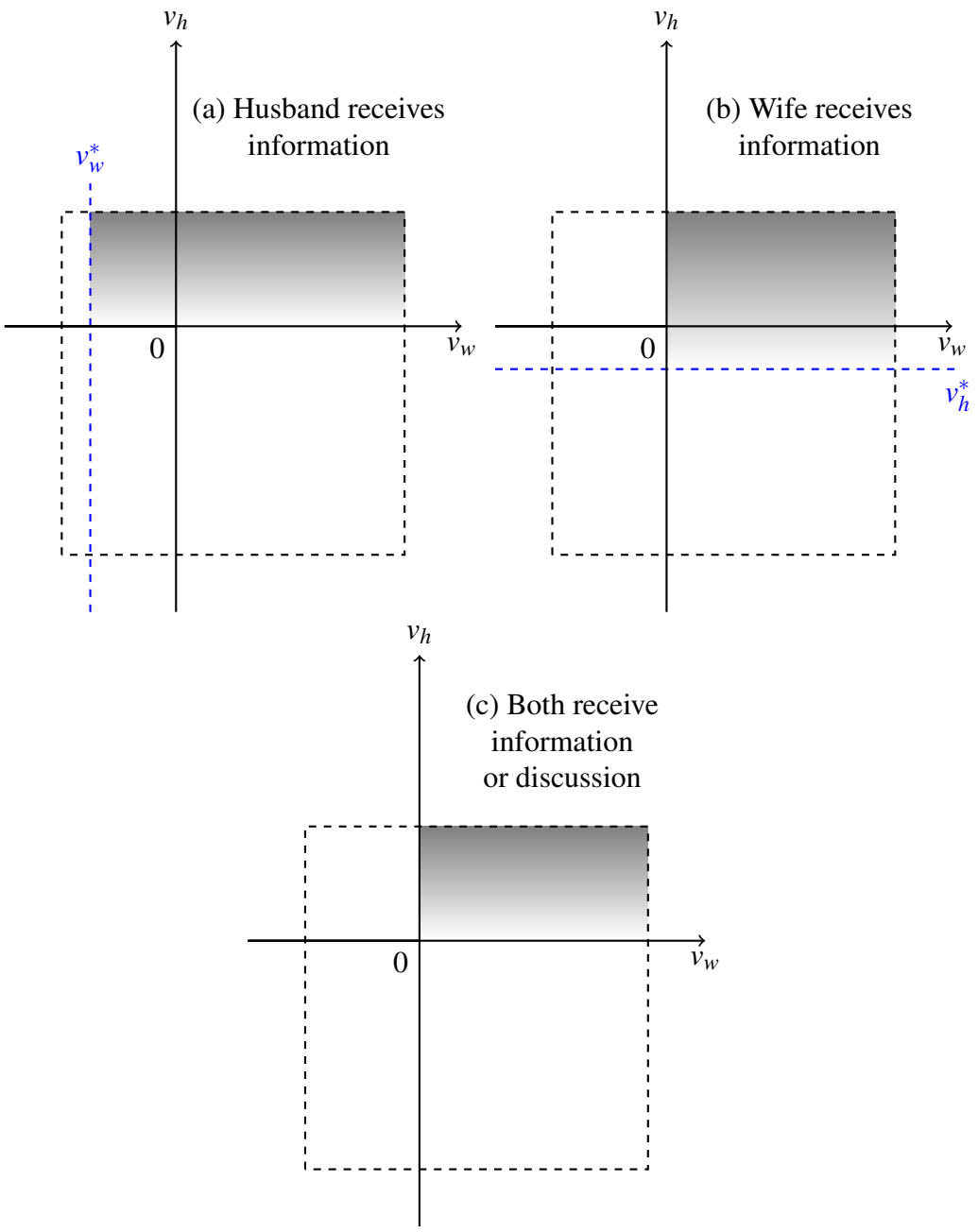
Notes: The figure shows the mean predicted enrollment rate for each treatment according to our full sample of 70 experts on intra-household economics. Each expert was told the true enrollment rate (19%) of the treatment in which the husband gets the job ticket and the wife doesn't know, while the other five bars reflect predictions. Triangles denote the 25th and 75th percentile prediction for each treatment.

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



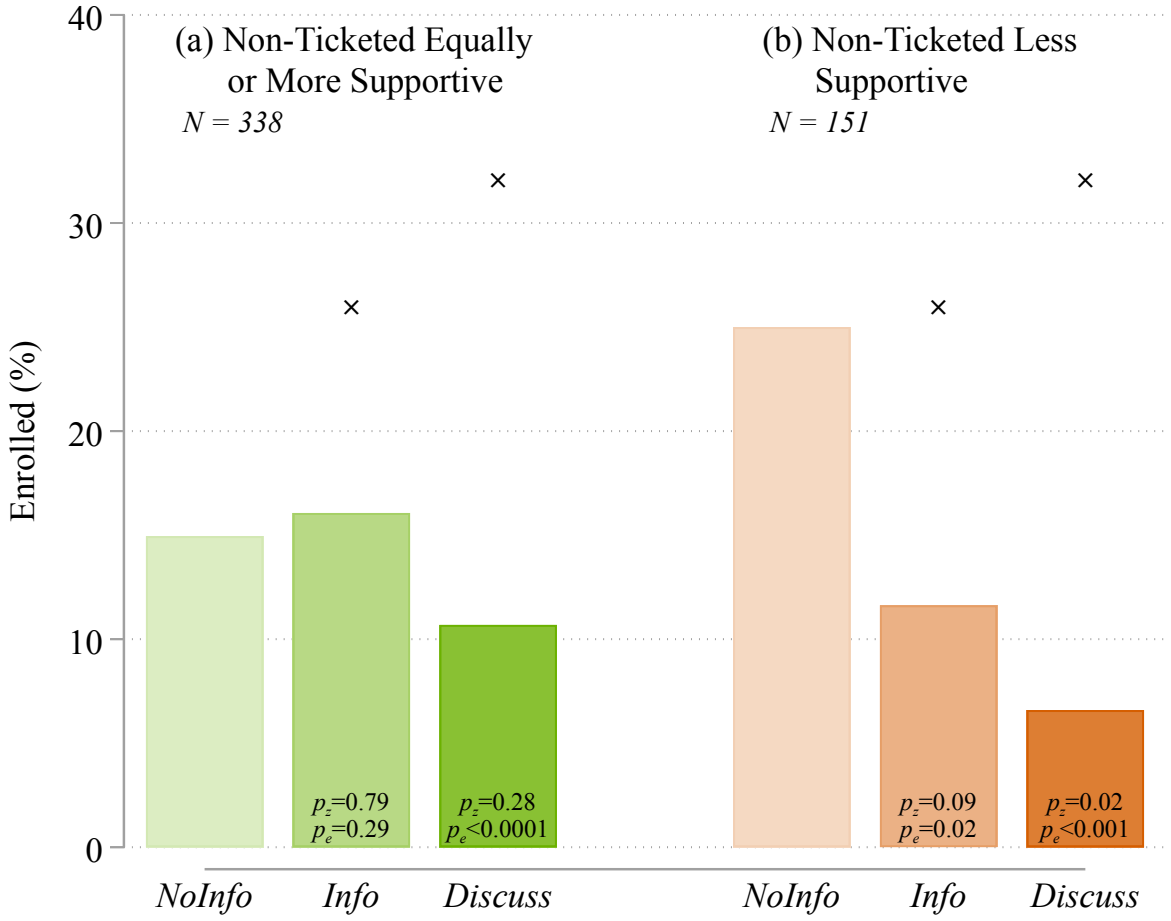
Notes: The figure visualizes enrollment rates across the six treatments, derived from a regression ($N = 495$) of enrollment on treatment indicators, strata fixed effects, and the following controls: (1) age of wife, (2) age of husband, (3) indicator for wife has no education, (4) indicator for husband 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), and missingness dummies for each of these variables. 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 7: Enrollment Regions Under Alternative Model



