Learning in the Household*

John J. Conlon† Malavika Mani Gautam Rao
Matthew Ridley Frank Schilbach

May 25, 2021

Abstract

We study social learning between spouses using an experiment in Chennai, India. We vary whether individuals discover information themselves or must instead learn what their spouse discovered via a discussion. Women treat their 'own' and their husband's information the same. In sharp contrast, men's beliefs respond less than half as much to information that was discovered by their wife. This is not due to a lack of communication: husbands put less weight on their wife's signals even when perfectly informed of them. In a second experiment, when paired with mixed- and same-gender strangers, *both* men and women heavily discount their teammate's information relative to their own. We conclude that people have a tendency to underweight others' information relative to their own. The marital context creates a countervailing force for women, resulting in a gender difference in learning (only) in the household.

Keywords: information, gender, communication, social learning, experiment

*We thank workshop audiences at ASSA, Berkeley, Harvard, MIT, Northeastern, UNC, Warwick, Zurich and the Workshop on Subjective Beliefs for helpful comments and feedback. We also thank Nava Ashraf, Ned Augenblick, Oriana Bandiera, Rafael Batista, Leonardo Bursztyn, Arun Chandrasekhar, Katie Coffman, Stefano DellaVigna, Ben Golub, David Laibson, Shengwu Li, Ulrike Malmendier, Michel Maréchal, Sendhil Mullainathan, Tommy O'Donnell, Ricardo Perez-Truglia, Matthew Rabin, Chris Roth, Dmitry Taubinsky and David Yanagizawa-Drott for helpful suggestions and comments. We thank Sangeetha Ramanathan and the entire team at the Behavioral Development Lab in Chennai for excellent research assistance. We thank all our study participants for their time and patience. This project was funded by the Wellspring Philanthropic Fund. We received IRB approval from MIT, protocol #1810538700. The experiment was pre-registered on the AEA registry, number AEARCTR-0004253.

†Conlon: Harvard (johnconlon@g.harvard.edu); Mani: Columbia (mm5819@columbia.edu); Rao: Harvard and NBER (grao@fas.harvard.edu); Ridley: MIT (mridley@mit.edu); Schilbach: MIT and NBER (fschilb@mit.edu)

1 Introduction

Household members frequently have access to independent information and have many opportunities to share information with each other. The household could thus be a significant venue of social learning in the economy. Recent work has shown that strategic motives can obstruct the flow of information between spouses (Ashraf et al., 2020). In many situations, however, spouses have *common* objectives: to wisely invest their money, to consult a competent doctor, or to send their child to a good school. Making good decisions in these cases requires spouses to pool information. Yet, we have little understanding of how well spouses learn from each other, whether such learning differs by gender, and what psychological mechanisms pose barriers to learning in the household and beyond.

We design experiments to measure how well spouses learn from each other when incentives are aligned. We recruit 400 married couples and 500 strangers in Chennai, India, to participate in a social-learning task. We test (i) whether people respond similarly to information uncovered by themselves and by their spouse; (ii) how this varies by gender; (iii) whether inefficient learning is due to a lack of communication or because communicated information is used incorrectly; and (iv) whether spouses learn differently from each other than mixed- and same-gender strangers working in teams.

We randomize across experimental rounds whether participants learn information entirely on their own or partially via their spouse. The goal in each round is to guess the share of red balls in an urn. Before making their guesses, participants receive independent, noisy signals in the form of two sets of draws from the urn. In the *Individual* round, participants privately draw both sets of signals themselves and play the game alone. Men and women are equally good at the task in this condition and have similar levels of confidence. In the *Discussion* round, each spouse instead privately draws only one set of signals. The couple then have a face-to-face discussion—giving them a chance to pool information—after which each spouse makes a private guess. One guess made in the experiment is randomly chosen to be paid off based on its accuracy, and the payoff is split equally between the spouses.

Our three empirical approaches—non-parametric, reduced-form, structural—impose different assumptions but yield consistent results. We focus here on the reduced-form approach, which simply asks how much the average guess changes in response to an additional red (as opposed to white) draw—we call this the "weight" placed on signals.

Information pooling implies equal weights on signals drawn oneself and by one's spouse, since the order and number of draws are held constant by design.[1]

We first compare guesses in the *Individual* and *Discussion* rounds, played in randomized order. Husbands put 58 percent less weight ($p<0.01$) on information their wives gathered—available to them via discussion—than on information they gathered themselves. In contrast, wives barely discount their husband's information (by 7 percent), and we cannot reject that wives treat their husband's information like their own ($p=0.61$). The difference in husbands' and wives' discounting of each other's information is statistically significant ($p=0.02$).

The lower weight husbands place on their wives' information is not due of a lack of communication from wives to husbands. In another experimental treatment—the *Draw-sharing* round—husbands put less weight on their wife's information even when it is *directly* conveyed to them by the experimenter (absent any discussion). In this case, husbands discount information collected by their wives by a striking 98 percent compared to information collected by themselves ($p<0.01$), while wives again treat their spouses' information nearly identically to their own. Lack of communication between spouses or husbands' mistrust of (say) wives' memory or ability thus cannot explain husbands' behavior. Rather, husbands treat information their wives gathered as innately less informative than information they gathered themselves. In contrast, wives treat their own and their husbands' information equally.

Husbands' discounting of their spouse's information is costly. When their wife randomly receives more draws than them, husbands' guesses in the *Discussion* and *Draw-sharing* rounds earn in expectation 9 percent (or 0.3 standard deviations) less than in the *Individual* round ($p=0.03$). For comparison, an additional year of education is associated with a 2 percent increase in expected earnings. Wives instead perform similarly in both types of rounds.

To examine whether the gender difference documented above extends beyond married couples, we conduct a second experiment with 500 adults who played an identical task in pairs of strangers. In both mixed- and same-gender pairs, men and women *both* respond more strongly to their own information than to their teammate's. Thus, the underweighting of others' information appears to be a more general phenomenon.

[1] Of course, joint deliberation may also help people better process the information and thus change the weight placed on each signal. Other experimental treatments, described below, remove the possibility of joint deliberation.

Husbands treat their wives (information) as they treat strangers; wives instead put more weight on their husband's information than on strangers' information. This difference between husbands and wives, and between spouses and pairs of strangers, is not explained by differences in observable characteristics such as the players' relative age, marital status, relative ability, or confidence.

Our findings cannot be explained by factors such as confusion, risk aversion, social signaling, reputational concerns or competitive behavior. Since the underweighting of others' information occurs even when the information is perfectly communicated by the experimenter, it also cannot be driven by differences in ability or confidence, propensity to contribute information (Coffman, 2014), or reluctance to ask for or provide information (Chandrasekhar et al., 2018; Banerjee et al., 2018). To establish the limits of our findings, we conducted a smaller third experiment with pairs of strangers, using variants of the *Draw-sharing* condition. We find significant evidence of underweighting of others' information with (i) 50 percent higher stakes, and (ii) when participants directly observe their teammate drawing their signals, in person.

We discuss possible interpretations of the underweighting of others' information even when perfectly communicated. First, it could be a form of egocentric bias (Ross et al., 1977) or an ownership effect wherein people consider their own information to be innately more valuable (Kahneman et al., 1991) or worthy of attention (Hartzmark et al., 2021). Second, information from personal experience may be more vivid than information conveyed by others (Malmendier and Nagel, 2011; D'Acunto et al., 2021). Third, participants may be misapplying an otherwise-reasonable heuristic, if one's own information is usually far more precise and relevant than information provided by others.

We cannot say exactly why wives and husbands behave differently, except that these patterns are not due to gender differences *per se* (e.g. differences in self-confidence, assertiveness, or competitiveness; see Niederle and Vesterlund 2011; Exley and Kessler 2021; Exley et al. 2020), since men and women treat strangers similarly. Instead, the marital context itself appears to generate gender differences in behavior, for instance due to norms of wives deferring to their husbands or repeated experiences of husbands possessing better information.

Our study contributes to a large literature on household decision-making, particularly in developing countries. Standard models of household decision-making assume that spouses have identical beliefs but may have differing preferences (see Pollak 2019 for a review). Several recent papers have relaxed the assumption of perfect information pooling, exploring situations where one spouse may have strategic reasons to hide information from the other (e.g. Ashraf 2009; Ashraf et al. 2014; Ambler 2015; Lowe and Mckelway 2019). Others have measured information asymmetries (Afzal et al., 2018) and tested whether information is transmitted within the household (e.g. Fehr et al. 2019; Ashraf et al. 2020; Apedo-Amah et al. 2020). To our knowledge, ours is the first to study how couples pool information when their incentives are fully aligned. Our results suggest a gender asymmetry in learning in the household. If this holds more generally, we would expect to find lower 'pass-through' of information from wives to husbands than vice versa, at least in contexts similar to those we study.

Second, our study adds to the literature on the role of gender in group judgments and decision-making.[2] Existing evidence shows that women are less likely to contribute their ideas, particularly in stereotypically-male tasks (Coffman, 2014; Cooper and Kagel, 2016) and in mixed-gender groups (Bordalo et al., 2019; Chen and Houser, 2017). When women do contribute information, they are often perceived as less competent or worse communicators, even conditional on ability (Beaman and Dillon, 2018; Coffman et al., 2021a; Mengel et al., 2019). We show that—in a gender-neutral task—gender differences in learning can emerge in the household context, even when there are no gender differences outside it. This echoes Abbink et al. (2020), who show that women in Bangladesh are more likely than men to delegate risky lab decisions to their spouse, despite no gender differences when playing with strangers.

Third, we contribute to the literature on social learning. Mobius et al. (2015) describe two classes of barriers to social learning: diffusion (whether private information reaches others) and aggregation (how individuals weigh information). We provide new evidence for a potentially far-reaching bias in information aggregation which may hinder social learning: individuals treat data gathered by others as much less informative than data gathered themselves. Many lab studies of observational learning find that people put more weight on private information than on what can rationally be inferred from the

[2]There is a long literature in psychology and a smaller one in economics on how people respond to advice (see Bonaccio and Dalal 2006 for a review) and on whether group or individual decision-making tends to produce more accurate judgments (see, e.g., Cooper and Kagel 2005; Minson et al. 2018). A related strand of this literature finds that group discussions and resulting judgments tend to focus on shared rather than private information (Stasser and Titus, 1985; Gigone and Hastie, 1993). Another shows that individuals put more weight on their own initial opinions than on those of their conversation partner or advisors (e.g., Bonaccio and Dalal 2006; Minson et al. 2011; Liberman et al. 2012).

actions of others (e.g. Weizsäcker 2010).[3] Though these findings are consistent with our interpretation, they are also consistent with other (sometimes rational) explanations, such as mistrust of others' ability (De Filippis et al., 2017), overconfidence (Angrisani et al., 2018), attenuation bias from imperfect measurement of social networks (Mobius and Rosenblat, 2014), altruism (March and Ziegelmeyer, 2020), base-rate neglect (Benjamin et al., 2019), or other behavioral biases (Guarino and Jehiel, 2013). Our finding that agents down-weight others' *signals*, not just their actions or beliefs, along with other features of our experiment, rules out these and other explanations.[4] Rather, people just treat their own information differently.

2 Setting, Recruitment, and Study Sample

2.1 Setting, recruitment, and screening

All experimental sessions were conducted in our lab in Chennai, India. The sessions with couples took place between June and November 2019, and the sessions with strangers were conducted between July and December 2019. We recruited participants on a rolling basis, with about 2 to 5 pairs of people completing the experiment on a given day. Recruitment stopped when we reached our pre-specified target of 400 couples and 500 unrelated individuals who completed the experiment.

Recruitment of couples. We recruited couples from low- to middle-income communities within a reasonable travel time of the lab. Surveyors went door-to-door to advertise an academic study on 'how decisions are made in the household'. Potential participants were informed that they would spend 2-3 hours at the study office and could expect to earn Rs. 300-560 ($4-7.75) per couple, plus a payment of Rs. 100 ($1.40) to cover travel expenses. Participants were required to be 'married couples' who could

[3]See Stone and Zafar (2014) for an example of similar underweighting of public information in the context of sports rankings.

[4]Drehmann et al. (2005) include in their online experiment a treatment similar to Anderson and Holt (1997) but where subjects can see previous decision-makers' signals (a or b) as well as their choice of urns (A or B). Using their data, we find (analyses not shown) that, even conditional on the total number of a and b signals available to the agent (her signal and all previous players' signals), she is much more likely to choose urn A if "her" signal was a than if "her" signal was b. These results, while suggestive of and consistent with our finding of intrinsic under-weighting of "others" compared to "own"' information, cannot rule out that subjects put more weight on their own information because they receive it last, as models of base-rate neglect would in fact predict.

come to the lab together.[5] No more specific information was provided at this point.

In total, 419 couples came to our lab. 11 couples were screened out due to low understanding of the task, as evidenced by failure to answer basic comprehension questions correctly. 8 couples dropped out mid-way through the experiment for other reasons. This left us with the target sample of 400 couples.

Recruitment of strangers. Our non-couples or 'strangers' sample participated in a separate round of experimental sessions, where they played in both mixed- and same-gender pairs. We recruited individuals unknown to each other prior to our study from the same neighborhoods as the couples sample. Recruiters followed the same procedure as for the couples sample, with the exception that participants were recruited individually. In total, 508 individuals (254 men and 254 women) were enrolled. Of these, 4 men and 4 women were excluded before starting the experiment as they either did not understand the task (as measured by comprehension questions) and/or lacked sufficient numeracy. This left 500 individuals (250 men and 250 women) who form our non-couple sample.

2.2 Descriptive statistics

The first two columns of Table 1 describe the couples in our study. The average husband is about 36 years old, has been married for 12 years, and has eight years of education. The average wife is about four years younger, but has the same years of education. There is substantial heterogeneity across individuals, as indicated by the large standard deviations. Eighty-five percent of participants self-report that they are literate (in Tamil, the local language), and 41 percent are numerate (as measured by correctly answering "Can you tell me what 3×9 is?"). Husbands are somewhat more numerate than wives and more likely to work outside the home, work more hours, and earn more conditional on working. These gender differences in labor-force participation are broadly in line with patterns in our study setting.

Our secondary sample of strangers (Table 1, cols 3 and 4) is similar to the couples on demographics except that they also include single adults of both genders (although

[5]In practice, by 'married couples' we mean cohabiting couples who identified as married. Given the cultural context, it was inappropriate to ask people to reveal other forms of romantic relationships. It is also uncommon in our study context for same-sex couples to publicly identify as such, and thus our sample consists entirely of couples consisting of one man and one woman.

a majority are married). They have similar education levels, ages, literacy, numeracy and labor-market participation as the couples.

3 Experimental Design

Figure 1(a) illustrates the overall structure of the experiment. Participants play five rounds—with different treatments—of a balls-and-urns task building on a large literature studying individual learning (Benjamin, 2019). The goal in each round is to guess the number of red balls in an urn containing 20 red and white balls. Participants are informed that the number of red balls is drawn uniformly from 4 to 16 in each round. We explain this distribution with the help of the scale in Figure A.I(a).[6]

In each round, participants receive independent signals about the composition of the urn. Concretely, they privately draw balls from the urn with replacement. Depending on the round, they either play the game entirely on their own or else can learn some of the signals from their teammate. Comparing learning across these rounds allows us to test for frictions in communication and information-processing which may interfere with social learning.

The number of draws in each set of signals is randomized—either 1, 5 or 9 draws—creating variation in how informed each participant is.[7] After receiving each set of signals (or potentially learning them through a discussion with their teammate), participants enter a private guess about the number of red balls in the urn. These guesses are incentivized: the closer their guess to the truth, the more participants can expect to be paid. The incentives of both teammates are aligned: one of the many guesses made by either participant in the pair is randomly picked to be paid out, and the earnings are split equally between the teammates (more details on incentives in section 3.2 below).

[6] We avoided more extreme distributions—fewer than 4 or over 16 red balls out of 20—as these were more likely to generate sets of signals with complete agreement between the teammates.

[7] To be precise, we randomly choose the number of draws in the two sets of signals received in each round with uniform probability from $\{(1,1),(1,5),(5,1),(5,5),(1,9),(9,1)\}$. We exclude cases with more than 10 draws total.

3.1 Experimental rounds

Figure 1(b) shows the structure of each round. Participants first play, in randomized order, an *Individual* round and a *Discussion* round.

***Individual* round.** The *Individual* round proceeds as follows. First, the participant draws a set of balls from the urn, followed by a guess of how many red balls are in the urn. Then, they draw a second set of balls from the urn and make a second (and final) guess. All drawing and guessing is done privately, without any opportunity to share information. This round serves as a control condition—a benchmark against which we compare the other conditions.

***Discussion* round.** The *Discussion* round differs from the *Individual* round in that, for each participant, their *teammate's* draws—accessible through a discussion—serve as their 'second' set of draws. Each person first makes one set of draws followed by a private guess, exactly as in the *Individual* round. Next, the couple are asked to hold a face-to-face discussion and decide on a joint guess. After their discussion and joint guess, each person makes one final, private guess. This final private guess will be compared to the final private guess in the individual round. We will also briefly examine the joint guess.

Couples can take as long as they like for the unstructured, face-to-face discussion. They are aware that they will enter a joint guess after the discussion and will then each have a chance to make a private guess. Thus, they have an incentive to pool information with their teammate. They also have an incentive to help their teammate deliberate and make better guesses conditional on information, as in Cooper and Kagel (2005, 2016). We record the audio of the discussion (with participants' consent) and later analyze the transcripts, as reported in Table A.I. We also randomize whether the experimenter is present for the discussion.