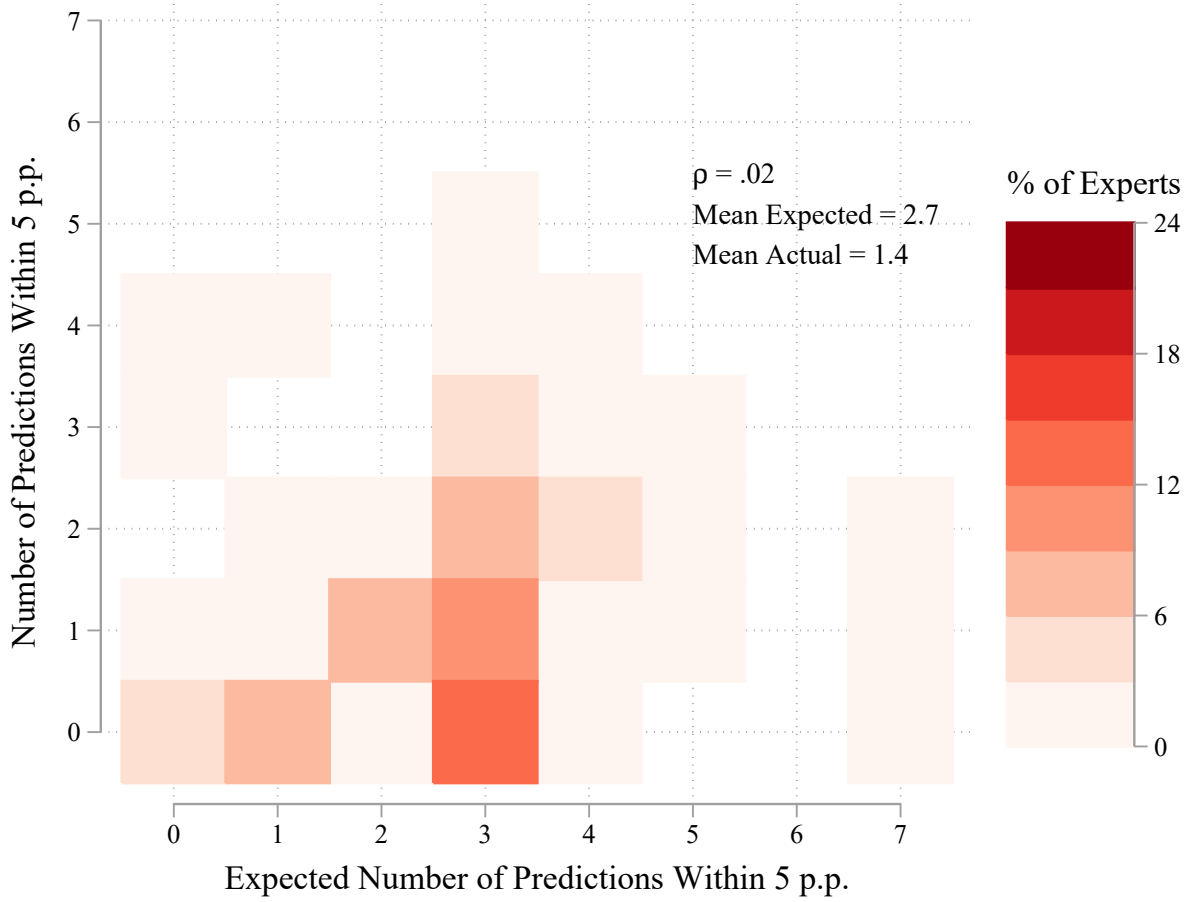
Notes: The figure is drawn for $v_h \sim U[-4,2], v_w \sim U[-2,4], v_h^* = -0.75, v_w^* = -1.5$. The black dashed line box shows the region in which the preferences of husbands and wives are independently and uniformly distributed. The blue dashed lines represent the level of utility below which the 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 4, we now assume full information diffusion, full veto power for each spouse in (c), and weakened veto power for the uninformed spouse in (a) and (b).

Figure 8: Joint Decision-Making Interventions Reduce Enrollment Only When Non-Ticketed is Less Supportive



Notes: The figure visualizes enrollment rates across the three information treatments, separately for two subsamples: couples in which the non-ticketed spouse rated women weaving as less appropriate at baseline than the ticketed spouse did (panel (b)); and the rest of the couples, in which the non-ticketed spouse rated women weaving as equally or more appropriate (panel (a)). The enrollment rates are then derived from two regressions, each regressing enrollment on indicator variables for $Info$ and *Discuss*, strata fixed effects, 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), and missingness dummies for each of these variables. 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.

Figure 9: Experts Are Overconfident



Notes: The figure shows a heatmap of the percentage of 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*), 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.

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, and missingness dummies for each of these variables. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Neither Husbands nor Wives Withhold Job Information

	Enrolled						Non-Ticketed Knows	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Info	-0.03	-0.03			-0.07	-0.10		
	(0.04)	(0.04)			(0.07)	(0.07)		
Info × Husband Gets Ticket			-0.02	-0.02			0.09	0.09
			(0.05)	(0.05)			(0.07)	(0.07)
Info × Wife Gets Ticket			-0.03	-0.04			0.13*	0.13*
			(0.05)	(0.05)			(0.07)	(0.07)
Observations	495	495	495	495	136	136	414	414
Omitted Group Mean	0.18	0.18	0.19	0.19	0.11	0.11	0.74	0.74
p(*Info* = Expert Mean)	0.03	0.02						
% Experts Rejected	46	56						
p(*Info* × Husb. = *Info* × Wife)			0.85	0.79			0.68	0.71
p(*Info* × Husb. = Expert Mean)			0.10	0.07				
p(*Info* × Wife = Expert Mean)			0.16	0.10				
Strata	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Discuss	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Discuss × Husband Gets Ticket	No	No	Yes	Yes	No	No	Yes	Yes
Husband Gets Ticket	No	No	Yes	Yes	No	No	Yes	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 their spouse received an enrollment ticket. Columns (5) and (6) include only the sample of couples where the ticketed spouse rates the appropriateness of women weavers more negatively than the non-ticketed spouse. % Experts Rejected is the percentage of 70 expert predictions of the effect of *Info* relative to *NoInfo* that can be rejected at a 5% significance level. 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 missingness dummies for each of these variables. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Discussion Reduces Enrollment

	Enrolled			
	(1)	(2)	(3)	(4)
Info	-0.03	-0.03	-0.08	-0.09
	(0.04)	(0.04)	(0.05)	(0.05)
Discuss	-0.07*	-0.09***	-0.10**	-0.14***
	(0.04)	(0.03)	(0.05)	(0.05)
Observations	495	495	289	289
Omitted Group Mean	0.18	0.18	0.19	0.19
p(*Discuss* = *Info*)	0.22	0.07	0.57	0.23
p(*Discuss*-*Info* = Expert Mean)	0.003	<0.001		
% Experts Rejected (*Discuss* vs. *Info*)	74	83		
p(*Discuss* = Expert Mean)	<0.001	<0.001		
% Experts Rejected (*Discuss* vs. *NoInfo*)	96	100		
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. Columns (3) and (4) include only the sample of disagreeing couples: where each spouse rates the appropriateness of women weavers differently. % Experts Rejected is the percentage of 70 expert predictions that can be rejected at a 5% significance level, first considering the effect of *Discuss* relative to *Info*, then the effect of *Discuss* relative to *NoInfo*. 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 missingness dummies for each of these variables. *** p<0.01, ** p<0.05, * p<0.1.

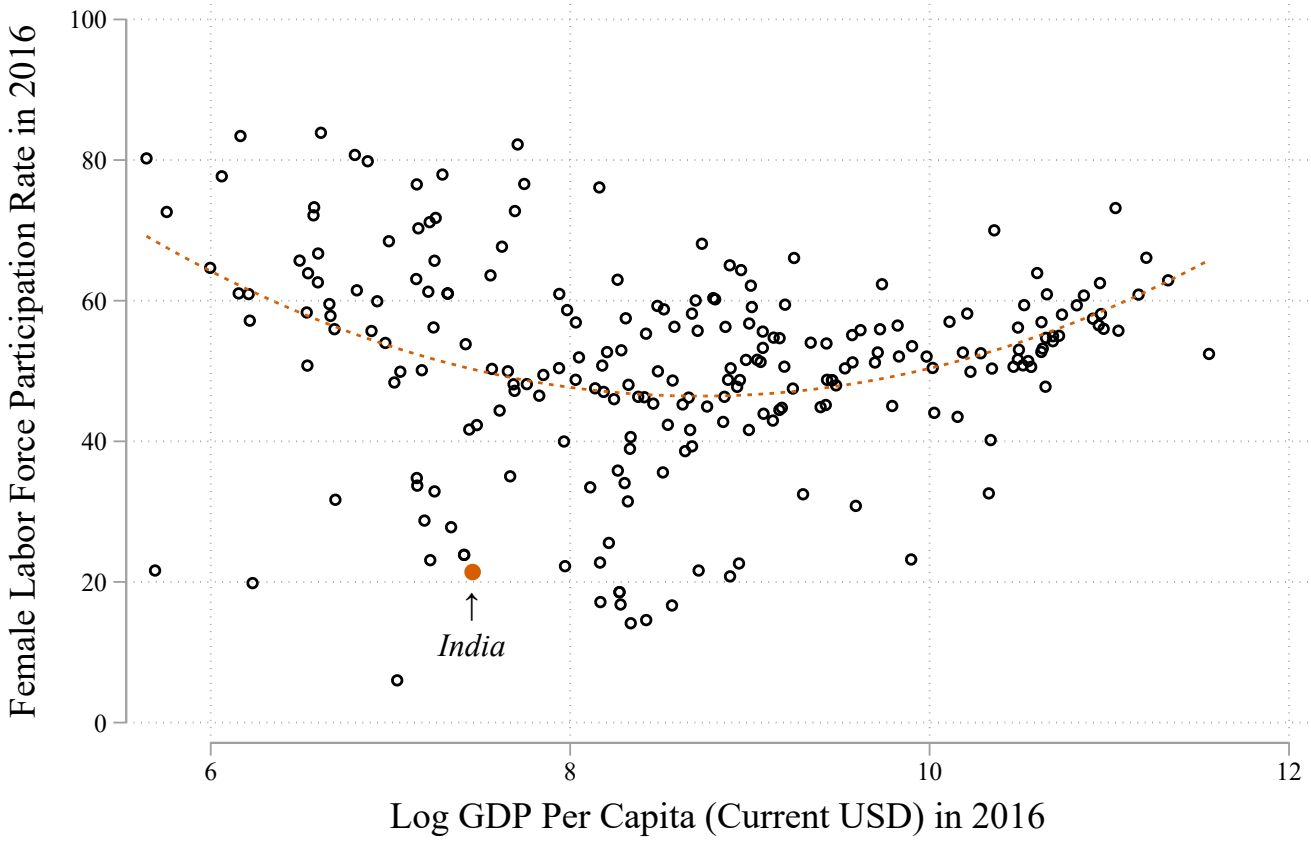
Table 5: Effects Are More Negative When Non-Ticketed is Less Supportive

	Enrolled			
	(1)	(2)	(3)	(4)
Info	0.01	0.01		
	(0.04)	(0.04)		
Discuss	-0.04	-0.05		
	(0.04)	(0.04)		
Info × Non-Ticketed Less Supportive	-0.12	-0.13		
	(0.09)	(0.08)		
Discuss × Non-Ticketed Less Supportive	-0.11	-0.16**		
	(0.08)	(0.08)		
Info or *Discuss*			-0.01	-0.02
			(0.04)	(0.04)
(*Info* or *Discuss*) × Non-Ticketed Less Supportive			-0.12	-0.14**
			(0.07)	(0.07)
Observations	489	489	489	489
p(*Info* × Non-Ticketed Less Supportive = *Discuss* × Non-Ticketed Less Supportive)	.94	.7		
Strata	Yes	Yes	Yes	Yes
Non-Ticketed Less Supportive	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes

Notes: Robust standard errors in parentheses. Enrolled is an indicator for enrolling on enrollment day. Non-Ticketed Less Supportive is an indicator for the non-ticketed spouse rating women weaving as less appropriate at baseline than the ticketed spouse did. 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 missingness dummies for each of these variables. *** p<0.01, ** p<0.05, * p<0.1.