Comparing the *Individual* and *Discussion* rounds. We compare each participant's final private guesses in the *Individual* and *Discussion* rounds. By design, participants have access to the exact same number of draws to inform their final private guess in these two rounds, provided they pool information in the discussion.[8] If a

[8] In order to allow a particularly sharp comparison between the *Individual* and *Discussion* rounds, we ensure that exactly the same number of draws are available to each individual by the end of the both rounds. For instance, suppose that the wife gets w draws first and h draws second in the *Individual* round, for a total of $w + h$ draws. We ensure that the husband in turn receives h and then w draws in the *Individual* round. To make the *Discussion* round comparable, we ensure that the wife receives

person responds equally to the second set of draws in both rounds, we can conclude that learning one's teammate's information through a discussion is just as good as receiving the information oneself. If participants are instead less sensitive to information collected by their teammate, this implies either a failure of communication or under-weighting of information provided by one's teammate.

To tell apart these channels, we implement additional treatments which shut down communication frictions to various degrees. In rounds 3 through 5, we implement another *Discussion* round and two *Info-sharing* rounds in randomized order, as described below and illustrated in Figure 1(b).

***Draw-sharing* round.** This round is identical to the *Discussion* round except that, after participants receive their first set of draws and enter their first guess, they are told their teammate's draws (both number and composition) directly by the experimenter, e.g. "Your spouse had five draws, of which three were red and two were white." They then make an additional private guess which can incorporate both sets of draws before moving on to the discussion, joint guess and final private guess.

Comparing the guess made after the draw-sharing (but before discussion) with the final guess in the *Individual* round allows us to directly test whether participants use information they gathered themselves in the same way as information collected by others but perfectly shared with them. In each case, there is no possibility of joint deliberation.[9] A failure to do so may imply some sort of ownership effect over information or greater vividness of one's 'own information', or a general heuristic of perceiving information learned from others as being less reliable. Comparing the post-discussion guess in the *Draw-sharing* round with the *Discussion* round holds fixed the possibility of joint deliberation while testing whether communication frictions in discussion inhibit information pooling.

***Guess-sharing* round.** The *Guess-sharing* round is the same as the *Draw-sharing* round except that the experimenter informs each person of their spouse's private guess (made based on their own draws only), rather than their spouse's draws. The experimenter also shares the number of draws this guess was based on, e.g. "Your spouse had

w draws and the husband receives h draws, such that, if they pool information in their discussion, each again has access to $w + h$ draws to inform their final private guess. (h, w) are randomized across couples. In the other rounds, (h, w) are randomized independently within-couple across rounds.

[9] Note that this comparison requires controlling for order effects, since the *Individual* round is always in the first two rounds, while the *Draw-sharing* round falls in rounds 3-5.

5 draws and, after seeing these draws, they guessed that the urn contains 12 red balls." Thus, while in the *Draw-sharing* round we directly transmit the signal received by one's spouse or teammate, in the *Guess-sharing* round we transmit the action (guess) taken based on that signal. This round parallels more closely a large previous literature which investigates social learning based on others' actions (Weizsäcker, 2010). When observing actions, beliefs about others' competence might affect how these actions are interpreted and how much is learned about the signals. With *Draw-sharing*, i.e. directly sharing the signal, such beliefs should not be relevant.

The patterns of results in the *Guess-Sharing* round are in general similar or more extreme than those that we find in the *Discussion* and *Draw-Sharing* rounds. Because these patterns can be explained by rational mistrust of others' guessing, we leave analysis of the *Guess-Sharing* round to Appendix A.1.

Strangers experiment. One might reasonably expect married couples to be better at pooling information than strangers paired in teams. On the other hand, social norms regarding gender roles in marriage or past experiences of household decision-making may distort learning. To learn (i) whether spouses pool information differently than teams of comparable strangers; and (ii) whether any gender differences between spouses are specific to the marital context, we repeat the experiment with pairs of strangers with similar demographics, as in Abbink et al. (2020). The pairs of strangers play the same five rounds of the task as above, the order of which was similarly randomized. However, participants play one of the two *Discussion* rounds—picked at random—in same-gender pairs, and the other four rounds in mixed-gender pairs.[10] This additionally allows us to also learn if discussions work better in same-gender teams, and whether any gender differences in learning persist in a same-gender environment or if mixed-gender environments instead create gender differences in behavior as in Babcock et al. (2017).

3.2 Incentives to pool information and make accurate guesses

We create incentives for participants to pool information and make accurate guesses. Specifically, we reward one randomly-chosen guess from each pair for its accuracy. To

[10] To accomplish this, we recruited four participants—two men and two women—at a time. For all but the same-gender *Discussion* round (and *Individual* round, which participants played alone), each man was matched with the same woman. One of the two *Discussion* rounds was randomly chosen to be the same-gender *Discussion* round, for which players were matched with the other person of the same gender in their group.

ensure that incentives are aligned within the pair, we divide the payoff equally between the two participants irrespective of who made the guess. Each participant receives their half in a separate envelope at the end of the experiment.[11] Each person thus has an incentive to make every guess from their team as accurate as possible. Neglecting to ask your teammate for information or withholding information from your teammate reduces your own expected payoff.

The incentives provided are easy for participants to understand: a penalty per ball away from the truth. Formally, each guess is incentivized by a piece-wise linear loss function. On top of their participation fee, each couple receives an amount in Rupees (Rs.) equal to $\max\{(210 - 30 \times |g - r|), 0\}$, where g is the guess and r the true number of red balls for the randomly-selected guess.[12] These incentives are sizable. Rs. 210 is about \$3 and Rs. 30 is about \$0.40, while average daily earnings per capita are about Rs. 350 (\$5). Further, as described in Section 5.4, randomizing higher stakes for half the rounds in a follow-up experiment does not change our findings. The incentive scheme was explained to participants using the illustration shown in Appendix Figure A.I(b).

We can calculate what a risk-neutral Bayesian who maximizes expected earnings would guess given a set of signals.[13] However, our analysis does *not* assume that participants are Bayesian or risk-neutral. Instead, we compare guesses made by the same individual *across* rounds to determine if participants place equal weight on their own and their teammate's information. Even if participants are risk averse or deviate from Bayesian updating in the many ways documented in the literature (Benjamin, 2019), they should still respond similarly to their own and their teammate's signals, since the order of receiving the signals, the number of draws and the prior are held equal across rounds by the experiment.

[11]Of course, married couples—but not strangers—might redistribute the earnings after leaving the experiment. Even so, each spouse should be at least weakly better off by pooling information.

[12]We do not interpret the guesses as subjective beliefs targeted by a scoring rule, but as actions which participants have an incentive to tailor to the signals they receive. An alternative approach would be to directly elicit feature of participants' posterior distribution using a proper scoring rule (Palfrey and Wang, 2009). We avoided this due to the difficulty of explaining robust scoring rules to even higher-education populations (Danz et al., 2020). That said, our incentive scheme is a proper scoring rule for the median of the Bayesian posterior under risk neutrality, due to its absolute value form. The exception is following rare extreme draws (mostly red or white) where the truncation of the loss function at zero incentivizes shading the guess towards 50% red.

[13]A risk-averse Bayesian would shade their guesses towards 50-50, thus appearing less sensitive to signals. A noisy decision-maker would also appear less sensitive, since censoring of the number of red balls at 4 and 16 causes noisy choices to look more conservative.

3.3 Complexity and comprehension

We designed the task with the intention to balance two goals. First, given relatively low education and numeracy levels in our sample, it was meant to be easy to understand and feasible for most participants. We therefore avoided eliciting probabilistic beliefs or employing complex scoring rules. Similarly, we used a uniform distribution since it was easy for participants to understand. We also provided training in the task before the first round. Participants individually play two unincentivized practice rounds with two guesses in each, and receive two 'tips' on making good guesses.[14] The vast majority understood the tasks, as measured by excellent performance on comprehension checks (Table A.II).

In addition, the simple environment of our experiment does not require participants to use others' actions to make (potentially complex) inferences about their information. In the *Individual* round, participants simply receive all information themselves. In the *Discussion* and *Draw-Sharing* round, they are either told their teammate's information by the experimenter or can directly ask their teammate for it. This setup is in contrast to studies where participants observe other participants' decisions and must both infer the underlying signals as best they can and then make decisions based on those inferences (Goeree et al., 2007; Reshidi, 2020; Chandrasekhar et al., 2020).

Second, the task was designed to be sufficiently complex—as in many learning problems in the field—to create some ambiguity and wiggle room for biases and heuristics to enter participants' decision-making. Making the task trivial—e.g. with just one or two balls in the urn—would potentially eliminate all biases since there would be little disagreement in the teammates' signals and the correct action would become obvious to everyone. In such cases, even couples who often have difficulty communicating or tend to ignore each other's information might appear to perfectly pool information.

3.4 Individual performance and beliefs about ability

Before testing the key hypotheses, it is helpful to describe a few basic facts about participants' performance in the *Individual* round.

[14]The first tip explains that it makes sense to guess there are more red than white balls if you draw more red than white, and vice-versa. The second tip is that "the more balls you draw, the more confident you can be in your guess".

Actual performance. Figure 2(a) plots participants' actual average performance—the expected earnings from their guesses—in the individual round against the *number* of draws they received. Reassuringly, expected earnings increase with the number of draws, implying that participants benefit from more signals. The figure also provides two extreme benchmarks: a risk-neutral Bayesian who saw exactly the same signals, and someone who randomly guesses uniformly between 4 and 16. Participants' performance lies roughly halfway between the random guesser and the risk-neutral Bayesian.

Gender differences in performance. Figure 2(a) and Table 1 show no significant gender differences in performance. Men and women have nearly identical expected earnings in the individual rounds (Rs. 122 vs. Rs. 120). In 48% of couples, the wife outperforms the husband.

Gender and beliefs about ability. After completing the experimental rounds, we asked participants to privately predict their own and their teammate's average expected earnings.[15] Men and women are both—equally—overconfident about their own ability, as reported in Figure 2(a) and Table 1. Men correctly believe that their wives are as good as they are, but intriguingly women *incorrectly* predict that their husbands are better than them. Women's inflated views of their husband's ability relative to their own does not extend to other men: when asked about their male teammates in the strangers experiment, or about men versus women 'in general', women do not think that men outperform women.

Altogether, we interpret these facts as suggesting that our experimental task is not particularly gendered. Men and women are equally good and, with the exception of wives' beliefs about their husbands, largely believe that they are equally good. This is worth noting, since a recent literature has shown that the gender stereotype of particular tasks affects beliefs (Bordalo et al., 2019), belief-updating (Coffman et al., 2021b), and contributions to problem-solving (Coffman, 2014; Coffman et al., 2021a).

[15]These predictions were incentivized. For each participant, either their prediction of their own earnings or their prediction of their teammate's earnings was randomly picked to be paid off. Participants earned Rs. 50 if their prediction was within Rs. 30 of the truth, and otherwise earned nothing. It was not revealed to participants which guess was paid off. A participant's guesses were not revealed to their teammate.

4 Empirical Framework

Our goal is to test (i) whether individuals respond similarly to signals drawn themselves versus by their spouse; (ii) how this varies by gender; (iii) whether failures to learn are due to communication or information-processing frictions; and (iv) whether spouses learn differently from each each other than strangers working in teams.

We present three types of empirical analyses to answer these questions. First, we present non-parametric results by simply plotting average guesses in different rounds against a measure of the signals drawn (specifically, red draws minus white draws). This has the advantage of imposing minimal structure but does not allow for straightforward tests of the hypotheses. Second, we present reduced-form regressions which impose a linear relationship between signals and guesses, allowing for formal tests. Finally, we estimate a structural model of quasi-Bayesian belief updating, which incorporates the full structure of the signals, and provides results which are interpretable as deviations from a Bayesian benchmark.

4.1 Non-parametric approach

The non-parametric approach seeks to show patterns in the data with minimal assumptions. To do so, we relate participants' private guesses to the signals they received in the different rounds. Recall that participants always receive their first set of signals on their own. The variation in whether participants draw the signals themselves or must learn them from their teammate applies only to the second set of signals. We thus plot the average guess that participants make in their second private guess as a function of the second set of signals. For simplicity, we summarize the information content of the second set of signals by the net number of red draws (i.e., red minus white). That is, if a participant saw 4 red draws and 1 white draw, we would classify the signal as being 3 net red draws.[16]

Figure 2(b) shows such a plot for the *Individual* round, which serves as the control condition. It reveals that guesses relate sensibly to the signals: the higher the number of net red draws, the higher the number of red balls that participants guess. For

[16] This simplification loses some information, e.g. it does not capture the total number of draws. A signal with 1 net red could come from a single draw of a red ball or from 9 draws with 5 red and 4 blue. A Bayesian should react differently to these two signals. The structural model does not share this weakness.

comparison, the figure also provides the simulated average guesses of a risk-neutral Bayesian provided with the exact same draws. The empirical and Bayesian curves lie fairly close to each other, indicating that participants on average do a reasonable job. The empirical curve is slightly flatter than the risk-neutral Bayesian, which could occur for a number of reasons. Participants may simply be 'conservative' in their updating and under-react to signals in general, as often found in the literature (Benjamin, 2019). Another possibility is that guesses are noisy, which—because possible guesses only ranged from 4 to 16—leads the average participant to *appear* conservative.[17] Finally, risk aversion would also push participants towards appearing conservative by inducing them to shade their guesses towards 50-50 (10 red). We do not seek to fully disentangle these explanations, as our focus is instead on testing whether guesses respond differently to information depending on the round by contrasting behavior *across* rounds.

4.2 Reduced-form approach

Next, we formally test whether participants respond more to information they gathered themselves compared to information their teammate gathered. To accomplish this in a reduced-form way, we impose a linear relationship between signals and the resulting guesses and test for differences in this relationship across treatments. We define $\textit{First Info}_{irt}$ and $\textit{Second Info}_{irt}$ as the net number of red draws (i.e., red minus white draws) in the first and second set of signals, respectively, for individual i in round number r and treatment (i.e. round type) t. We then estimate using OLS:

$$\begin{aligned} \textit{Guess}_{irt} = \alpha + \beta_1 \cdot \textit{First Info}_{irt} + \beta_2 \cdot \textit{Second Info}_{irt} \\ + \boldsymbol{\beta_3} \cdot \mathbf{T}_{irt} \cdot \textit{Second Info}_{irt} + \epsilon_{irt} \end{aligned} \tag{1}$$

where $\textit{Guess}_{irt}$ is i's private guess—after having a chance to learn both sets of signals—of the number of red balls in round number r and treatment t. β_1 captures the "weight" that participants put on their first set of draws, averaging across all treatments.[18] β_2 is the weight they put on their second set of signals in the *Individual* round, where they

[17] To see this, note that if the Bayesian guess is 14, there are mechanically more possible guesses that are less extreme than Bayesian (i.e., toward 50-50) than that are more extreme.

[18] We average across all treatments for *First Info* since the treatment only applies to the second set of signals. In a robustness check (Table A.X), we allow β_1 to vary by treatment. This does not change our conclusions.

gather these signals themselves.[19] $\mathbf{T}_{irt}$ is a vector of indicators for whether a particular guess corresponds to the *Discussion*, *Draw-sharing (pre-discussion)* or *Draw-Sharing (post-discussion)* treatments.[20] $\boldsymbol{\beta}_3$ is a vector of coefficients capturing the *additional* weight placed on the second set of signals in treatment t relative to the *Individual* round. Finally, while suppressed above for ease of exposition, we also allow the weights on first and second signals to vary arbitrarily by round number by including round-number dummies interacted with *First Info* and *Second Info*. Standard errors are clustered at the level of the team.

The treatment effects are thus captured by $\boldsymbol{\beta}_3$. With perfect information-pooling, it should not matter whether the signals were drawn by oneself or by one's teammate, i.e. $\boldsymbol{\beta}_3 = 0$. If instead an element of $\boldsymbol{\beta}_3 < 0$, this implies that participants in the corresponding treatment do not fully learn their teammate's information or under-weight this information relative to their own.

4.3 Structural approach

In our third empirical approach, we estimate a simple model of quasi-Bayesian updating. The model accounts for the full information content of the signals and incorporates the censoring in our data. It also allows us to compare learning against a normative benchmark: a risk-neutral Bayesian. To do so, we assume that, after observing the two sets of draws/signals d_1 and d_2 (e.g., d_1 might equal $\{\text{Red}, \text{Red}, \text{White}, \text{Red}, \text{White}\}$), participants update their beliefs about the state of the world s (the number of red balls in the urn) according to a modified version of Bayes' Rule:

$$Posterior(s|d_1, d_2) \propto Prior(s) * P(d_1|s)^{\omega_1} * P(d_2|s)^{\omega_2} \tag{2}$$

where $Prior(s)$ is the participant's prior about the probability of state s, and $P(d_i|s)$ is the conditional probability of observing a set of draws d_i conditional on state s. Recall that participants are told each state is equally likely, and there are 13 possible states $s \in \{4, 5, ..., 16\}$, so $Prior(s) = \frac{1}{13}$. Next, ω_1 and ω_2 are the weights that the participant puts on the first and second set of draws respectively. Note that when $\omega_1 = \omega_2 = 1$,

[19] Since the number of draws received in the first and second signals is the same on average, a Bayesian playing alone should weight them equally on average, i.e. $\beta_1 = \beta_2$.

[20] Recall that the different types of rounds are played in randomized order as shown in Figure 1. In this section, we treat the pre-discussion and post-discussion guesses in the *Draw-Sharing* round as different treatments.

equation (2) reduces to Bayes' rule.