A Appendix [For Online Publication]

Figure A1: Female Labor Force Participation: India Versus the World

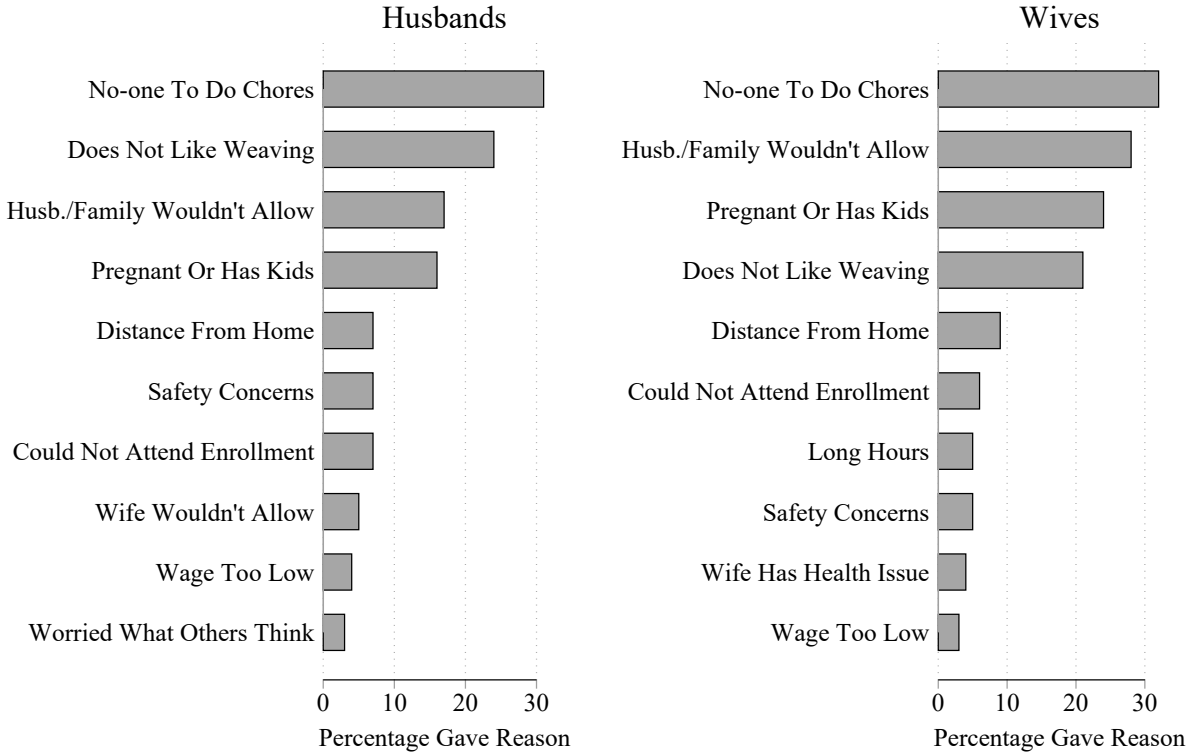


Notes: This figure plots the 2016 female labor force participation rate (% of female population aged 15+) against 2016 log GDP per capita (measured in current USD), along with a quadratic fitted line. The data covers 227 countries and derives from the World Bank's World Development Indicators. The variable codes are SL.TLF.CACT.FE.ZS (for female labor force participation) and NY.GDP.PCAP.CD (for GDP per capita).

Figure A2: A Hand-Knotted Loom

Notes: Obeetee manufactures and exports both hand-knotted and hand-tufted carpets. Hand-knotted carpets are made fully by hand, while hand-tufted carpets involve the use of a mechanized tool. The photo shows two men working on a hand-knotted carpet. In our experiment, women work only on hand-knotted carpets.

Figure A3: The Most Common Reasons for Non-Enrollment

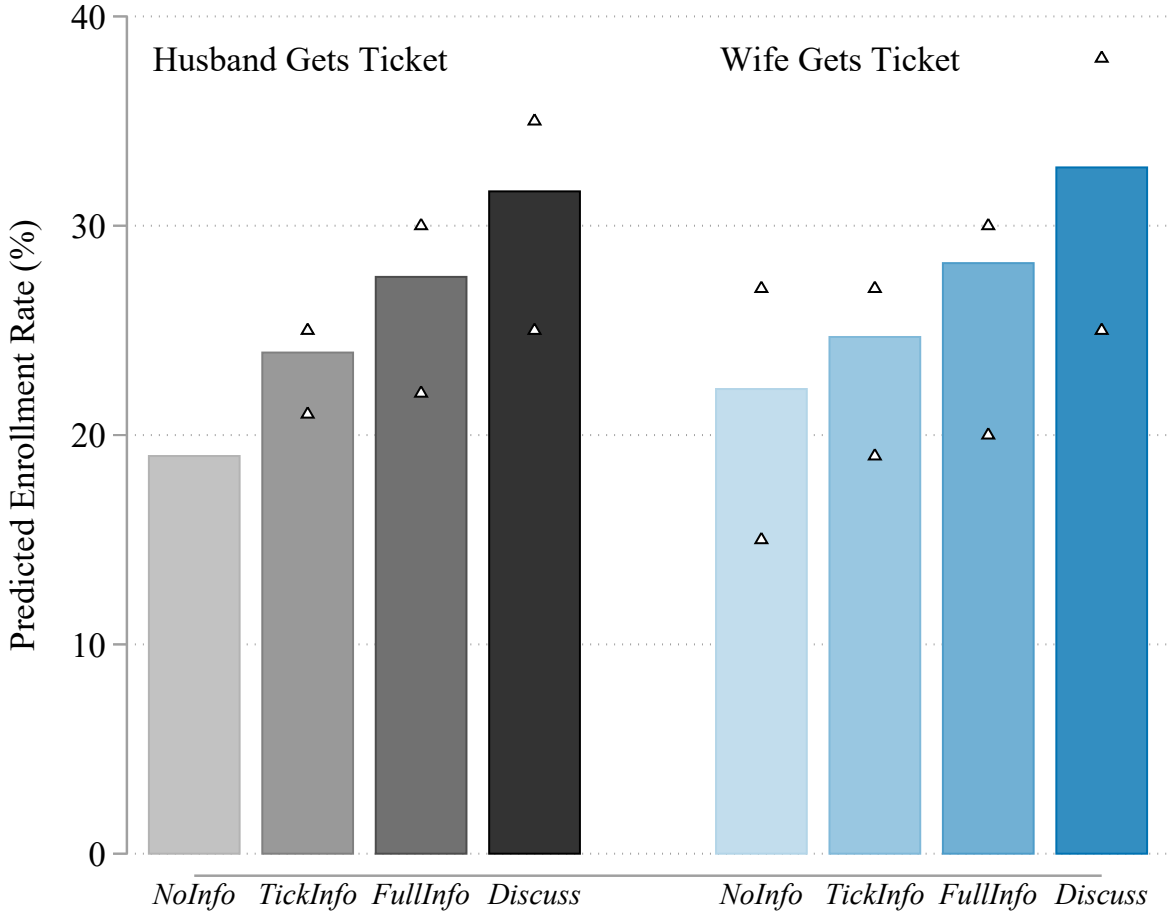


Notes: The figure shows the ten most common reasons for non-enrollment given by 334 husbands and 339 wives during the endline survey. Specifically, during the endline survey we asked wives and husbands of non-enrolling households: *What are the main reasons your household decided not to enroll in this 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, since these spouses did not have any say in the decision not to enroll.

Figure A4: An Enrollment Ticket

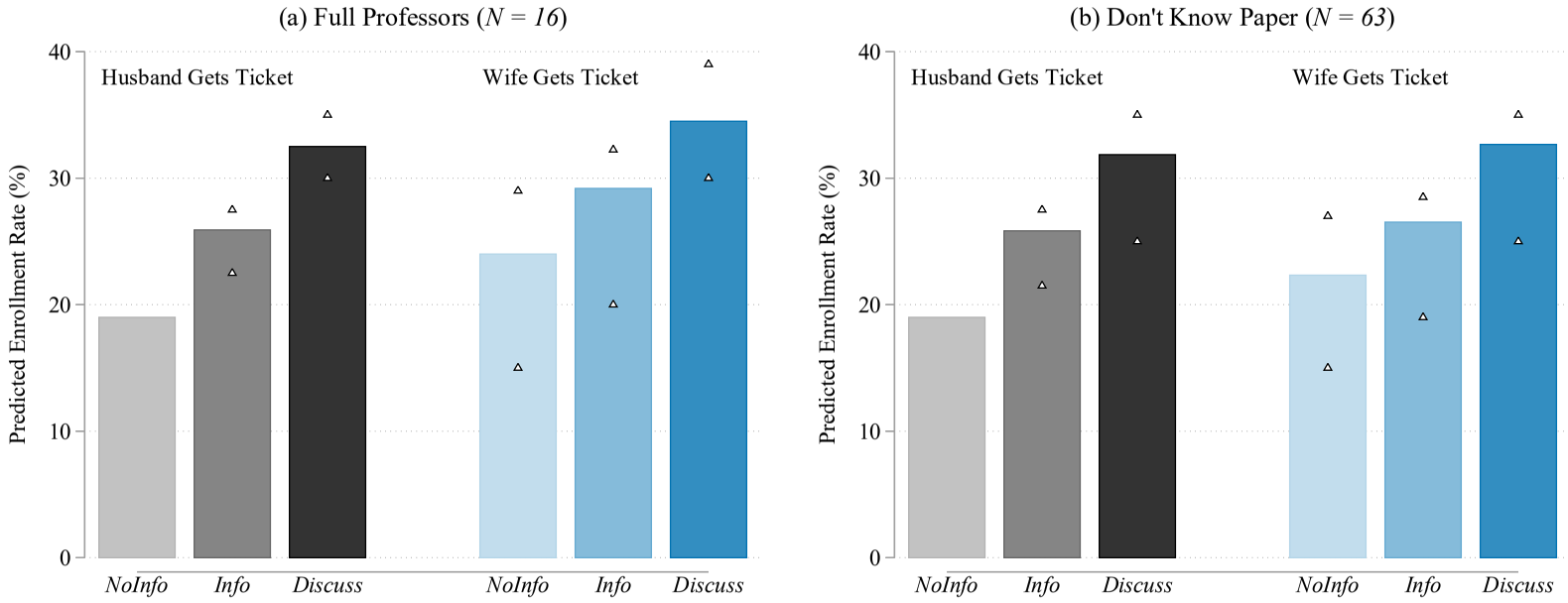
Notes: This is a photo of an enrollment ticket. The first line says "enrollment ticket." The second provides the names of the husband and wife to which this ticket corresponds. The names are blurred for confidentiality. The third line has the unique identification number we assigned to this couple. The 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 A5: Expert Predictions Including *Info* Sub-Treatments



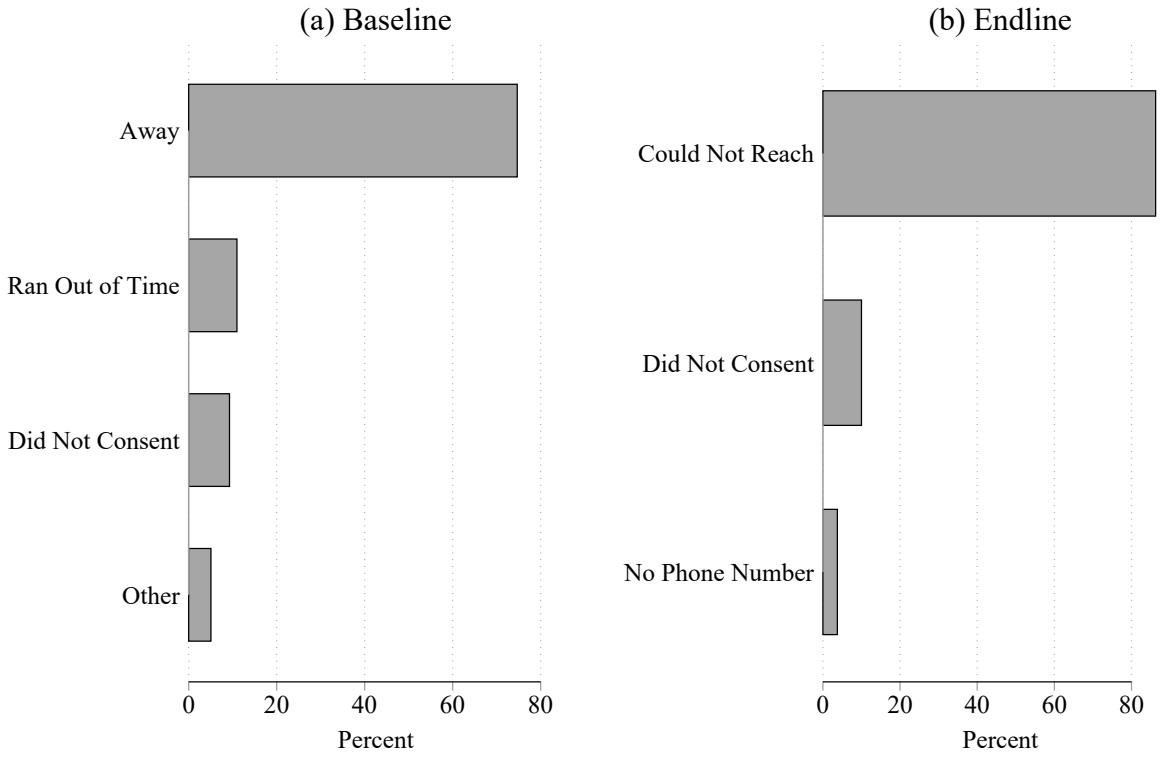
Notes: The figure shows the mean predicted enrollment rate for the full eight treatment cells according to our full sample of 70 experts on intra-household economics. Each expert was told the true enrollment rate (19%) of the treatment in which the husband gets the job ticket and the wife doesn't know, 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 A6: Robustness of Expert Predictions