We allow ω_1 and ω_2 to differ from the Bayesian benchmark in several ways. First, participants can simply put more or less weight on each piece of information than a rational agent with full information would, allowing for intrinsic over- or under-reaction to either or both signals. Next, they can put more or less weight on signals depending on the chronological order of the round, allowing for learning during the course of the experiment. Finally, we allow them to place different weight on signals depending on how they learned about them, e.g. placing greater weight on information they gathered themselves. We choose a functional form to closely match the reduced-form specifications above:

$$\begin{aligned} \omega_{1rt} &= \beta_1 + \mu_{1r} \\ \omega_{2rt} &= \beta_2 + \mu_{2r} + \boldsymbol{\beta_3} \cdot \mathbf{T}_{rt} \end{aligned} \tag{3}$$

The participant always draws the first set of signals themselves, so ω_1 is only allowed to vary with the chronological order in which the round occurred. Thus β_1 is the weight participants place on the first set of draws in the first round, and μ_{1r} captures how this weight changes with round order. β_2 captures the weight participants put on their second signal in the *Individual* round, with μ_{2r} again representing order effects. As before, $\mathbf{T}$ is the vector of treatment indicators and the elements of $\boldsymbol{\beta_3}$ indicate the additional weight participants put on the second set of signals in other treatments relative to the *Individual* treatment. A negative value in the vector $\boldsymbol{\beta_3}$ indicates that participants put less weight on the second set of draws in the corresponding treatment compared to the individual round where they gather these signals themselves. As before, the set of treatments is $t \in \{$*Discussion*, *Draw-Sharing* (pre-discussion), *Draw-Sharing* (post-discussion)$\}$.

In addition to systematically biased updating, we allow for noisy choice. Doing so allows us to account for heterogeneity in guesses conditional on signals (i.e., not everyone with the same signals makes the same guess). We assume that agents are risk-neutral but calculate the expected payoff of each possible guess with noise. In particular, we can define $Earnings(g, s)$ to be the earnings, given the experimental incentives, that a participant would earn (more precisely, split with their spouse) if they made guess g and the true state was s.[21] We assume that the agent calculates the expected payoff

[21] The incentives we employed imply that $Earnings(g, s) = \max\{0, 210 - 30 * |g - s|\}$.

of each guess g using their (potentially biased) updating rule given by equation 2 plus a random additive error term. That is, we assume the perceived expected payoff from making guess g given signals d_1 and d_1 is given by

$$EP(g|d_1, d_2) = \sum_{s=4}^{16} Posterior(s|d_1, d_2) Earnings(g, s) + \alpha \epsilon_{i,g}$$

The agent then chooses the guess that maximizes this perceived expected payoff. For simplicity, we assume $\epsilon_{i,g}$ is iid Type 1 extreme value. The parameter α then governs the extent of noisy choice.[22] We estimate the model by maximum likelihood.[23]

5 Results: Learning from own vs. others' information

5.1 Non-parametric results

We first provide a qualitative illustration of our main findings by plotting participants' guesses against the second set of signals in each treatment. Reassuringly, as discussed in Section 4.1, the (average) number of red balls guessed by participants increases in the number of "net red" draws in the *Individual* rounds, implying that participants respond to information they receive. We will examine how this responsiveness differs across treatments.

Figure 3 depicts the relationship between participants' guesses and the "net red" draws in their second set of signals in the *Discussion* and *Draw-sharing* rounds, comparing each to the *Individual* round. For husbands (Panel A), the blue curve representing the *Discussion* round (top left) is distinctly flatter than the grey curve showing the *Individual* round, indicating that husbands' beliefs are less sensitive to information

[22] See Goeree et al. (2007) for an example of a similar model of noisy discrete choice in a balls-and-urns decision problem.

[23] In particular, given the assumptions above, the probability that an agent with signals d_1 and d_2 will choose guess g is $P(i \text{ guesses } g|d_1, d_2) \propto \exp\left(\frac{1}{\alpha}\left[\sum_{s=4}^{16} Posterior(s|d_1, d_2) Earnings(g, s)\right]\right)$. We then choose parameters that maximize the joint likelihood of observing all the choices in our data. We calculate standard errors by bootstrapping the data, drawing couples (or groups of strangers) with replacement from the data. Throughout, we report bootstrapped standard errors for legibility but denote significance using bootstrapped confidence intervals (e.g., an estimate is significant at the 5% level if the center 95% of bootstrapped estimates do not include zero).

that their wives gathered compared to information they gathered themselves. As we saw in Figure 2(b), participants' guesses even in the *Individual* round are themselves less responsive (on average) to information than a Bayesian's would be. The even lower sensitivity to signals in the *Discussion* round suggests that husbands learn less effectively from discussing with their wives compared to learning on their own, which we will later show is costly in terms of expected earnings.

Strikingly, the curve is even flatter in the *Draw-sharing* round (top right), in which wives' information is *directly* communicated to their husbands by the experimenter. Despite having been given *all* decision-relevant information about their wives' draws directly, husbands react to this information much less than they do to information that they collected themselves.[24] This result suggests that the key friction is not communication (i.e. husbands never learning the information from their wives) but instead husbands underweighting information uncovered by their wife, as discussed in Section 5.2 in more detail.

In contrast, wives (panel B) do *not* appear to put less weight on information that their husbands discovered compared to information they gathered themselves. For wives, both the *Discussion* and *Draw-sharing* rounds look similar to the *Individual* round. This implies that women react to their husbands' information as they do to their own when it is directly shared with them in the *Draw-sharing* rounds. And they are able to effectively learn that information from their husbands in the *Discussion* rounds.

Figure A.II shows the equivalent figures for men and women in the strangers sample. Both men and women in this sample appear to respond less to information that comes from their teammate, particularly in the *Draw-sharing* round. The finding above that women—unlike men—equally weight their own and their spouse's information thus does not appear to be due to gender differences in information-processing in general. Instead, it is specific to the marital context: men treat their wives' (information) like strangers, while women give their husbands' (information) greater weight than they do to strangers.[25] We discuss interpretations of these findings in the following sections.

[24]Note that this graph depicts the second private guess—after being informed of one's teammate's draws but before having a chance to discuss with them. This provides a clean comparison with the individual round: the only difference is drawing the signals oneself versus being informed of the signals one's teammate drew.

[25]The *Discussion* round results in Figure A.II pool data from both the same-gender and mixed-gender *Discussion* rounds, but the *Draw-Sharing* rounds in the strangers sample, in which both men

5.2 Couples experiment

To formally test for differences across treatments and gender, we turn to our reduced-form and structural estimates, which yield overall similar results. Figure 4 plots the average "weights" husbands and wives place on the second set of signals, by treatment, using estimates from equation 1. Panel A shows that—in their final private guesses after the discussion—husbands put significantly less weight on the second set of signals in the *Discussion* round compared to the *Individual* round (p<0.01). This implies they respond less to information collected by their wives compared to their 'own' information. Either they do not learn their wives' signals in the first place (communication frictions), or they discount them relative to their own signals upon learning them (information-processing frictions).

Panel A also shows that husbands put close to zero weight on their wife's information in the *Draw-sharing* round, right after it is *directly* shared with them (third bar). By design, this behavior cannot be explained by failure to communicate information or by mistrust of the spouse's memory or motives. Adding a face-to-face discussion with their wife after being informed of her draws somewhat increases husbands' weight on their wife's signals (fourth bar), but the weight remains identical to the *Discussion* round, and significantly below the weight on their own signals (p<0.01). Thus, eliminating communication frictions does not increase the weight husbands put on their wife's information. Husbands' underweighting of their wife's information must therefore be due to differences in how they process information discovered by their wife relative to their own information.

Wives, in stark contrast to husbands, do *not* under-weight their spouse's information relative to their own. Panel B shows that wives place nearly identical weights on their own information in the *Individual* round and their husbands' information in the *Discussion* round (p=0.61), and the *Draw-sharing* round, both pre-discussion (p=0.94) and post-discussion (p=0.35). Thus, wives are able to learn their husband's signals in the discussion and place equal weight on them and their own signals.

The corresponding regression estimates are presented in Table 2 (cols 1 to 3). $\beta_{3,1}$ captures the treatment effect of the *Discussion* treatment. The estimate pooling husbands and wives shows a clear result (col 1): participants put 38 percent ($-0.20/0.53$)

and women respond substantially less to their teammate's information, were always done in opposite-gender pairs.

less weight on information collected by their spouse in the *Discussion* round relative to information they collected themselves in the *Individual* round (p<0.01). This effect is primarily driven by husbands (col 2), who put 58 percent less weight on information gathered by their wives compared to information they gathered. In contrast, the effect for wives (col 3) is relatively small at 7 percent and statistically insignificant. Husbands and wives thus place a very different 'discount' on information collected by their spouses: the $\beta_{3,1}$ coefficients in columns 2 and 3 are significantly different from each other (p=0.02).

The $\beta_{3,2}$ and $\beta_{3,3}$ terms in Table 2 capture the differential weights participants put on their spouse's information in the *Draw-sharing* round. Husbands and wives (pooled) put 53 percent ($-0.28/0.53$) less weight on their spouse's draws in the pre-discussion *Draw-Sharing* guess compared to in the *Individual* round (p<0.01). Again, this effect is almost entirely driven by husbands who put a striking 98 percent ($-0.53/0.54$) less weight on their wife's information compared to their own (p<0.01). This difference is less pronounced and not statistically significant after the discussion when pooling both spouses' guesses (25 percent, p=0.14) but still sizable and highly significant for husbands (65 percent, p<0.01). In contrast, women treat their second set of own information essentially the same as their spouses' information.

The structural estimates mirror the reduced-form results (cols 4-6 of Table 2). Husbands and wives (pooled) put 53% less weight ($-0.73/1.37$, p<0.01) on their spouse's signals in the *Discussion* round. As before, this effect is driven primarily by husbands who put 78% less weight (p<0.01). Wives put 25% less weight on their spouse's signals, though this difference is not statistically significant (p=0.39).[26]

Comparing husbands and wives. The very different behavior of husbands and wives does not appear to be explained by differences in observable characteristics other than gender (Table A.III). Pooling both spouses' guesses from the *Discussion* and *Draw-Sharing* rounds, we first confirm that wives place significantly higher weight on their spouse's information than husbands do (0.47 vs. 0.31, col 1). We then test whether this additional weight shrinks when controlling for a rich set of observables. Controlling for relative age, perceptions of who the household decision-maker is, perceived and

[26]Of course, the parameters of this quasi-Bayesian model have a different scale and interpretation than the reduced-form results discussed above. But perfect information pooling implies $\beta_{3,1} = 0$ in both cases. Appendix A.3 shows that the reduced-form and structural estimates are consistent with each other: data simulated using the structural model produces the same reduced-form results as the empirical data.

actual guessing ability, and comprehension—by including interactions of these controls with the signals—does not reduce the gap between wives' and husbands' weights on their spouse's information (cols 2 through 7). Instead, wives' higher weights on their spouse's information may be driven by gender differences in learning (e.g. women may not under-weight others' information in general) or by something specific to gender norms in marriage in our study context (e.g. "women should defer to their husbands").[27]

Earnings implications. A direct measure of performance in the experiment is the expected earnings from guesses. Figure 5 plots average expected earnings from guesses as a function of the number of draws in the second set of signals. It compares earnings in the *Individual* round with pooled earnings from the treatment *Discussion* and *Draw-sharing* (both pre- and post-discussion) rounds. Panel A shows that husbands' earnings decline relative to the *Individual* round as the number of draws received by their wife increases. When their wife only receives one draw, underweighting her information is not significantly costly. When she has 5 draws—on average, as many as he does—his earnings decline by Rs. 5.7 (p=0.12). When she is better informed than him, having received 9 draws, his earnings decline by Rs. 11 (9 percent or 0.3 standard deviations, p=0.03). In contrast, and consistent with the finding that wives do not under-weight their husband's information, Panel B shows that women earn the same on average regardless of whether they receive all draws themselves or instead must rely on some draws from their husband. Even when their husband receives 9 draws, we cannot reject that women do just as well as in the *Individual* round (p=0.75).

Turning to regressions in Table 3, husbands earn significantly less in the *Discussion* and *Draw-sharing* rounds than in the *Individual* round (column 2), while the differences for wives are small in magnitude and not significant (column 3). The next three columns ask how these differences depend on the number of draws (i.e., on the informativeness of signals). We see, as expected, that for both sets of signals more draws significantly increase expected earnings (about Rs. 2.5 per extra draw). However, husbands earn significantly less for each additional draw their wife makes in the *Draw-Sharing* round, both pre- and post-discussion. As expected given the patterns in Figure 5, we see no significant differences in earnings per draw across treatments for wives.

[27]The final column of Table A.III shows that the experimenter being present ("Public Discussion") does not significantly affect either the wife's or the husband's weight on their spouse's information. Note that the guesses considered so far are always made privately, so this variation may be expected to have more bite in joint guesses, which we discuss in Section 5.5.

Summary. We find that spouses pool information in an asymmetric way. Wives learn their husband's information in discussions and weight it equally with their own. In contrast, husbands put much less weight on their wife's information, even when it is perfectly shared with them, and appear to treat data their wives generated as less informative. Husbands' guesses thus earn less when part of the information is gathered by their spouse, particularly when the spouse is relatively well-informed. We next investigate whether these results are specific to married couples or if under-weighting and gender differences in social learning emerge more generally in teams.

5.3 Strangers experiment

Panel C of Figure 4 plots the weights that men place on the second set of signals when playing with strangers. Men's behavior in this experiment is strikingly similar to that of husbands towards their wives (Panel A). Compared to their own signals, men put significantly less weight on the signals drawn by their teammate. We find this pattern in the *Discussion* round (p<0.01) and in the *Draw-sharing* round both pre- and post-discussion (p<0.01 and p=0.04, respectively). This suggests that men treat strangers' information just as they treat their wife's: they substantially discount it, even when it is directly communicated to them.

Panel D tells a slightly different story for women playing with strangers. Recall that wives did not discount their husbands' information in any round. In contrast, we find evidence that women *do* discount their teammate's information in the experiment with strangers. In the *Discussion* round, played once each with male and female teammates, we find suggestive evidence that women discount their teammate's information (p=0.12). In the *Draw-sharing* round, which women play only with male teammates, women significantly discount their teammate's information pre-discussion (p=0.02) and suggestively post-discussion (p=0.10). Pooling across the three treatments, there is clear evidence of women discounting their teammate's information (p<0.01). Prima facie, this suggests that women do not treat others as they treat their husbands, in terms of social learning. It is not that women generally weight their own and others' information equally, but they do when the 'other' is their spouse.

Turning to the regressions (Table 4), the second set of draws get less than half as much weight when accessible only via discussion with a stranger (col 1). This effect is more pronounced for men (col 2) compared to women (col 3) but this gender difference is not statistically significant (p=0.22). The point estimates imply that men discount their teammate's information by 70 percent (p<0.01) while women discount their teammate's information by 33 percent (p=0.12). In the structural estimates (cols 4 to 6), which account for noisy choice as well as the sample size of the draws, the estimated average discounting of teammate's information is even more pronounced (73 percent), with slightly more pronounced discounting for men (76 percent) than for women (68 percent), both highly statistically significant (p<0.01).[28]

Similarly, in the *Draw-sharing* rounds, strangers on average put 84 percent less weight on their teammate's information in the pre-discussion guess (col 1). The estimates are statistically significant for men and women in both the reduced-form and structural estimates (cols 2 to 6). Having a face-to-face discussion mitigates this effect. While still sizable in magnitude (47%), the estimate for women is only marginally significant post-discussion in the reduced-form estimates (p=0.10). In the structural estimates, the post-discussion effect is sizeable and significant for both men and women.

For the *Discussion* round, we can compare behavior in same- and mixed-gender pairs. We find no evidence of systematic differences in discounting of their teammate's information in same- compared to mixed-gender pairs (Table A.IV).

Earnings implications. Figure 5 Panels C and D show earnings by male and female strangers across treatments. Broadly similar to husbands playing with their wives, strangers earn less in the *Discussion* and *Draw-sharing* rounds compared to the *Individual* round, when their teammate receives as many or more draws than them. Turning to regressions, Table A.V shows that men and women both generate lower expected earnings when part of the information is collected by their teammate (col 1), though only men's pre-discussion guesses in the *Draw-sharing* round is statistically significant (cols 2 and 3). Strangers also earn significantly less per additional second-signal draw in the *Discussion* and the pre-discussion *Draw-sharing* rounds compared to the *Individual* round (col 4).

Comparing couples and strangers. To formally test for gender differences between couples and pairs of strangers, we pool the *Discussion* and *Draw-sharing* rounds

[28] One reason the structural estimates generally provide sharper results than the reduced form is that the structural model properly accounts for the number of draws (i.e. the sample size) each teammate gets. Empirically, it is precisely when a participant's teammate receives most of the draws that the gap between the *Individual* and *Draw-sharing* round is largest. The structural model better incorporates this pattern in the data.

of both samples and estimate how the weights placed on one's teammate's info vary by sample (couples vs. strangers) and by gender (Table A.VI). While women place a significantly higher weight than men on their teammate's information overall (0.41 vs. 0.28 for men, col 1), this effect is driven more by women playing with their husbands rather than women in general (col 2).[29] Meanwhile, the difference between husbands playing with a stranger and husbands playing with their wife is smaller and insignificant (0.24 vs. 0.31 playing with their spouse, col 2).

The differences between women playing with their husbands and others cannot easily be explained by differences in observable characteristics. In Table A.VII, we again pool the two samples and consider the *Discussion* and *Draw-sharing* rounds. We then run a series of regressions in which we control for interactions between teammate's signals and various demographic variables, measures of ability and beliefs, and household decision-maker measures. None of these additional interaction terms affect the significance or magnitude of being a woman in a couple. That is, observable characteristics such as age or beliefs about ability cannot explain why women weight only their spouse's information (and not others') more than men do.

5.4 Interpretation, confounds, and robustness

To recap, our experiments uncover two main results. First, participants put less weight on information that their teammate discovered, even if this information is perfectly communicated to them. Second, the exception to this pattern is that wives weight their husband's information equally to their own.

Interpretation. One interpretation of the first result is that people perceive information they uncovered themselves as "theirs" and put more weight on it by virtue of ownership. This broadly relates to the literature on ownership effects and the endowment effect (e.g. Kahneman et al., 1991; Hartzmark et al., 2021) and also to 'egocentric bias', whereby individuals think that their own views are more commonly held than they actually are (Ross et al., 1977).