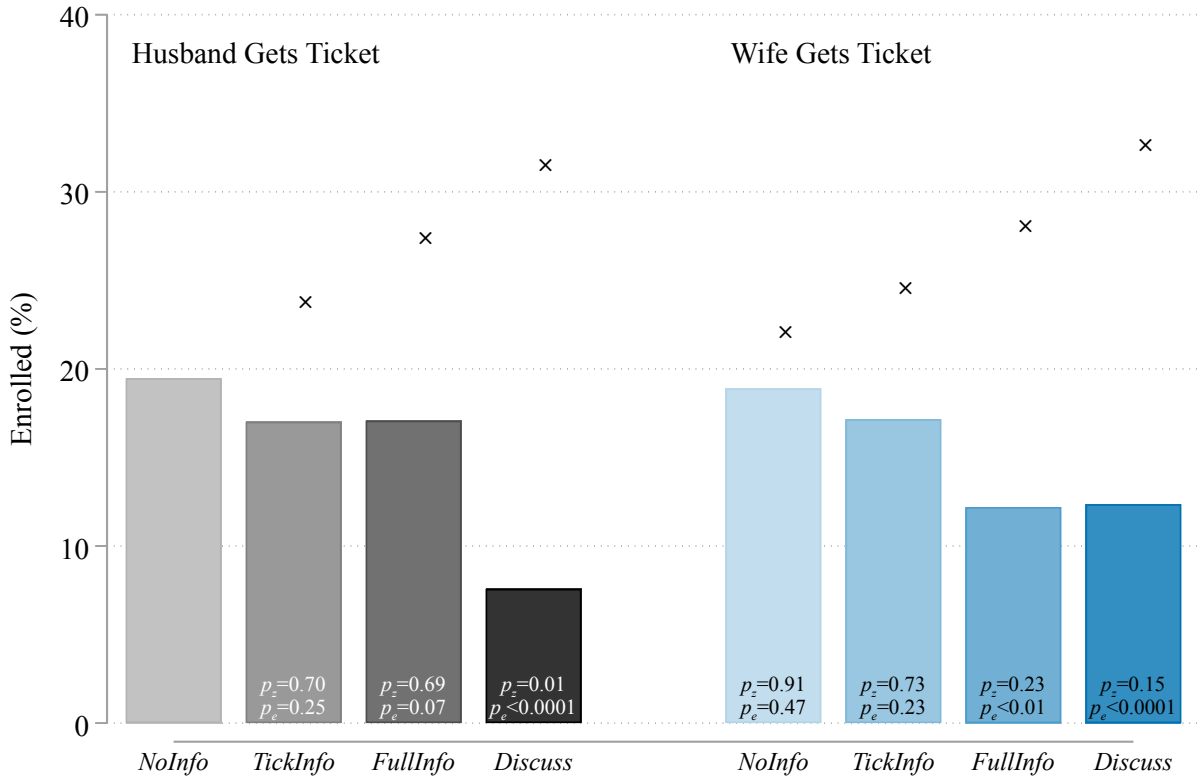
Notes: The figure shows the mean predicted enrollment rate for each treatment cell according to experts on intra-household economics. Panel (a) shows the predictions of the 16 experts who are Full Professors. Panel (b) shows the predictions of the 63 experts that answered *No* to the question *Have you 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 A7: 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 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, strata fixed effects, 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), and missingness dummies for each of these variables. 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.

Table A1: No Detectable Enrollment Effect of *Full Info* Versus *Ticket Info*

	Correct Answer About...					
	Pay	Pay Cut	Start Time	End Time	Job Desirability	Enrolled
	(1)	(2)	(3)	(4)	(5)	(6)
Full Info	0.33***	0.44***	0.29***	0.22***	0.30	-0.04
	(0.08)	(0.09)	(0.08)	(0.07)	(0.24)	(0.06)
Observations	143	143	128	128	78	171
Ticket Info Mean	0.51	0.36	0.52	0.61	0.85	0.18
p(*Full Info* = *Ticket Info*)	<0.01	<0.01	<0.01	<0.01	0.21	0.51
Strata	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. All columns include only those assigned to the *Info* treatment. Columns (1) to (5) include only the non-ticketed spouses, i.e. those affected by the two sub-treatments within *Info*. The unit of observation in these columns is the individual. For column (6), the unit of observation is the couple. The outcomes in columns (1) to (5) come from the endline survey. Columns (1) to (4) are dummy variables equal to one if the respondent knew the initial pay for the women's weaving program, the lower pay level from the second month if targets were not met, the start time of the working day, and the end time. 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. Enrolled is an indicator for the wife of a couple enrolling on enrollment day. The start and end time knowledge questions were not included for the first center's respondents, while the Job Desirability question was asked only for the final three centers. Couple-level controls included throughout 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 missingness dummies for each of these variables. *** p<0.01, ** p<0.05, * p<0.1.

Table A2: 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 A3: 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	0.41	-0.01	0.07	0.06*	0.09	0.03	-0.01	-0.02	0.01	0.02	-0.11	0.01
	(0.04)	(0.55)	(0.59)	(0.05)	(0.04)	(0.18)	(0.05)	(0.05)	(0.05)	(0.03)	(0.04)	(0.09)	(0.09)
Discuss	-0.02	0.77	0.44	0.03	0.03	0.19	0.00	0.03	-0.03	0.10**	0.04	-0.11	0.25**
	(0.04)	(0.57)	(0.62)	(0.05)	(0.04)	(0.17)	(0.06)	(0.05)	(0.05)	(0.04)	(0.04)	(0.09)	(0.10)
Wife Ticket	0.02	0.22	0.48	0.02	-0.04	-0.09	-0.04	-0.00	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
Joint p-value	0.44	0.60	0.73	0.57	0.18	0.66	0.72	0.88	0.65	0.04	0.81	0.40	0.03
Outcome Mean	.32	26	29	.39	.13	2.3	.4	.32	.28	.13	.82	1.1	.81
Strata	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. 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 A4: Wife's Past Employment Does Not Predict Enrollment

	Enrolled			
	(1)	(2)	(3)	(4)
Wife Employed	-0.005	0.008	-0.015	-0.002
	(0.047)	(0.048)	(0.046)	(0.046)
Observations	495	495	495	495
Outcome Mean	0.15	0.15	0.15	0.15
Strata	No	Yes	Yes	Yes
Controls	No	No	Yes	Yes
Treatments	No	No	Yes	Yes

Notes: Robust standard errors in parentheses. Enrolled is an indicator for enrolling on enrollment day. Wife Employed is an indicator for whether the wife has worked for income in the last three months. All columns include an indicator for Wife Employed being missing (five observations). 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), and missingness dummies for each of these variables. Treatments are indicators for *Info*, *Discuss*, and Wife Gets Ticket. *** p<0.01, ** p<0.05, * p<0.1.

Table A5: 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	0.25	-0.07	0.03	0.07*	0.08	0.02	-0.02	0.01	0.02	0.01	-0.16	-0.03
	(0.04)	(0.04)	(0.59)	(0.64)	(0.06)	(0.04)	(0.20)	(0.06)	(0.05)	(0.06)	(0.04)	(0.05)	(0.10)	(0.10)
Discuss	0.05	0.02	0.91	0.33	-0.00	0.03	0.19	-0.02	0.05	-0.03	0.10**	0.07	-0.07	0.24**
	(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)
Wife Ticket	0.01	0.01	0.25	0.28	0.04	-0.05	-0.05	-0.02	-0.02	0.04	0.07**	-0.01	0.10	-0.03
	(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
Joint p-value	0.13	0.46	0.54	0.92	0.76	0.16	0.79	0.89	0.58	0.69	0.04	0.41	0.25	0.04
Outcome Mean	.17	.16	26	29	.4	.13	2.3	.4	.32	.28	.12	.82	1.2	.83
Strata	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. 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 A6: Discussion Effect by Village

	Enrolled					
	(1)	(2)	(3)	(4)	(5)	(6)
Discuss	-0.15	-0.24	0.01	-0.05	-0.12**	-0.11
	(0.15)	(0.20)	(0.04)	(0.09)	(0.06)	(0.09)
Observations	64	52	90	88	110	91
Omitted Group Mean	0.48	0.33	0.00	0.11	0.14	0.21
Strata	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Info	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. Enrolled is an indicator for enrolling on enrollment day. Each column includes only couples from one of the six villages. 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 missingness dummies for each of these variables. *** p<0.01, ** p<0.05, * p<0.1.

Table A7: Heterogeneity by Non-Ticketed Being Less Supportive Does Not Reflect Heterogeneity by Individual Preferences

	Enrolled	
	(1)	(2)
Info	-0.02	-0.03
	(0.04)	(0.05)
Discuss	-0.07*	-0.09**
	(0.04)	(0.04)
Info × Wife's Weaving Preference (0 to 2)	0.02	0.02
	(0.04)	(0.04)
Discuss × Wife's Weaving Preference (0 to 2)	0.01	0.02
	(0.04)	(0.04)
Info × Husband's Weaving Preference (0 to 2)	-0.02	-0.04
	(0.05)	(0.05)
Discuss × Husband's Weaving Preference (0 to 2)	-0.04	-0.04
	(0.05)	(0.05)
Observations	489	489
p(*Info* × Wife's Weaving Pref. = *Discuss* × Wife's Weaving Pref.)	.83	.96
p(*Info* × Husband's Weaving Pref. = *Discuss* × Husband's Weaving Pref.)	.75	.95
Strata	Yes	Yes
Wife's and Husband's Weaving Preference Level Variables	Yes	Yes
Controls	No	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). 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, and missingness dummies for each of these variables. *** p<0.01, ** p<0.05, * p<0.1.

Table A8: Who Had Most Influence over the Enrollment Decision?

	Me	Equal Influence	My Spouse
	(1)	(2)	(3)
Info	-0.02	0.05	-0.02
	(0.06)	(0.06)	(0.06)
Discuss	-0.05	0.05	-0.00
	(0.06)	(0.06)	(0.06)
Observations	407	407	407
Omitted Group Mean	.33	.31	.36
Strata	Yes	Yes	Yes
Controls	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 whether the respondent said *Me*, Equal Influence is an indicator for whether the respondent said *Me and my spouse had same influence*, and My Spouse is an indicator for whether the respondent said *My spouse*. The sample includes only ticketed spouses, as non-ticketed spouses in *NoInfo* that did not know about the ticket (a selected sample) were not asked the question. 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 missingness dummies for each of these variables. *** p<0.01, ** p<0.05, * p<0.1.

Table A9: Discussion Effect not Driven by Argument

	Spousal Disagreement		Spousal Inconsiderateness	
	(1)	(2)	(3)	(4)
			(a) Wives	
Info	-0.03	-0.06	-0.28**	-0.30**
	(0.12)	(0.12)	(0.13)	(0.13)
Discuss	-0.04	-0.05	-0.28**	-0.27**
	(0.12)	(0.12)	(0.13)	(0.13)
Observations	240	240	217	217
Omitted Group Mean	1.3	1.3	1	1
p(*Discuss* = *Info*)	.89	.92	.96	.83
			(b) Husbands	
Info	0.04	0.02	-0.03	-0.05
	(0.13)	(0.13)	(0.10)	(0.11)
Discuss	0.08	0.17	-0.17*	-0.12
	(0.13)	(0.13)	(0.09)	(0.09)
Observations	239	239	223	223
Omitted Group Mean	1.1	1.1	.39	.39
p(*Discuss* = *Info*)	.77	.27	.12	.45
Strata	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes
Obs. Unit	Individual	Individual	Individual	Individual

Notes: Robust standard errors in parentheses. Spousal Disagreement is the level of reported disagreement between spouses about the job opportunity, measured on a 0-2 scale. Spousal Inconsiderateness is the extent to which one's spouse was inconsiderate of one's own opinion about the job opportunity, measured on a 0-2 scale. The two questions were asked only for the final three centers. Panel (a) uses only the reports of wives, and Panel (b) uses only the reports of 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, (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 missingness dummies for each of these variables. *** p<0.01, ** p<0.05, * p<0.1.

Table A10: Discussion Effect not Driven by Adherence to Norms around Female Employment

	Enrolled	
	(1)	(2)
Info	-0.06	-0.04
	(0.06)	(0.06)
Discuss	-0.09	-0.09*
	(0.06)	(0.05)
Info × High Female Employment Village	0.06	0.02
	(0.08)	(0.07)
Discuss × High Female Employment Village	0.00	-0.05
	(0.07)	(0.07)
Observations	490	490
p(*Info* × High Female Emp. Vil. = *Discuss* × High Female Emp. Vil.)	.45	.3
Strata	Yes	Yes
Wife Employed	Yes	Yes
Info × Wife Employed	Yes	Yes
Discuss × Wife Employed	Yes	Yes
Controls	No	Yes

Notes: Robust standard errors in parentheses. Enrolled is an indicator for enrolling on enrollment day. Wife Employed is an indicator for whether the wife has worked for income in the last three months. High Female Employment Village is an indicator for living in a village with an above median fraction of wives employed. 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), and missingness dummies for each of these variables. *** p<0.01, ** p<0.05, * p<0.1.

Table A11: Discussion Effect not Driven by Deeper Deliberation

	Number Of Discussions		Enrolled		Enrolled And Stayed 1 Month	
	(1)	(2)	(3)	(4)	(5)	(6)
Info	0.05	0.17			0.02	0.02
	(0.36)	(0.36)			(0.03)	(0.03)
Discuss	0.98**	1.02**			-0.04	-0.05*
	(0.40)	(0.43)			(0.03)	(0.03)
Days Before Enrollment			0.01	0.02		
			(0.01)	(0.01)		
Observations	829	829	330	330	495	495
Omitted Group Mean	3	3			.094	.094
Outcome Mean			.17	.17		
p(*Discuss* = *Info*)	.0066	.018			.033	.012
Strata	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes
Wife Indicator	Yes	Yes	No	No	No	No
Sample	All	All	Exclude *Discuss*		All	All

Notes: Standard errors in parentheses clustered at couple-level for columns (1) and (2), otherwise robust standard errors. Regression is at the individual-level for columns (1) and (2), otherwise couple-level. Number Of Discussions is the endline-reported number of discussions spouses had about the job opportunity. Days Before Enrollment is the number of days in between ticket receipt and the enrollment day. Enrolled is an indicator for enrolling on enrollment day. Enrolled within 1 month is an indicator for enrolling within 1 month of enrollment day. 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 missingness dummies for each of these variables. *** p<0.01, ** p<0.05, * p<0.1.