A related possibility is that the own information in our experiment was simply more vivid and salient. For instance, the act of physically drawing the balls one-by-one might be more compelling than learning the same information through discussion

[29] Relative to men playing with strangers, women playing with strangers also place somewhat more weight on their teammate's information, although this difference is not significant (col 2).

or a dry message from the experimenter. This may be a general feature of learning from experience versus learning from others, related to the economics literature on experience effects (e.g. Malmendier and Nagel, 2011; D'Acunto et al., 2021). A third round of our experiment, conducted with strangers and presented in Appendix A.2, sheds some light on this hypothesis. In one treatment, participants directly observe their teammate drawing balls from the urn while sitting next to them *in the same booth*. Strikingly, people underweight their teammate's information relative to their own even in this treatment (p=0.03), although the weight placed on teammates' information is somewhat higher than in the draw-sharing round (p=0.10). This is consistent with at least some role for vividness, but still leaves a substantial gap relative to uncovering information oneself.

A final interpretation is that participants are using a reasonable heuristic from their everyday lives. People may think that information they directly acquire—seen with their own eyes—tends to be less noisy or more pertinent than information shared by others. Thus, it may be appropriate to put more weight on one's own information in many real-world situations. Participants may then extend this heuristic to situations, such as our experiment, where it does not apply. But is the extent of underweighting of others' information seen in our experiment quantitatively sensible? Our structural estimates suggest that husbands underweight their wife's information relative to their own by 78 percent in the *Discussion* round and almost entirely in the *Draw-sharing* round pre-discussion. Wives' information would thus need to be at most a quarter as 'informative' as husbands' own in typical situations to justify the observed degree of underweighting.

Our interpretation of the second result is that women weighting their husband's information equally is directly tied to the marital context. One possibility is that a social norm of deferring to one's husband—but not to one's wife—may counteract the general tendency to underweight others' information. Note that any such norm would have to be an internalized one, since the guesses we focus on so far are all *private* guesses. Another possibility is that women follow a reasonable heuristic that their husbands (compared to strangers) possess information that is particularly relevant to them. Of course, this raises the question of why husbands do not display the same heuristic towards their wives.

Confounds. The experimental design attempts to render several alternative explanations and confounds irrelevant. For example, participants' beliefs about their teammate's ability to engage in Bayesian updating should not matter. As long as one's teammate remembers the draws they saw mere minutes ago, one should be able to learn their information and weight it appropriately.[30] In the *Draw-sharing* round, of course, one need not even rely on one's teammate to remember their signals. Unlike in many information cascade experiments, agents do not need to infer others' signals by interpreting their actions (e.g. Goeree et al. 2007 and De Filippis et al. 2017), since they can simply ask for (or receive) their teammate's signals. Nor do participants have incentives to follow their own info in order to altruistically signal it to later agents (March and Ziegelmeyer, 2020). Instead, they have simple, economically-meaningful incentives to share information in each direction, since one guess by either teammate is randomly picked to be paid out, and the earnings are divided equally between them.

Risk aversion. Risk aversion cannot explain our main results. Risk-averse participants have a slight incentive to shade their answer towards the median of the distribution (10 red balls), given our incentive structure. Women are often found to be more risk averse than men on average (Eckel and Grossman, 2008), and so one might expect women to update less than men in our experiment. However, men and women play very similarly in the *Individual* rounds—recall Figure 2(b)—and, importantly, our research design rests on comparing guesses *across* treatments. Gender differences in risk aversion should show up in the degree of belief updating in the *Individual* rounds rather than in differences in belief updating across the *Individual* and other rounds.

Trust in the experimenter. One concern with the *Draw-sharing* round is that participants may mistrust the experimenter (relative to their teammate) and thus may rationally place lower weight on information given by the experimenter. However, participants could verify the experimenter's message in the face-to-face discussion with their teammate, such that any deception would be easily unveiled. No concerns along these lines were expressed in debriefing or emerged in the transcripts of the unstructured discussion between the teammates. Finally, as described above, significant underweighting of teammate's information occurs even when participants directly observe them making the draws (Appendix A.2).

Reputational concerns vis-à-vis teammates. One may worry that participants' guesses may be influenced by possible repercussions from their spouse for not using their in-

[30]In any case, recall that men correctly believed that their wives are as good as them at the task, as we reported in Section 3.4. Thus, beliefs about relative competence cannot explain why husbands do not weight their wives' information equally.

formation adequately, or for deviating from the jointly chosen guesses. To address this concern, we did not reveal any post-discussion private guesses, even if they were selected for payment. This ensures that people remain incentivized to guess what they truly think and not, for instance, agree with their spouse to avoid future retribution.[31] Nor do any such concerns seem plausible in the case of the strangers sample, since participants did not know each other and would not expect to have any future interactions.

Competition. Despite the aligned monetary incentives, participants may enjoy making better guesses than their teammate. Moreover, gender differences in preferences for competition are well-established in the literature (Niederle and Vesterlund, 2011). However, while such a force might lead to withholding of information (or even deception) in discussions, it should not affect the *Draw-Sharing* results. In that round, even if a participant's desire to guess accurately is motivated by a sense of competitiveness with their teammate, the best way to guess accurately is to use all information at hand—including what their teammate drew. To generate our results with a utility-based explanation, one would have to assume that participants (except wives) derive utility specifically from making guesses consistent with their own signals only, which is close to our suggested mechanism of egocentric bias or ownership effects.

Increasing stakes. Participants face significant stakes in our experiment. The third experiment described in Appendix A.2 shows that further increasing the stakes by 50 percent in randomly-selected rounds led to almost identical (in fact, insignificantly lower) weight placed on teammate's information (Figure A.VI). The stakes in the higher-stakes condition ($4.50) are close to the daily per capita income ($5) in the study sample. Of course, major household decisions with substantially larger stakes might well lead to better information pooling.

Robustness to design variations. The third experiment also tested other experimental variations which might eliminate underweighting of others' information and establish the limits of our findings. As reported in Figure A.V, we find significant underweighting of teammate's information in all of the following conditions: (i) observing one's teammate drawing their signals in person (p=0.03); (ii) a modified *Draw-sharing* round where the participant is informed of their teammate's draws ball-by-ball instead of in summary form (p<0.01); and (iii) not entering a first guess before receiving one's partner's information (p=0.02). The only condition in which underweighting is not

[31] Pre-discussion guesses were only revealed if selected for payment.

significant is when the participant received their teammate's information before receiving their own draws (p=0.20), but this is partly because of lower power, as only half the participants could be in this condition. None of these treatments are statistically significantly different from the *Draw-sharing* round. Nonetheless, we treat these results with caution, since the experiment was forced to stop in March 2020 at 146 pairs of strangers—well short of the target of 400 pairs—due to Covid-19 restrictions, and thus statistical power is somewhat limited.

5.5 Individual vs. joint decisions

The private guesses discussed so far shed light on the extent to which people learn from each other, and whether information flows freely within households and teams. Our experiment also allows us to study how people make *joint decisions*—in the form of joint guesses—which may ameliorate or exacerbate the biases documented above.

We first examine how the joint guess weights the information each participant uncovered. Perhaps surprisingly given the above results, couples' joint guesses put very similar weights on husbands' and wives' information (Figure A.IIIa). Couples put slightly more weight on wives' information compared to husbands' information in the *Discussion* round but do the opposite in the *Draw-sharing* round, though neither difference is statistically significant (p=0.54 and p=0.42, respectively). The weights on men and women among strangers look very similar to the weights among spouses (Figure A.IIIb). One interpretation of these patterns is that group decisions can mitigate biases in decision-making that occur in individual choices. While the teammates may not agree on the best guess—as evidenced by the fact that only half of participants make the same private final guess as the joint guess—a process of bargaining and compromise might ensure equal weights in the joint guess.

Table A.VIII shows the corresponding regressions and reports heterogeneity by relative ability, beliefs about ability, and by a proxy for household decision-making power. Consistent with Figure A.IIIa, wives' and husbands' information receives very similar weight in the joint guess (col 1). The joint decision weights more heavily the signals of the spouse who is better at the task. This is true for objective measures of performance (col 2) as well as for the subjective beliefs of the couple (col 3). Finally, the signals drawn by the households' "primary household decision-maker" (as reported by husbands) also receive significantly more weight (col 4). While all these patterns seem intuitive—higher ability and greater household bargaining weights lead to greater weight on one's information—it is worth emphasizing that they should *not* matter. As long as participants can recall their draws, such that their information can be pooled with their spouse's, joint decisions should *not* weight their information differently.

The relevance of ability, perceived ability and household decision-maker status in determining the weights on information in the joint guess hint that the process of deciding on the joint guess in the *Discussion* round does not always involve sharing the underlying signals. If only initial guesses are shared, it may be rational to weight more heavily those with higher ability.[32] Indeed, using transcripts of the recorded discussions, we find evidence along these lines (Table A.I). Forty percent of husbands and wives each share their first guess with their spouse in the discussion, while only about 30 percent share the exact composition of draws.

Given that husbands' and wives' information receive equal weight in joint guesses, joint decisions earn significantly more than individual guesses made by both husbands and wives in the same round, though the effect is larger for husbands (Table A.IX, cols 1 to 3). Guesses naturally yield higher expected earnings when they are based on more draws (cols 4 to 6). However, additional draws by the wife are more than twice as valuable in joint decisions compared to in husbands' final private guesses (col 5, $p<0.01$). We do not see significant differences for wives, since they already weight their husband's information equally. Joint guesses, by giving equal weight to wives' information, appear to improve decision-making for husbands who would otherwise neglect this information.

6 Conclusion

We designed a simple two-person social-learning task to study whether married couples efficiently pool their private information through a face-to-face discussion, when they have incentives to do so. We have three main findings. First, husbands substantially underweight their wife's information relative to their own, while wives equally weight their own and their husband's information. Second, this underweighting of oth-

[32] Draw-sharing appears to mitigate these heterogeneity patterns (cols 7 to 10). While we find that the information of spouses who perform better in the *Individual* round gets more weight in joint decisions (p=0.01, column 6), there are no clear pattern of heterogeneity by spouses' perceived performance (col 7) or by the gender of the main household decision-maker (col 8).

ers' information is not due to communication frictions, but instead occurs even when information is noiselessly communicated. Husbands treat information generated by their wives as being intrinsically less informative than information generated themselves through their own experiences. Third, both men and women underweight others' information when they are paired with strangers of either gender. This implies a general tendency to underweight others' information relative to one's 'own' information, with a counteracting tendency for women to weight their husband's information highly.

Our paper implies that—even absent strategic motives—households may not pool information. Moreover, they may fail to learn from each other in an asymmetric way, even in gender-neutral domains. If these findings prove to be true more generally, then a policy-maker who wants to ensure that both spouses in a household acquire some knowledge should not assume that informing or training one spouse will suffice. Moreover, information provided to women may be particularly discounted, as in BenYishay et al. (2020).

Beyond the household, the systematic underweighting of others' information we document could be a powerful and previously underappreciated barrier to social learning. Existing evidence already shows that, when faced with complex social learning problems, people deviate from Bayesian inference by, e.g. overweighting private signals relative to public information (Weizsäcker, 2010), neglecting selection (Enke, 2020), neglecting correlation in signals (Eyster et al., 2015; Enke and Zimmermann, 2019), and averaging beliefs as in DeGroot learning (Chandrasekhar et al., 2020). Moreover, social image and stigma can cause people to avoid seeking information from others in the first place (Chandrasekhar et al., 2018; Banerjee et al., 2018). We show that, even in a relatively simple setup with perfectly observed signals, people strongly underweight information learned from others relative to their own information. This could play a role in any number of settings where social learning is possible, from technology adoption to learning about health behaviors and investment opportunities.

Our study has numerous shortcomings which point to useful avenues for future research. First, it would be valuable to collect more data on differences in economically-important beliefs within the household, as in e.g. Fehr et al. (2019) and D'Acunto et al. (2021). Second, it will be important to study information pooling within the household using more natural field experiments and with higher stakes, as in Ashraf et al. (2020). Third, we studied a relatively ungendered domain in which men and women had similar ability and similar beliefs about own ability. Given the well-documented importance of gender stereotypes in beliefs and learning (e.g. Coffman et al. 2021a,b), it would be interesting to study if individuals weight their spouse's information more highly when it is in a domain congruent with the spouse's gender. Finally, we studied learning in a context where participants unearthed new information by themselves, as in learning by experience. It will be important to learn if people also weight information they learned from others more highly than information their spouse in turn learned from others. We hope that future research will make make progress in all these directions.

References

Abbink, Klaus, Asad Islam, and Chau Nguyen, "Whose voice matters? An experimental examination of gender bias in intra-household decision-making," *Journal of Economic Behavior & Organization*, 2020, *176*, 337–352.

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

Ambler, Kate, "Don't tell on me: Experimental evidence of asymmetric information in transnational households," *Journal of Development Economics*, 2015, *113*, 52–69.

Anderson, Lisa R and Charles A Holt, "Information Cascades in the Laboratory," *American Economic Review*, 1997, *87* (5), 847–862.

Angrisani, Marco, Antonio Guarino, Philippe Jehiel, and Toru Kitagawa, "Information redundancy neglect versus overconfidence: a social learning experiment," Technical Report, cemmap working paper 2018.

Apedo-Amah, Marie, Habiba Djebbari, and Roberta Ziparo, "Gender, information and the efficiency of household production decisions: An experiment in rural Togo," 2020.

Ashraf, Nava, "Spousal Control and Intra–Household Decision Making: An Experimental Study in the Philippines," *American Economic Review*, 2009, *99* (4), 1245–77.

_ , Erica Field, and Jean Lee, "Household Bargaining and Excess Fertility: An Experimental Study in Zambia," *American Economic Review*, 2014, *104* (7), 2210–37.

_ , Erica M Field, Alessandra Voena, and Roberta Ziparo, "Maternal Mortality Risk and Spousal Differences in the Demand for Children," *NBER Working Paper Number 28220*, 2020.

Babcock, Linda, Maria P. Recalde, Lise Vesterlund, and Laurie Weingart, "Gender Differences in Accepting and Receiving Requests for Tasks with Low Promotability," *American Economic Review*, 2017, *107* (3), 714–47.

Banerjee, Abhijit, Emily Breza, Arun G. Chandrasekhar, and Benjamin Golub, "When Less is More: Experimental Evidence on Information Delivery During India's Demonetization," *NBER Working Paper Number 24679*, 2018.

Beaman, Lori and Andrew Dillon, "Diffusion of agricultural information within social networks: Evidence on gender inequalities from Mali," *Journal of Development Economics*, 2018, *133*, 147–161.

Benjamin, Dan, Aaron Bodoh-Creed, and Matthew Rabin, "Base-Rate Neglect: Foundations and Implications," *Working Paper*, 2019.

Benjamin, Daniel J., "Chapter 2 - Errors in Probabilistic Reasoning and Judgment Biases," *Handbook of Behavioral Economics: Applications and Foundations 1*, 2019, *2*, 69–186.

BenYishay, Ariel, Maria Jones, Florence Kondylis, and Ahmed Mushfiq Mobarak, "Gender gaps in technology diffusion," *Journal of Development Economics*, 2020, *143* (August 2019).

Bonaccio, Silvia and Reeshad S. Dalal, "Advice taking and decision-making: An integrative literature review, and implications for the organizational sciences," *Organizational Behavior and Human Decision Processes*, 2006, *101* (2), 127–151.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer, "Beliefs about Gender," *American Economic Review*, 2019, *109* (3), 739–73.

Chandrasekhar, Arun G, Benjamin Golub, and He Yang, "Signaling, Shame, and Silence in Social Learning," *NBER Working Paper Number 25169*, 2018.

Chandrasekhar, Arun G., Horacio Larreguy, and Juan Pablo Xandri, "Testing Models of Social Learning on Networks: Evidence From Two Experiments," *Econometrica*, 2020, *88* (1), 1–32.

Chen, Jingnan and Daniel Houser, "Gender Composition, Stereotype and the Contribution of Ideas," *Working Paper*, 2017.

Coffman, Katherine Baldiga, "Evidence on Self-Stereotyping and the Contribution of Ideas," *The Quarterly Journal of Economics*, 2014, *129* (4), 1625–1660.

Coffman, Katherine, Clio Bryant Flikkema, and Olga Shurchkov, "Gender Stereotypes in Deliberation and Team Decisions," *Working Paper*, 2021.

_ **, Manuela Collis, and Leena Kulkarni**, "Stereotypes and Belief Updating," *Working Paper*, 2021.

Cooper, David J and John H Kagel, "Are two heads better than one? Team versus individual play in signaling games," *American Economic Review*, 2005, *95* (3), 477–509.

— **and —** , "A failure to communicate: an experimental investigation of the effects of advice on strategic play," *European Economic Review*, 2016, *82*, 24–45.

D'Acunto, Francesco, Ulrike Malmendier, Juan Ospina, and Michael Weber, "Exposure to Grocery Prices and Inflation Expectations," *Journal of Political Economy*, 2021, *129* (5).

Danz, David, Lise Vesterlund, and Alistair J Wilson, "Belief Elicitation: Limiting Truth Telling with Information on Incentives," *NBER Working Paper Number 27327*, 2020.

De Filippis, Roberta, Antonio Guarino, Jehiel Philippe, and Toru Kitagawa, "Updating Ambiguous Beliefs in a Social Learning Experiment," *Working Paper*, 2017.

Drehmann, Mathias, Jörg Oechssler, and Andreas Roider, "Herding and Contrarian Behavior in Financial Markets: An Internet Experiment," *American Economic Review*, 2005, *95* (5), 1403–1426.

Eckel, Catherine C. and Philip J. Grossman, "Chapter 113: Men, Women and Risk Aversion: Experimental Evidence," *Handbook of Experimental Economics*, 2008, *1*, 1061–1073.

Enke, Benjamin, "What you see is all there is," *The Quarterly Journal of Economics*, 2020, *135* (3), 1363–1398.

— **and Florian Zimmermann**, "Correlation neglect in belief formation," *The Review of Economic Studies*, 2019, *86* (1), 313–332.

Exley, Christine L and Judd B Kessler, "The Gender Gap in Self-Promotion," *NBER Working Paper Number 26345*, 2021.

Exley, Christine L., Muriel Niederle, and Lise Vesterlund, "Knowing When to Ask: The Cost of Leaning In," *Journal of Political Economy*, 2020, *128* (3), 816–854.

Eyster, Erik, Matthew Rabin, and Georg Weizsacker, "An experiment on social mislearning," *Available at SSRN 2704746*, 2015.

Fehr, Dietmar, Johanna Mollerstrom, and Ricardo Perez-Truglia, "Your Place in the World: The Demand for National and Global Redistribution," Technical Report, National Bureau of Economic Research 2019.

Gigone, Daniel and Reid Hastie, "The common knowledge effect: Information sharing and group judgment.," *Journal of Personality and social Psychology*, 1993, *65* (5), 959.

Goeree, Jacob K., Thomas R. Palfrey, Brian W. Rogers, and Richard D. McKelvey, "Self-Correcting Information Cascades," *The Review of Economic Studies*, 2007, *74* (3), 733–762.

Guarino, Antonio and Philippe Jehiel, "Social Learning with Coarse Inference," *American Economic Journal: Microeconomics*, 2013, *5* (1), 147–74.

Hartzmark, Samuel M, Samuel Hirshman, and Alex Imas, "Ownership, Learning, and Beliefs," *Working Paper*, 2021.

Kahneman, Daniel, Jack L Knetsch, and Richard H Thaler, "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias," *Journal of Economic Perspectives*, 1991, *5* (1), 193–206.

Liberman, Varda, Julia A. Minson, Christopher J. Bryan, and Lee Ross, "Naïve realism and capturing the "wisdom of dyads"," *Journal of Experimental Social Psychology*, 2012, *48* (2), 507–512.

Lowe, Matt and Madeline Mckelway, "Bargaining Breakdown: Intra-Household Decision-Making and Female Labor Supply," *Working Paper*, 2019.

Malmendier, Ulrike and Stefan Nagel, "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?," *The Quarterly Journal of Economics*, 2011, *126* (1), 373–416.

March, Christoph and Anthony Ziegelmeyer, "Altruistic Observational Learning," *Journal of Economic Theory*, 2020, *190*, 105–123.

Mengel, Friederike, Jan Sauermann, and Ulf Zölitz, "Gender Bias in Teaching Evaluations," *Journal of the European Economic Association*, 2019, *17* (2), 535–566.

Minson, Julia A., Jennifer S. Mueller, and Richard P. Larrick, "The Contingent Wisdom of Dyads: When Discussion Enhances vs. Undermines the Accuracy of Collaborative Judgments," *Management Science*, 2018, *64* (9), 4177–4192.

_ , Varda Liberman, and Lee Ross, "Two to tango: Effects of collaboration and disagreement on dyadic judgment," *Personality and Social Psychology Bulletin*, 2011, *37* (10), 1325–1338.

Mobius, Markus and Tanya Rosenblat, "Social Learning in Economics," *Annual Review of Economics*, 2014, *6* (1), 827–847.

_ , Tuan Phan, and Adam Szeidl, "Treasure Hunt: Social Learning in the Field," *NBER Working Paper Number 21014*, 2015.

Niederle, Muriel and Lise Vesterlund, "Gender and Competition," *Annual Review of Economics*, 2011, *3*, 601–630.

Palfrey, Thomas R. and Stephanie W. Wang, "On Eliciting Beliefs in Stategic Games," *Journal of Economic Behavior and Organization*, 2009, *71* (2), 98–109.

Pollak, Robert A., "How Bargaining in Marriage Drives Marriage Market Equilibrium," *Journal of Labor Economics*, 2019, *37* (1), 297–321.

Reshidi, Pëllumb, "An Experiment on Social Learning with Information Sequencing," *Working Paper*, 2020.

Ross, Lee, David Greene, and Pamela House, "The "false consensus effect": An egocentric bias in social perception and attribution processes," *Journal of Experimental Social Psychology*, 1977, *13* (3), 279–301.

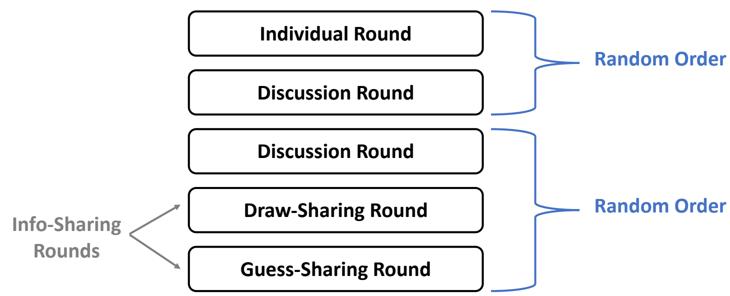
Stasser, Garold and William Titus, "Pooling of unshared information in group decision making: Biased information sampling during discussion.," *Journal of personality and social psychology*, 1985, *48* (6), 1467.

Stone, Daniel F. and Basit Zafar, "Do we follow others when we should outside the lab? Evidence from the AP Top 25," *Journal of Risk and Uncertainty*, 2014, *49*, 73–102.

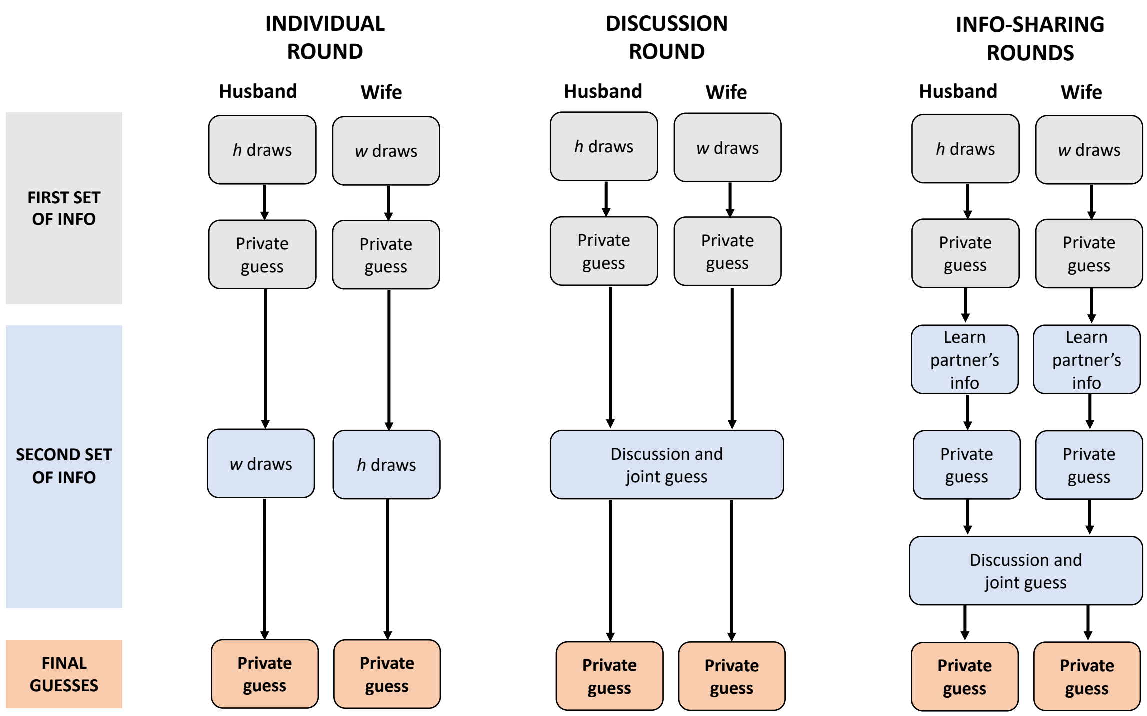
Weizsäcker, Georg, "Do We Follow Others When We Should? A Simple Test of Rational Expectations," *American Economic Review*, 2010, *100* (5), 2340–60.

Figure 1: Experimental Design

Panel A: Overall Design



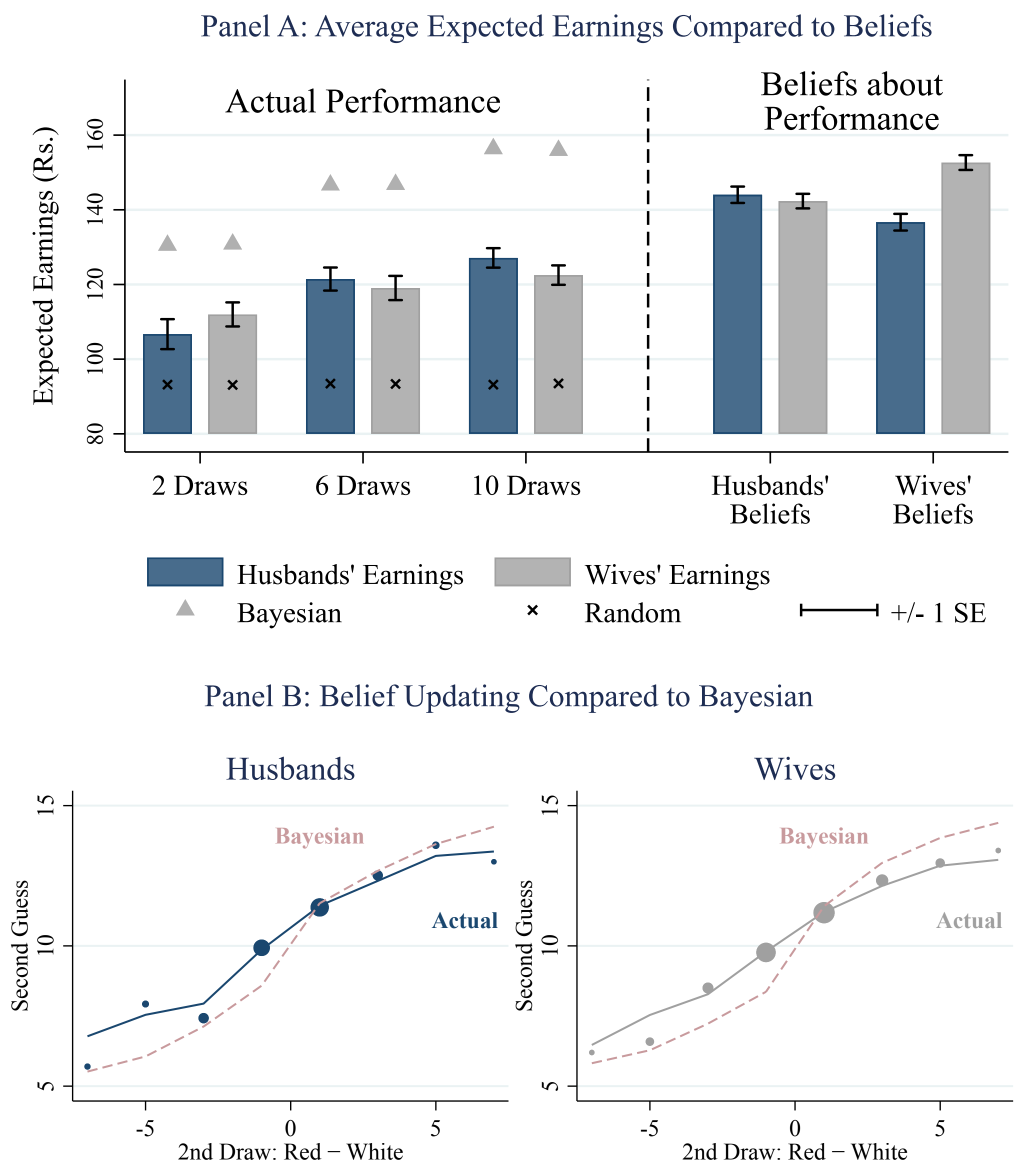
Panel B: Structure of Individual, Discussion, and Info-Sharing Rounds



Panel A shows the five rounds of the couples experiment. All couples complete all five rounds. The order of the first two rounds (*Individual* and *Discussion*) is randomized. Similarly, the order of the following three rounds (*Discussion*, *Guess-Sharing*, and *Draw-Sharing*) is randomized.

Panel B describes the structure of the different rounds. In the *Individual* round, each spouse gets two sets of private draws from the urn and makes a private guess after each set of draws. The *Individual* round is the only round with no discussion. In the *Discussion* round, each spouse makes one set of draws followed by a private guess. The two spouses are then asked to discuss and make a joint guess. Next, each spouse makes a final private guess. Both versions of the *Info-Sharing* rounds are identical to the *Discussion* round, except for that they include additional information sharing before the discussion and joint guess. In the *Draw-Sharing* round, each spouse is informed about their partner's draws earlier in the round, and then asked to make a private guess. In the *Guess-Sharing* round, each spouse is informed about their partner's guess earlier in the round, and then asked to make a private guess.

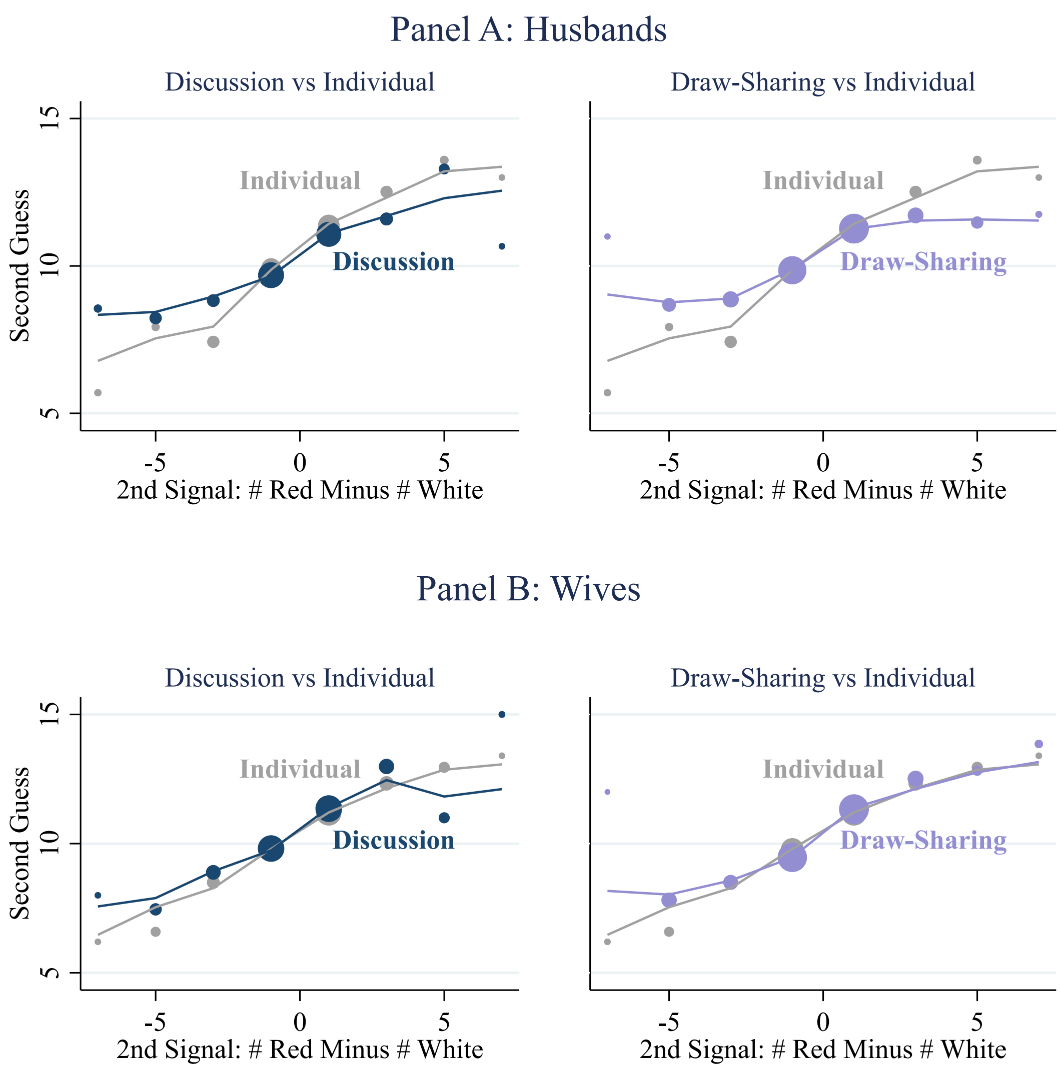
Figure 2: Actual and Perceived Performance



Panel A shows spouses' actual and their perceived performance in the game. The left panel shows the average expected earnings of the final guesses in the *Individual* round by the total number of draws in the round. Blue and gray bars indicate the mean expected earnings for husbands and wives, respectively. Triangles indicate the mean expected earnings for Bayesian risk-neutral optimal guesses conditional on the same draws, and crosses indicate the mean expected earnings if guessed randomly (i.e., uniformly among all possible guesses). Bands show ± one standard error. The right panel shows spouses' predictions of how much their own and their spouse's guesses would earn on average. These predictions were incentivized by a Rs. 50 reward for being within Rs. 30 of the actual average. Blue and grey bars show spouses' predictions of husbands' and wives' average expected earnings, respectively.

Panel B shows the average second private guess in the *Individual* round depending on the net number of red draws (i.e., red draws minus white draws) in participants' second signal. The blue and grey curves show locally-weighted means (lowess) for husbands and wives, respectively. The dotted lines show the average of what a risk-neutral Bayesian would have guessed given the same signals.