Table A12: The Predictions of Experienced Researchers Are Less Accurate

	Confidence (0 to 7)			Mean Prediction Accuracy		
	(1)	(2)	(3)	(4)	(5)	(6)
Associate Professor	.625			-.887		
	(.636)			(1.84)		
Full Professor	-.458			-3.96**		
	(.445)			(1.76)		
Other Position	.0694			-6.45*		
	(.542)			(3.29)		
Has Google Scholar Profile		-.712	-.606		-8.99	-11.1
		(1.55)	(1.61)		(8.46)	(8.15)
Google Scholar Citation Ranking (1 to 61)		.00014	-.00103		.0964*	.124**
		(.0111)	(.0117)		(.0527)	(.0612)
Above-Median Time on Survey			.143			-3.46
			(.414)			(2.63)
Confidence						.65
						(.514)
Observations			70			
Outcome Mean		2.73			-13.94	

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 Theory Details

B.1 Micro-founding the net utility parameters

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

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

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:

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

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:

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

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:

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

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

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

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 $0 \leq h_i \leq 1$.[38] Substituting in to the budget constraint and simplifying, we can solve for total earned

[38] 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

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:

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

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

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

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 Section 4 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).

$h'_w > 0$ 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.

B.2 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 4) 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 $v_w < 0, v_h < c, \beta v_h + (1-\beta) v_w > 0$:

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

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

C Expert Survey

See next page.

Lowe & McKelway -- Expert Survey

Household Decision-Making and Women's Labor Supply in India: An Expert Survey

This survey asks you to predict the results of an experiment on household decision-making about female labor supply in India conducted by Matt Lowe (UBC) and Madeline McKelway (Stanford). We selected you for this survey based on your expert knowledge of the economics of the household.

The survey takes about 5 to 8 minutes to complete. Your participation is voluntary. Your individual responses will remain confidential. Upon completion of data collection, we will send respondents our completed manuscript, including results from this survey.

We will not ask for any identifying information in this survey. However, we have previously collected information (e.g., name, institution) that was publicly available online. So, we will be able to link your answers with previously collected identifying information. Your data will be de-identified before being analyzed, with only the PIs having access to the full data. Data will be stored on Dropbox in the US and Canada in the longer term, so your information is subject to the US Patriot Act. Upon publication, only de-identified information will be shared if needed.

Feel free to contact the authors with any questions or comments (matt.lowe@ubc.ca, mckelway@stanford.edu). If you have any concerns or complaints about your rights as a research participant and/or your experiences while participating in this study, contact the Research Participant Complaint Line in the UBC Office of Research Ethics at 604-822-8598 or, if long distance, e-mail RSIL@ors.ubc.ca or call toll free 1-877-822-8598.

Would you like to participate in the survey?

- ○ Yes
- ○ No

This survey will ask you to predict the results of an experiment. The experiment was conducted in 2016 and the results have been circulated. We want only those who have not seen or heard the experimental results to make predictions.

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 (Stanford)?

Note: if you need more information about the experiment to answer, select “don’t know”; we will ask this question again after providing details on the experiment.

- ○ Yes, and I remember at least some of the results
- ○ Yes, but I do not remember any of the results at all
- ○ No

○ Don't Know

Study Background: Weaving Program for Women in Uttar Pradesh

India's female labor force participation rate (FLFP), at 20%, is one of the lowest in the world, and substantially lower than the 46% average across low- and middle-income countries (World Bank, 2020). Intra-household constraints may help explain this: husbands in India have a great deal of control over their wives' labor supply and report more conservative views about women's work than wives (Bernhardt et al. 2018; Field et al. 2021).

We conducted an experiment on household decision-making about female labor supply in rural Uttar Pradesh, India. Uttar Pradesh is one of India's poorest states and one that adheres to particularly regressive gender norms.

We partnered with one of India's largest carpet producers. The firm employs many male carpet weavers in this area, but recently began a program to train and employ women. The program offers four months of paid training in weaving followed by employment. It occurs in all-female weaving centers located in villages. Each center recruits women from its village and has space for 20 women.

Figure 1: A Female Weaver

We conducted our experiment in conjunction with the opening of six new women's weaving centers. Our survey team went door-to-door in the six villages to identify women eligible for the program and for our experiment. Any women who were married, aged 18-30, available to participate in the study, and whose husbands were also available were included in the experiment. Just 13% of women in our sample were working for income at baseline. Husbands in our sample reported less support than their wives for women's employment, both in general and in the firm's program.

Experimental Design

We conducted baseline surveys with each husband and each wife separately and in private. Our experiment randomly varied what each spouse was told about the firm's program after completing the survey. Our outcome of interest is the wife's enrollment in the firm's program.

We first printed enrollment tickets for 90% of couples, chosen randomly. The experiment was conducted amongst these 495 couples for whom tickets were printed. Each ticket had the name of the couple and a unique identification number. Women could not enroll in the program without a ticket and only the woman for whom a particular ticket was printed could enroll using the ticket. **The delivery of the ticket and job information varied according to two, cross-randomized treatments, visualized in the figure below. The first determined which spouse received the ticket.** The spouse who received the ticket was also given details on the program and enrollment process, and told what his or her spouse would be told by a surveyor.

The second treatment determined what information was given to the non-ticketed spouse. In the first condition, he or she received no information; in the second, he or she was informed that a ticket had been printed for the wife, and had been or would be given to his or her spouse; in the third, the non-ticketed spouse was given this information plus details on the program and enrollment process; finally, in the fourth condition, the two spouses were given details on the program and enrollment process together, the two were then given three minutes to

discuss the opportunity (this discussion occurred in the presence of surveyors who remained silent during the conversation), and just after the discussion, the ticket was given to the ticketed spouse.

	Ticketed Spouse:	
Non-Ticketed Spouse:	**Husband** gets ticket and job details	**Wife** gets ticket and job details
No Info.		
Ticket Info.		
Full info.		
Discussion		

Figure 2: Experimental Design

Crucially, we designed the experiment so that **ticketed spouses in the "No Info" condition had plausible deniability, meaning they could withhold the information that there was a work opportunity for the wife without fear of getting caught**. We did this by *not* printing tickets for 10% of couples. In the remaining 90%, any spouse who was assigned to receive information about the ticket was also told that some couples would be surveyed but not receive a ticket. Thus, if only one spouse was informed about the ticket, he or she could plausibly deny its existence.

Program Enrollment

Women could enroll in the program by going to their local weaving center on enrollment day. They were required to present their enrollment ticket and to come with their husbands. The requirement that husbands attend ensured women could not enroll without their husbands' support, and makes enrollment a joint household decision. Enrollment occurred 1-8 days after tickets and information had been given.

As mentioned previously, the experiment was conducted in 2016 and the results have been circulated. We want only those who have not seen or heard the experimental results to make predictions. We asked previously about your awareness of our results, but it is possible your answer has changed after reading more about the experiment, so we ask again.

Have you seen or heard results from this experiment?

- O Yes, and I remember at least some of the results
- O Yes, but I do not remember any of the results at all
- O No
- O Don't Know

Predicting the Results

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?**

	Husband gets ticket and job details	Wife gets ticket and job details
Non-Ticketed Spouse has No Information	19	

	Husband gets ticket and job details	Wife gets ticket and job details
Non-Ticketed Spouse has Ticket Information		
Non-Ticketed Spouse has Full Information		
Spouses have Discussion		

Confidence in Predictions

How many of your seven predictions do you expect are 5 percentage points or less from the actual enrollment rate in each treatment group?

- 0
- 1
- 2
- 3
- 4
- 5
- 6
- 7