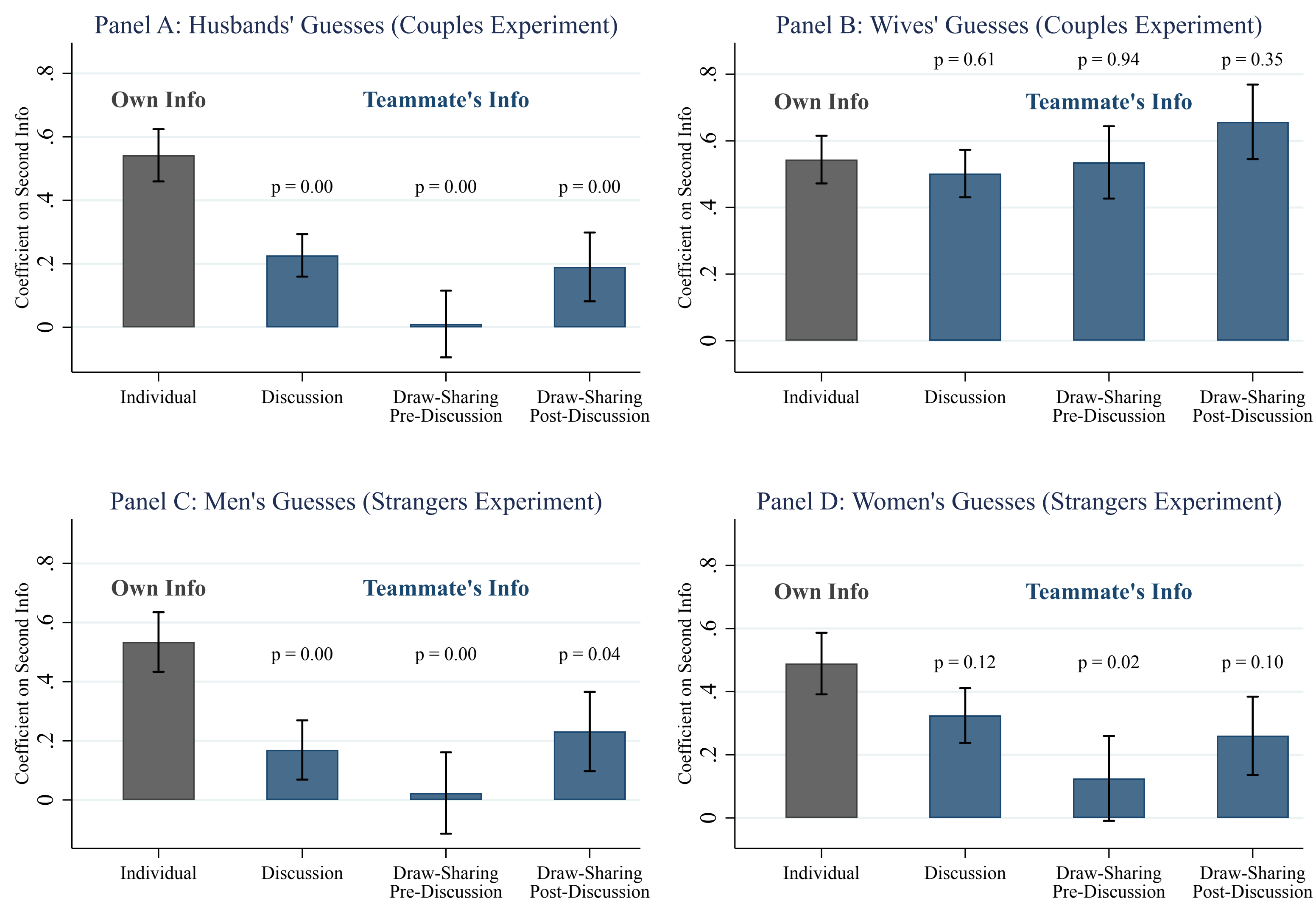
Figure 3: Spouses' Guesses in Individual, Discussion, and Draw-Sharing Rounds



Notes: This figure shows the average second private guess of husbands (Panel A) and wives (Panel B).

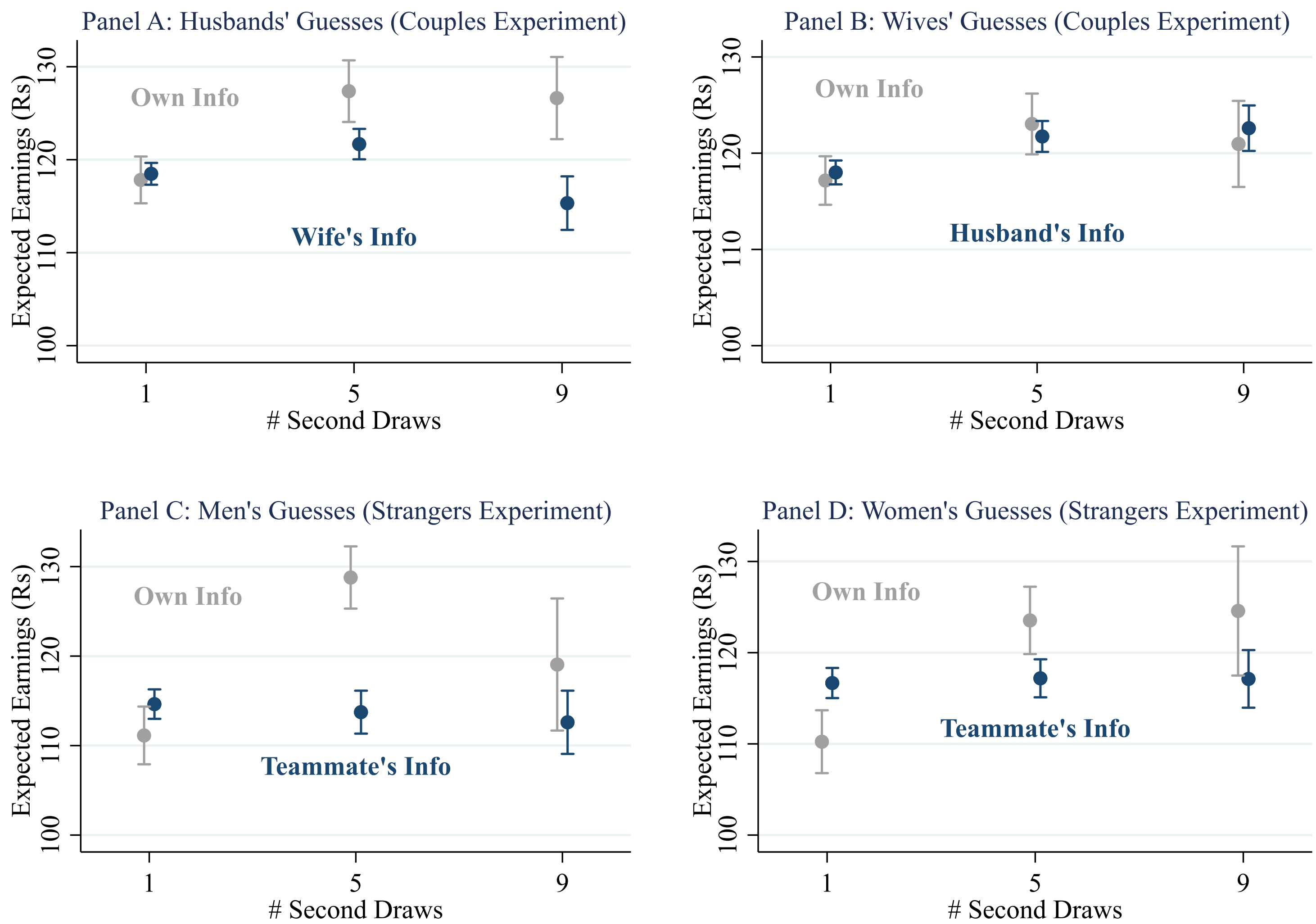
- The x-axis denotes the net number of red draws (i.e. the number of red draws minus the number of white draws) in the second signal of the round.
- The gray dots indicate average guesses in the *Individual* Round, where participants made the second set of draws themselves.
- The dark-blue dots in the graphs on the left indicate guesses in the *Discussion* Round, where the second set of draws had to be communicated to the participant via discussion.
- The lavender dots in the graphs on right indicate average guesses in the *Draw-Sharing* round, after the respondent is told of his/her spouse's draws by the experimenter (but before the joint discussion).

Figure 4: Weights on Own vs. Others' Information



Notes: This figure shows the weights participants put on different pieces of information. The upper two panels show weights for husbands (*Panel A*) and wives (*Panel B*) in the couples experiment; the lower two panels show weights for men (*Panel C*) and women (*Panel D*) in the strangers experiment. For each of these groups, we estimate equation 1 and then display the sum of $\beta_2 + \beta_{3t}$ for each of the following four types of private guesses: (a) *Individual*, where participants collect all information on their own; (b) *Discussion*, in which participants collect the first set of information on their own and the second set is only accessible via discussion; (c) *Draw-Sharing (pre-discussion)*, where participants receive the second set of information directly from the experimenter but before any discussion with their teammate; (d) *Draw-Sharing (post-discussion)*, in which participants receive the second set of information directly *and* have the chance to discuss it with their teammate. For each of the dark-blue bars, we show the p-value of testing whether the weight in that round equals the corresponding weight in the *Individual* round (grey bar).

Figure 5: Average Expected Earnings by Number of Draws



Notes: This figure shows the average expected earnings (in rupees) of the second private guesses depending on the number of draws in the second signal. The figure shows expected earnings for husbands (Panel A) and wives (Panel B) in the couples sample and by men (Panel C) and women (Panel D) in the strangers sample. The x-axis denotes the number of draws in the second signal of the round. Gray dots indicate expected earnings of final private guesses in the *Individual* Round, where participants made the second set of draws themselves. Dark-blue dots in the graphs on the left indicate earnings from final private guesses pooling all guesses in which the second set of draws had to be communicated to the participant via discussion, i.e. the *Discussion* round (both pre and post discussion) and the *Draw-sharing* round. Bands show ± one standard error.

Table 1: Sample Characteristics

	Couples		**Strangers**	
	Husbands	Wives	Men	Women
Marriage & Age				
Married	1.00	1.00	0.56	0.85
Years married | Married	12.33	12.23	13.00	15.09
	(8.47)	(8.45)	(7.65)	(8.66)
Age	36.46	31.86	34.92	34.39
	(9.10)	(8.34)	(8.69)	(8.48)
Education				
Highest grade attended	7.86	8.11	7.77	7.26
	(3.31)	(3.29)	(3.54)	(3.44)
Read Tamil	0.86	0.83	0.77	0.75
Multiplied correctly	0.48	0.33	0.52	0.36
Work				
Works (at least 1 day/week)	1.00	0.42	1.00	0.54
Daily work hours | Works	8.23	5.56	7.93	4.40
	(2.74)	(3.61)	(3.18)	(3.65)
Days working per week | Works	5.73	5.90	5.27	5.75
	(1.05)	(1.15)	(1.26)	(1.31)
Daily earnings (in Rs.) | Works	571	280	577	282
	(269)	(196)	(300)	(210)
Ability at task				
Actual ability (exp. earnings in Rs.)	122	120	117	119
	(37)	(36)	(37)	(38)
Belief of own ability (in Rs.)	144	137	139	144
	(44)	(45)	(45)	(46)
Belief of partner's ability (in Rs.)	142	153	123	123
	(39)	(40)	(47)	(53)
Who in general is better at the task?				
Men	0.21	0.22	0.13	0.14
Women	0.40	0.39	0.27	0.26
About the same	0.39	0.39	0.60	0.59
Number of participants	397	399	250	250

Notes: This table shows averages of key background characteristics for the couples and strangers samples. Standard deviations for non-binary variables are in parentheses. Columns 1 and 2 describe our main experimental sample of 400 married couples; columns 3 and 4 describe our secondary sample of 500 individuals. "Highest grade attended" refers to the highest school grade attended out of 12. Tamil is the local language. "Multiplied correctly" equals 1 if the participant knew the answer to "What is 3×9?" " | " means "conditional on". Earnings are in Indian Rupees (US\$1 $\approx$ 70 Rupees). Actual ability refers to the expected earnings of participants' final guesses in the *Individual* round. Four people in the couples sample did not complete the demographic survey at the end of the experiment, so are excluded from this table.

Table 2: Couples: Reduced-Form and Structural Estimates

	Reduced-Form Coefficients			**Structural Parameters**		
	Pooled	Husbands	Wives	Pooled	Husbands	Wives
	(1)	(2)	(3)	(4)	(5)	(6)
β_1: First Info	0.43***	0.50***	0.37***	0.82***	0.82***	0.71***
	(0.03)	(0.05)	(0.05)	(0.16)	(0.16)	(0.26)
β_2: Second Info	0.53***	0.54***	0.54***	1.37***	1.29***	1.32***
	(0.05)	(0.08)	(0.07)	(0.27)	(0.28)	(0.44)
$\beta_{3,1}$: Second Info X Discussion	-0.20***	-0.32***	-0.04	-0.73***	-1.00***	-0.33
	(0.06)	(0.08)	(0.08)	(0.21)	(0.24)	(0.38)
$\beta_{3,2}$: Second Info X Draw-Sharing (pre-discussion)	-0.28***	-0.53***	-0.01	-0.99***	-1.32***	-0.31
	(0.09)	(0.12)	(0.12)	(0.26)	(0.31)	(0.55)
$\beta_{3,3}$: Second Info X Draw-Sharing (post-discussion)	-0.13	-0.35***	0.11	-0.29	-0.86**	0.44
	(0.09)	(0.12)	(0.12)	(0.29)	(0.37)	(0.58)
α: Constant/ Structural Noise Parameter	10.29***	10.14***	10.39***	67.63***	60.82***	71.46***
	(0.12)	(0.15)	(0.15)	(4.72)	(5.20)	(5.61)
Observations	4000	2000	2000	4000	2000	2000
p-value: $\beta_{3,1}$ equal across genders		0.02			0.14	
p-value: $\beta_{3,2}$ equal across genders		0.00			0.13	
p-value: $\beta_{3,3}$ equal across genders		0.00			0.08	
Includes Info X Order FEs	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table show reduced-from and structural estimates for the couples sample's *Discussion* and *Draw-sharing* rounds.

Reduced-form coefficients: Columns 1 to 3 show reduced-form results, estimating equation 1 by OLS. The dependent variable is participants' private guess. "First Info" indicates the net number of red draws (i.e., red draws minus white draws) in the first set of signals. Similarly, "Second Info" indicates the net number of red draws in the second set of draws. "Discussion" is an indicator that equals one for the final private guess in the *Discussion* round, when the second set of draws were drawn by the participant's spouse and then (potentially) communicated to her through discussion. "Draw-Sharing (pre-discussion)" indicates the second private guess in the *Draw-Sharing* round, after the participant was directly told their spouse's information but before discussing it with their spouse. "Draw-Sharing (post-discussion)" indicates the third private guess in the *Draw-Sharing* round, after the discussion. All regressions include order fixed effects interacted with "First Info" and "Second Info." Standard errors are clustered at the couple level. *, **, and *** indicate significance at the $p<0.10$, 0.05, and 0.01 levels.

Structural parameters: Columns 4 to 6 show estimates of the structural model described in Section 4.3. "First Info" and "Second Info" indicate the weights placed on the first and second set of signals in the agents' quasi-Bayesian updating rule. The interaction terms indicate the difference in weight on the second signal in *Discussion* and *Draw-Sharing* rounds (see notes for reduced-form estimates for more detail). Bootstrapped standard errors (clustered at the couple level) in parentheses. *, **, and *** indicate significance at the $p<0.10$, 0.05, and 0.01 levels.

Table 3: Couples' Expected Earnings by Type of Guess and Number of Draws

	Pooled (1)	Husbands (2)	Wives (3)	Pooled (4)	Husbands (5)	Wives (6)
Discussion	-2.48	-4.47*	-0.49	0.01	-0.59	0.43
	(2.06)	(2.71)	(2.83)	(2.94)	(3.83)	(4.20)
Draw-Sharing (pre-discussion)	-4.92	-8.32**	-1.52	4.03	3.47	4.24
	(3.03)	(3.90)	(4.02)	(3.89)	(5.01)	(5.33)
Draw-Sharing (post-discussion)	-0.44	-4.22	3.33	2.98	3.64	2.06
	(3.16)	(3.87)	(4.02)	(3.76)	(4.90)	(5.53)
# First Draws				2.50***	2.62***	2.36***
				(0.31)	(0.38)	(0.41)
# Second Draws				2.54***	3.26***	1.77**
				(0.53)	(0.71)	(0.76)
# Second Draws X Discussion				-0.66	-1.04	-0.24
				(0.59)	(0.81)	(0.85)
# Second Draws X Draw-Sharing (pre-discussion)				-2.33***	-3.18***	-1.44
				(0.75)	(1.03)	(1.07)
# Second Draws X Draw-Sharing (post-discussion)				-0.79	-2.05*	0.45
				(0.66)	(1.05)	(1.02)
Constant	119.98***	121.08***	118.88***	100.78***	98.80***	103.19***
	(1.71)	(2.27)	(2.42)	(3.21)	(4.13)	(4.26)
Observations	4000	2000	2000	4000	2000	2000

Notes: This table compares spouses' expected earnings in the *Discussion* and *Draw-sharing* rounds to their earnings in the *Individual* round. The table shows OLS estimates of the following two equations:

$$\textit{Expected Earnings}_{irt} = \alpha + \mu_r + \mu_i + \beta \cdot \mathbf{T}_{irt} + \epsilon_{irt} \tag{4}$$

$$\begin{aligned}\textit{Expected Earnings}_{irt} = \alpha + \mu_r + \mu_i + \beta \cdot \mathbf{T}_{irt} + \gamma_1 \textit{\# First Draws}_{irt} + \gamma_2 \textit{\# Second Draws}_{irt} \\ + \gamma_3 \cdot \mathbf{T}_{irt} \cdot \textit{\# First Draws}_{irt} + \gamma_4 \cdot \mathbf{T}_{irt} \cdot \textit{\# Second Draws}_{irt} + \epsilon_{irt}\end{aligned} \tag{5}$$

where $\textit{Expected Earnings}_{irt}$ is the expected earnings from i's guess in round r and treatment t. We include round-fixed effects μ_r and individual-fixed effects μ_i in both regressions. As before, $\mathbf{T}_{irt}$ is a vector of indicators corresponding to the *Discussion*, *Draw-sharing (pre-discussion)* and *Draw-Sharing (post-discussion)* treatments. Equation 5 additionally includes $\textit{\# First Draws}_{irt}$ and $\textit{\# Second Draws}_{irt}$, which indicate the number of draws in the first and second set of signals, as well as interacting them with the vector of treatments $\mathbf{T}_{irt}$. All regressions include order fixed effects and individual fixed effects. Standard errors are clustered at the couple level. *, **, and *** indicate significance at the p<0.10, 0.05, and 0.01 levels.

Table 4: Strangers: Reduced-Form and Structural Estimates

	Reduced-Form Coefficients			Structural Parameters		
	Pooled (1)	Men (2)	Women (3)	Pooled (4)	Men (5)	Women (6)
β_1: First Info	0.51***	0.51***	0.49***	0.90***	0.79***	0.99***
	(0.05)	(0.07)	(0.07)	(0.25)	(0.23)	(0.26)
β_2: Second Info	0.51***	0.53***	0.49***	1.51***	1.27***	1.71***
	(0.06)	(0.10)	(0.10)	(0.41)	(0.45)	(0.44)
$\beta_{3,1}$: Second Info X Discussion	-0.26***	-0.37***	-0.16	-1.11***	-0.97***	-1.16***
	(0.08)	(0.12)	(0.11)	(0.30)	(0.41)	(0.40)
$\beta_{3,2}$: Second Info X Draw-Sharing (pre-discussion)	-0.43***	-0.51***	-0.36**	-1.51***	-1.22***	-1.73***
	(0.11)	(0.15)	(0.15)	(0.34)	(0.50)	(0.49)
$\beta_{3,3}$: Second Info X Draw-Sharing (post-discussion)	-0.26**	-0.30**	-0.23	-1.00***	-0.76**	-1.24**
	(0.10)	(0.15)	(0.14)	(0.36)	(0.51)	(0.51)
α: Constant/ Structural Noise Parameter	10.71***	10.73***	10.70***	69.22***	69.44***	67.59***
	(0.13)	(0.19)	(0.19)	(6.50)	(7.70)	(8.31)
Observations	2500	1250	1250	2500	1250	1250
p-value: $\beta_{3,1}$ equal across genders		0.22			1.00	
p-value: $\beta_{3,2}$ equal across genders		0.51			0.96	
p-value: $\beta_{3,3}$ equal across genders		0.72			0.94	
Includes Info X Order FEs	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows reduced-from and structural estimates for the strangers sample's *Discussion* and *Draw-sharing* rounds.

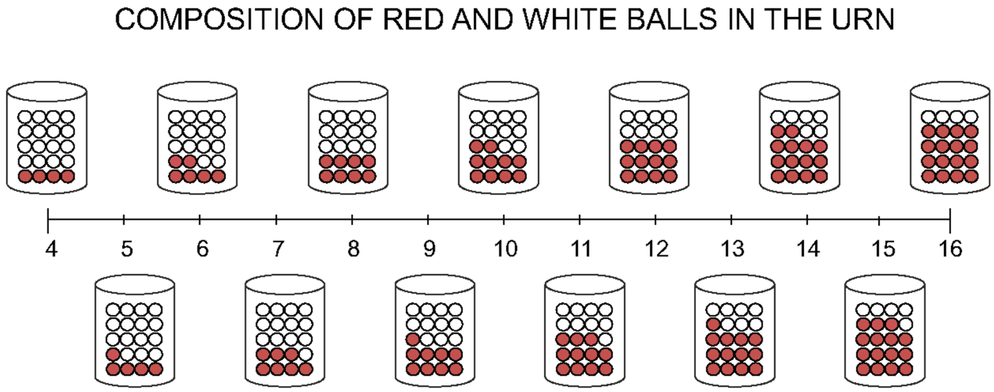
Reduced-form coefficients Columns 1 to 3 show reduced-form results, estimating equation 1 by OLS. The dependent variable is participants' private guess. "First Info" indicates the net number of red draws (i.e., red draws minus white draws) in the first set of signals. Similarly, "Second Info" indicates the net number of red draws in the second set of draws. "Discussion" is an indicator that equals one for the final private guess in the *Discussion* round, when the second set of draws were drawn by the participant's partner and then (potentially) communicated to her through discussion. "Draw-Sharing (pre-discussion)" indicates the second private guess in the *Draw-Sharing* round, after the participant was directly told their partner's information but before discussing it with them. "Draw-Sharing (post-discussion)" indicates the third private guess in the *Draw-Sharing* round, after the discussion. All regressions include order fixed effects interacted with "First Info" and "Second Info." Standard errors are clustered at the couple level. *, **, and *** indicate significance at the p<0.10, 0.05, and 0.01 levels.

Structural parameters: Columns 4 to 6 show estimates of the structural model described in Section 4.3. "First Info" and "Second Info" indicate the weights placed on the first and second set of signals in the agents' quasi-Bayesian updating rule. The interaction terms indicate the difference in weight on the second signal in *Discussion* and *Draw-Sharing* rounds (see notes for reduced-form estimates for more detail). Bootstrapped standard errors (clustered at the couple level) in parentheses. *, **, and *** indicate significance at the p<0.10, 0.05, and 0.01 levels.

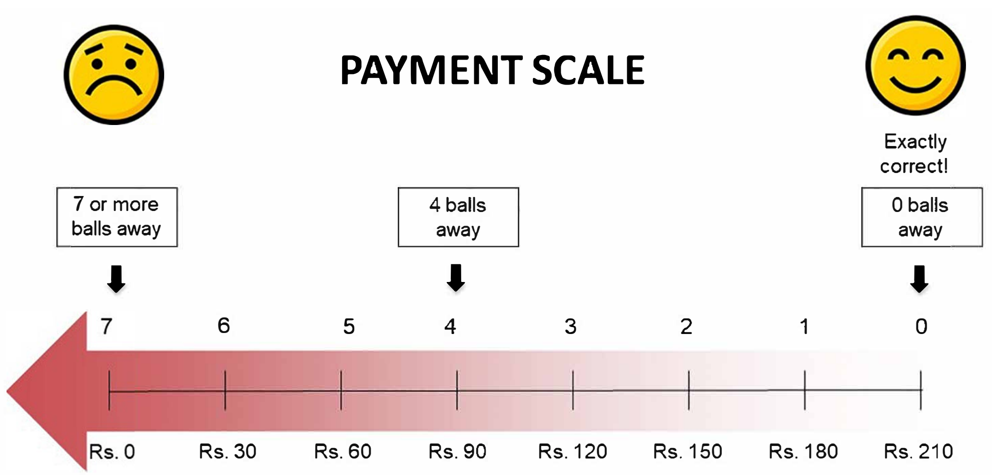
A Learning in the Household: Online Appendix

Figure A.I: Visual Aids

(a) **Guess Scale**



(b) **Payment Scale**

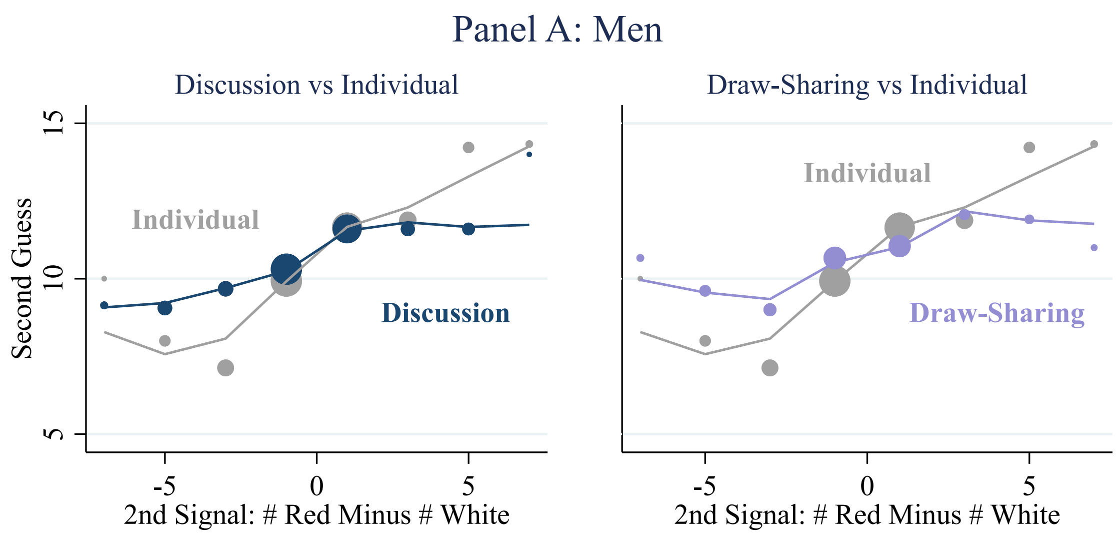


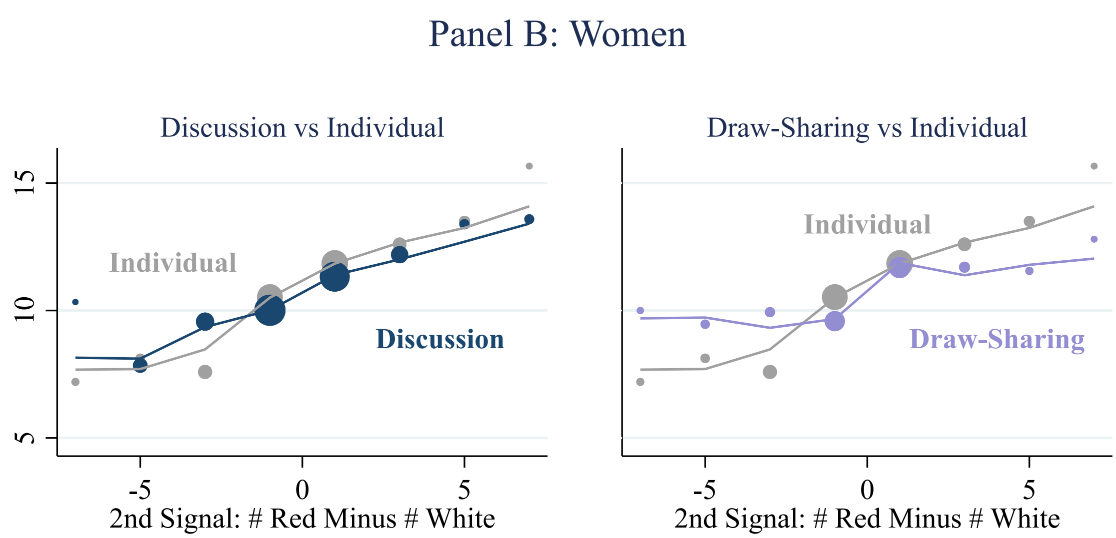
Notes: This figure uses visual used to explain the experiment to study participants.

Panel A: The figure shows the scale which participants used to make their guesses. It shows the 13 possible urn compositions ranging from 4 to 16 red balls (among 20 balls in total). We induced common priors: Participants were informed that in each round, each of these compositions were equally likely (probability 1/13 each). Participants guessed by placing a small token on top of the corresponding number.

Panel B: The figure shows the scale used to explain the incentives for accurate guessing to participants. For each pair of participants, one of their guesses was randomly selected to determine the pair's payment. On top of their participation fee, each couple receives an amount in Rupees (Rs.) equal to $\max\{(210 - 30 \times |g - r|), 0\}$, where g is the guess and r the true number of red balls for the randomly-selected guess. See more detail in Section 3.2.

Figure A.II: Strangers' Guesses in Individual, Discussion, and Draw-Sharing Rounds

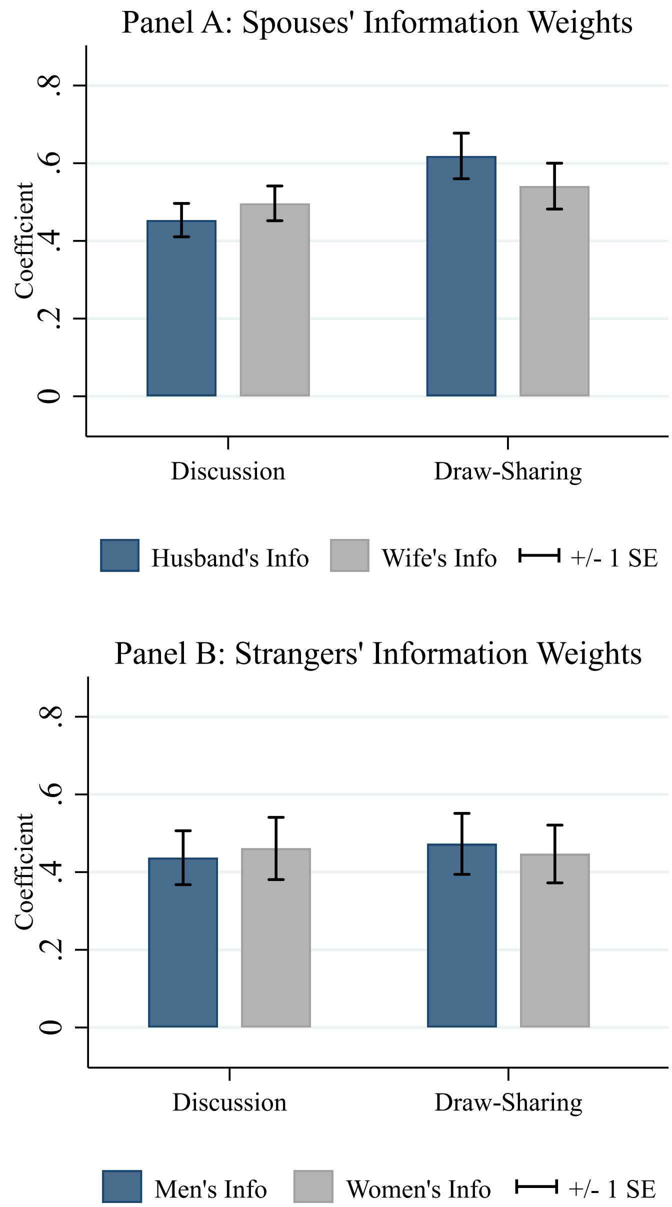




Notes: This figure shows the average second private guess of men (Panel A) and women (panel B) in the strangers sample.

- The x-axis denotes the net number of red draws (i.e. the number of red draws minus the number of white draws) in the second signal of the round.
- The gray dots indicate average guesses in the *Individual* Round, where participants made the second set of draws themselves.
- The dark-blue dots in the graphs on the left indicate guesses in the *Discussion* Round, where the second set of draws had to be communicated to the participant via discussion.
- The lavender dots in the graphs on right indicate average guesses in the *Draw-Sharing* round, after the respondent is told of his/her spouse's draws by the experimenter (but before the joint discussion).

Figure A.III: Weights in Joint Decisions



Notes: Each pair of bars above shows OLS estimates of β_1 and β_2 in the following equation:

$$Joint\ Guess_{irt} = \alpha + \beta_1 \cdot Husband's\ Info_{irt} + \beta_2 \cdot Wife's\ Info_{irt} + \epsilon_{irt}$$

$Husband's\ Info_{irt}$ is defined as the number of "net red draws" (i.e., red draws minus white draws) in the husband's set of signals. $Wife's\ Info_{irt}$ is number of net red draws in the wife's set of signals. The left pairs of bars show estimates for joint guesses in the *Discussion* rounds, and the right pairs show estimate for the *Draw-Sharing* round. Panel A shows estimates for the sample of married couples, and Panel B shows estimates for the sample of mixed-gender pairs of strangers. Whiskers denote standard errors clustered at the couple/group level.

Table A.I: Transcripts of Joint Discussions: Summary Statistics

	Couples		**Non-Couples**	
	Husbands	Wives	Men	Women
Spoke First	0.51	0.49	0.48	0.52
Spoke Last	0.57	0.43	0.53	0.47
Asked for Other's Information	0.14	0.12	0.33	0.39
Explained Task to Teammate	0.04	0.03	0.00	0.00
Shared Guess	0.40	0.42	0.29	0.29
Shared Number of Draws	0.25	0.28	0.19	0.20
Shared Composition of Draws	0.27	0.32	0.22	0.24
Suggested Final Guess	0.65	0.53	0.61	0.45
Entered Guess	0.60	0.42	0.50	0.50
Length of Discussion (mins)	1.63	1.63	0.92	0.92

Notes: This table shows averages of key characteristics of the discussion among participants for the couples and non-couples samples. These variables were constructed using transcripts of the discussions between participants before the joint guesses were made.

- Columns 1 and 2 describe our main experimental sample of 400 married couples; columns 3 and 4 describe our secondary sample of 500 individuals.
- For the couples sample, we pool the discussions across 4 rounds, and exclude the *Individual* round. For the non-couples sample, we pool the discussions across 3 rounds, and exclude the *Individual* round and *Discussion* round with same-gender pairs.
- "Shared Number of Draws" equals 1 if participants shared their total draws or mentioned the specific composition of their draws, ("I drew 4 red balls and 1 white ball"). "Shared Composition of Draws" equals 1 if participants shared the specific composition of draws ("I drew 4 red balls and 1 white ball") or mentioned that they drew more of one color ("I drew more red balls than white"). "Entered Guess" refers to the participant who placed the token on the sheet (A.I) as the joint guess.
- 83% of our transcripts were audible, so the remaining have been excluded from this table.

Table A.II: Comprehension and Memory

Question	Couples		Strangers	
	Husbands	Wives	Men	Women
A. Basic Design				
Number of balls	0.95	0.97	0.98	0.96
Colors of balls	1.00	0.99	1.00	1.00
B. Common Prior				
Possible $<$ 4 red	0.92	0.93	0.92	0.94
Possible $>$ 16 red	0.95	0.94	0.94	0.93
Who chooses number of red balls	0.84	0.87	0.79	0.83
Likelihood of each number	0.85	0.87	0.78	0.79
C. Signals				
Learn more from more balls	0.90	0.93	0.85	0.92
Possible have 4 draws	0.76	0.80	0.73	0.80
How number draws differs	0.55	0.58	0.46	0.49
How spouse's draws differ	0.63	0.65	0.57	0.66
D. Incentives				
Payment if 1 off	0.92	0.89	0.90	0.91
Payment if way off	0.89	0.85	0.85	0.86
Payment if 4 off	0.92	0.89	0.91	0.92
E. Memory				
Correctly remembered own guess	0.94	0.95	0.92	0.92
Correctly remembered # of own draws	0.99	0.97	0.96	0.98
Correctly remembered # of own red draws	0.84	0.86	0.85	0.85
F. Memory of Draws (from follow-up sample)				
Correctly remembered # of own draws			0.96	0.97
Correctly remembered # of own red draws			0.78	0.85
Correctly remembered # of partner's draws			0.89	0.90
Correctly remembered # of partner's red draws			0.72	0.64

Notes: This table shows summary statistics of participants comprehension of the task and their memory of previous draws and guesses. Cols 1 and 2 show our main experimental sample of 400 married couples; cols 3 and 4 show our secondary sample of 500 individuals. The questions asked were as follows:

- **Panel A:** shows answers to questions "How many balls are in the urn?" (correct answer: 20), and "What colors are the balls?" (red and white).
- **Panel B:** "Is it possible to have less than 4 /more than 16 red balls?" (no); "Who chooses how many balls are red?" (the computer), and "Are some numbers more likely than others?" (no).
- **Panel C:** "Do you learn more from one draw or five draws?" (five); "Can you get exactly 4 draws in any round?" (no); "Will you have the same or different numbers of draws across rounds?" (could be same or different); "Will your partner have the same or different number to you?" (could be same or different).
- **Panel D:** shows the fraction of people who could correctly indicate their payment on the scale if their guess was 1, 11, or 4 balls off.
- **Panel E:** shows the proportion of participants who correctly remember their own guess and draws in the *Guess-Sharing* and *Draw-Sharing* rounds when these questions were asked.
- **Panel F:** shows the fraction of people who correctly remember their own and their partner's draws in the third experimental round (Appendix A.2). "Correctly remembered # of own draws" and "Correctly remembered # of own red draws" correspond to results pooled across 4 rounds, including the *In-Person* round. "Correctly remember # of partner's draws" and "Correctly remembered # of partner's red draws" correspond to the *In-Person* round.

Table A.III: Explaining Differences Between Husbands and Wives

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Own Net Red	0.56***	0.67***	0.56***	0.59***	0.60***	0.56***	0.78***	0.61***
	(0.03)	(0.09)	(0.04)	(0.04)	(0.05)	(0.03)	(0.11)	(0.05)
Spouse's Net Red	0.31***	0.21*	0.34***	0.28***	0.25***	0.31***	0.13	0.32***
	(0.04)	(0.10)	(0.04)	(0.04)	(0.05)	(0.04)	(0.11)	(0.05)
Spouse's Net Red X Guesser Is Wife	0.16**	0.25*	0.16**	0.15**	0.13**	0.15**	0.25*	0.18*
	(0.05)	(0.10)	(0.05)	(0.05)	(0.05)	(0.05)	(0.10)	(0.07)
Spouse's Net Red X Guesser Is Older		0.11					0.13	
		(0.10)					(0.10)	
Spouse's Net Red X Guesser Thinks Sole HHDM			-0.09				-0.09	
			(0.05)				(0.05)	
Spouse's Net Red X Spouse Better				0.06			0.05	
				(0.05)			(0.05)	
Spouse's Net Red X Guesser Thinks Spouse Better					0.10		0.10	
					(0.06)		(0.06)	
Spouse's Net Red X Guesser Comprehension Index						0.05	0.05	
						(0.03)	(0.03)	
Spouse's Net Red X Public Discussion								0.04
								(0.07)
Spouse's Net Red X Public Discussion X Guesser Is Wife								-0.09
								(0.10)
Constant	10.56***	10.56***	10.56***	10.56***	10.55***	10.55***	10.55***	10.58***
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)
N	3200	3200	3200	3200	3200	3200	3200	2400

Notes: This table shows OLS regressions where the dependent variable is a person's individual guess.

- Regressions use data from the couples sample. These regressions stack the post-discussion guesses from both *Discussion* rounds and the pre- and post-discussion guesses in the *Draw-sharing* round, except for column 8, which does not include the pre-discussion guesses.
- All columns include an interaction between Partner's Net Red and an indicator for the guesser being a wife. Columns 2 to 7 add interactions between Partner's Net Red and other variables, to test whether heterogeneity along these variables explains the difference between wives and others. Column 8 tests whether the discussion being public affects the husband's or wife's weights.
- In each column we also control for the corresponding interactions between Own Net Red and the variables in that column (coefficients not shown).
- "Guesser Is Older" means the person guessing is older than their partner. "Guesser Thinks Sole HHDM" means the guesser considers themselves the sole household decision maker. "Partner Better" indicates that the guesser's partner earns more from their guesses on average; "Guesser Thinks Partner Better" indicates that the guesser thinks their partner earns more from their guesses on average. "Comprehension index" is the fraction of comprehension questions the guesser answered correctly, normalized by subtracting the mean and dividing by the standard deviation of the entire sample. "Public Discussion" means that the experimenter was present while the couple held their discussion.

Table A.IV: Comparing Mixed- and Same-Gender Pairs of Strangers

	Reduced-Form Coefficients			**Structural Parameters**		
	Pooled (1)	Men (2)	Women (3)	Pooled (4)	Men (5)	Women (6)
β_1: First Info	0.51***	0.51***	0.49***	0.86***	0.72***	0.99***
	(0.05)	(0.07)	(0.07)	(0.18)	(0.25)	(0.24)
β_2: Second Info	0.51***	0.53***	0.48***	1.43***	1.14***	1.72***
	(0.06)	(0.10)	(0.10)	(0.29)	(0.46)	(0.43)
$\beta_{3,1}$: Second Info X Discussion	-0.29***	-0.39***	-0.19	-1.05***	-0.84**	-1.13**
	(0.08)	(0.13)	(0.12)	(0.28)	(0.43)	(0.41)
$\beta_{3,2}$: Second Info X Discussion X Same-Gender Pair	0.06	0.06	0.06	0.01	-0.07	-0.06
	(0.07)	(0.10)	(0.11)	(0.16)	(0.29)	(0.26)
α: Constant/ Structural Noise Parameter	10.71***	10.73***	10.69***	66.15***	63.59***	67.76***
	(0.13)	(0.19)	(0.19)	(6.33)	(8.81)	(8.54)
Observations	1500	750	750	1500	750	750
Includes Info X Order FEs	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows reduced-form and structural estimates for strangers' choices.

Reduced-form coefficients: Cols 1 to 3 show reduced-form results for the strangers sample, estimating by OLS a variant on equation 1. In particular, it shows estimates of

$$\begin{aligned} Guess_{irt} = \alpha + \beta_1 \cdot First\ Info_{irt} + \beta_2 \cdot Second\ Info_{irt} \\ + \beta_{3,1} \cdot Discussion_{irt} + \beta_{3,2} \cdot Discussion_{rt} \cdot Same\text{-}Gender\ Pair_{irt} \end{aligned}$$

These regressions include only the *Discussion* and *Individual* rounds. "First Info" indicates the net number of red draws (i.e., red draws minus white draws) in the first set of signals. Similarly, "Second Info" indicates the net number of red draws in the second set of draws. "Discussion" is an indicator that equals one for the final private guess in the *Discussion* round, when the second set of draws were drawn by the participant's partner and then (potentially) communicated to them through discussion. "Same-Gender Pair" is an indicator for whether the participant's teammate is of the same gender as him/her. All regressions include order fixed effects interacted with "First Info" and "Second Info." Standard errors are clustered at the couple level.

Structural parameters: Cols 4 to 6 show estimates of a variant of the structural model described in Section 4.3.The model is identical to that in Section 4.3 except that the weight put on the second set of signals, ω_2 takes the following form:

$$\omega_{2rt} = \beta_2 + \mu_{2r} + \beta_{3,1} \cdot Discussion_{rt} + \beta_{3,2} \cdot Discussion_{rt} \cdot Same\text{-}Gender\ Pair_{rt}$$

where "Discussion" and "Same-Gender Pair" are defined as in the reduced-form OLS specification. "First Info" and "Second Info" indicate the weights placed on the first and second set of signals in the agents' quasi-Bayesian updating rule. The interaction terms indicate the difference in weight on the second signal in *Discussion* rounds and, in addition, in *Discussion* rounds when playing with a teammate of the same gender (see notes for reduced-form estimates for more detail). Bootstrapped standard errors (clustered at the couple level) in parentheses. *, **, and *** indicate significance at the p<0.10, 0.05, and 0.01 levels.

Table A.V: Strangers' Expected Earnings by Type of Guess and Number of Draws

	Pooled (1)	Men (2)	Women (3)	Pooled (4)	Men (5)	Women (6)
Discussion	-4.22	-5.69	-2.75	4.08	-1.87	9.77*
	(2.81)	(3.90)	(3.41)	(3.88)	(5.81)	(5.28)
Draw-Sharing	-8.85**	-13.23**	-4.46	0.46	-7.42	8.34
(pre-discussion)	(3.85)	(5.19)	(4.82)	(4.69)	(6.92)	(6.07)
Draw-Sharing	-6.18	-8.62	-3.74	-4.15	-8.29	0.11
(post-discussion)	(3.93)	(5.41)	(4.75)	(4.66)	(6.86)	(6.46)
# First Draws				2.37***	2.55***	2.14***
				(0.40)	(0.52)	(0.59)
# Second Draws				2.95***	1.78	4.07***
				(0.81)	(1.13)	(1.13)
# Second Draws X				-2.32**	-1.10	-3.46***
Discussion				(0.92)	(1.41)	(1.23)
# Second Draws X				-2.68***	-1.69	-3.64***
Draw-Sharing (pre-discussion)				(0.94)	(1.40)	(1.37)
# Second Draws X				-0.73	-0.22	-1.24
Draw-Sharing (post-discussion)				(0.97)	(1.43)	(1.42)
Constant	119.57***	118.38***	120.77***	100.67***	103.02***	98.55***
	(2.29)	(3.23)	(2.89)	(4.17)	(5.47)	(5.79)
Observations	2500	1250	1250	2500	1250	1250

Notes: Using the same approach as Table 3, this table compares expected earnings in the *Discussion* and *Draw-sharing* rounds to their earnings in the *Individual* round for the strangers sample. The table shows OLS estimates of the following two equations:

$$\textit{Expected Earnings}_{irt} = \alpha + \mu_r + \mu_i + \beta \cdot \mathbf{T}_{irt} + \epsilon_{irt} \tag{6}$$

$$\begin{aligned}\textit{Expected Earnings}_{irt} = \alpha + \mu_r + \mu_i + \beta \cdot \mathbf{T}_{irt} + \gamma_1 \#\ \textit{First Draws}_{irt} + \gamma_2 \#\ \textit{Second Draws}_{irt} \\ + \gamma_3 \cdot \mathbf{T}_{irt} \cdot \#\ \textit{First Draws}_{irt} + \gamma_4 \cdot \mathbf{T}_{irt} \cdot \#\ \textit{Second Draws}_{irt} + \epsilon_{irt}\end{aligned} \tag{7}$$

where $\textit{Expected Earnings}_{irt}$ is the expected earnings from i's guess in round r and treatment t. We include round-fixed effects μ_r and individual-fixed effects μ_i in both regressions. As before, $\mathbf{T}_{irt}$ is a vector of indicators corresponding to the *Discussion*, *Draw-sharing (pre-discussion)* and *Draw-Sharing (post-discussion)* treatments. Equation 7 additionally includes $\#\ \textit{First Draws}_{irt}$ and $\#\ \textit{Second Draws}_{irt}$, which indicate the number of draws in the first and second set of signals, as well as interacting them with the vector of treatments $\mathbf{T}_{irt}$. All regressions include order fixed effects and individual fixed effects. Standard errors are clustered at the couple level. *, **, and *** indicate significance at the p<0.10, 0.05, and 0.01 levels.

Table A.VI: Understanding Differences Between Couples and Strangers

	All rounds (Pre- & Post-Disc.)		**Draw-sharing round** (Pre-Discussion)		(Post-Discussion)		**Discussion rounds** (Post-Discussion)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Own Net Red	0.53***	0.49***	0.56***	0.51***	0.49***	0.32***	0.54***	0.57***
	(0.03)	(0.04)	(0.05)	(0.07)	(0.05)	(0.08)	(0.03)	(0.05)
Teammate's Net Red	0.28***	0.24***	0.18***	0.14	0.39***	0.38***	0.28***	0.21***
	(0.03)	(0.04)	(0.05)	(0.08)	(0.05)	(0.08)	(0.04)	(0.06)
Teammate's Net Red X	0.13***	0.09	0.16*	0.07	0.11	0.00	0.12*	0.15
Guesser Is Woman	(0.04)	(0.06)	(0.07)	(0.11)	(0.07)	(0.11)	(0.05)	(0.08)
Teammate's Net Red X		0.07		0.07		0.01		0.11
Guesser Is Husband In Couple		(0.06)		(0.10)		(0.11)		(0.08)
Teammate's Net Red X		0.14*		0.22*		0.20*		0.06
Guesser Is Wife In Couple		(0.05)		(0.10)		(0.09)		(0.06)
Constant	10.67***	10.67***	10.63***	10.63***	10.76***	10.76***	10.64***	10.64***
	(0.06)	(0.06)	(0.09)	(0.09)	(0.10)	(0.10)	(0.07)	(0.07)
N	5200	5200	1300	1300	1300	1300	2600	2600

Notes: This table shows OLS regressions in which the dependent variable is a person's individual guess. All regressions pool all data from both the couples and strangers sample.

- In cols 1 and 2, the sample includes all final guesses made by both men and women in the two *Discussion* rounds plus the pre- and post-discussion guesses from the *Draw-sharing* rounds.
- Cols 3 and 4 show results for the pre-discussion *Draw-sharing* guesses alone, and cols 5 and 6 show results from the post-discussion *Draw-Sharing* guesses alone. Cols 7 and 8 show results from the *Discussion* round guesses (after the joint discussion) alone.
- "Own Net Red" refers to the net red minus white draw in the set the person guessing drew his or herself, and "Partner's Net Red" to same in the set the guesser's partner drew. "Is Woman" equals one if the guesser is a woman, and "Is Wife In Couple" equals one if the person guessing is a woman *and* is in the couples sample, i.e. is playing with her husband.
- The regressions also include controls for Own Net Red interacted with Is Woman and Is Wife In Couple (coefficients not shown).

Table A.VII: Explaining Differences Between Couples and Strangers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Own Net Red	0.49***	0.52***	0.48***	0.52***	0.55***	0.50***	0.49***	0.66***
	(0.04)	(0.05)	(0.04)	(0.04)	(0.07)	(0.06)	(0.05)	(0.08)
Teammate's Net Red	0.24***	0.20***	0.25***	0.20***	0.13	0.26***	0.24***	0.10
	(0.04)	(0.05)	(0.04)	(0.05)	(0.07)	(0.06)	(0.04)	(0.09)
Teammate's Net Red X	0.07	0.04	0.09	0.07	0.12*	0.09	0.07	0.10
Guesser Is Husband In Couple	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.07)
Teammate's Net Red X	0.09	0.09	0.11	0.08	0.09	0.10	0.10	0.10
Guesser Is Woman	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Teammate's Net Red X	0.14*	0.17**	0.14**	0.14**	0.16**	0.14*	0.13*	0.20***
Guesser Is Wife In Couple	(0.05)	(0.06)	(0.05)	(0.05)	(0.05)	(0.06)	(0.05)	(0.06)
Teammate's Net Red X		0.08						0.09
Guesser Is Older		(0.05)						(0.05)
Teammate's Net Red X			-0.08					-0.06
Guesser Thinks Sole HHDM			(0.04)					(0.04)
Teammate's Net Red X				0.07				0.06
Teammate Better				(0.04)				(0.04)
Teammate's Net Red X					0.10			0.10
Guesser Thinks Teammate Better					(0.06)			(0.06)
Teammate's Net Red X						-0.04		-0.03
Guesser Is Married						(0.06)		(0.07)
Teammate's Net Red X							0.05*	0.05*
Guesser Comprehension index							(0.02)	(0.02)
Constant	10.67***	10.67***	10.67***	10.66***	10.66***	10.67***	10.66***	10.66***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
N	5200	5200	5200	5200	5200	5200	5200	5200

Notes: This table shows OLS regressions where the dependent variable is a person's individual guess. All regressions pool data from both the couples and strangers sample, and the sample includes all final guesses made by both men and women in the two *Discussion* rounds plus the pre- and post-discussion guesses from the *Draw-sharing* rounds.

- Column 1 repeats the specification in column 2 from Table A.VI. The other columns add interactions between Partner's Net Red and other variables, to test whether heterogeneity along these variables explains the difference between wives and others.
- When we add an interaction between Partner's Net Red and a variable, we control for the corresponding interaction(s) with Own Net Red (coefficients not shown).
- "Guesser Is Older" means the person guessing is older than their partner. "Guesser Thinks Sole HHDM" means the guesser considers themselves the sole household decision maker. "Partner Better" indicates that the guesser's partner earns more from their guesses on average; "Guesser Thinks Partner Better" indicates that the guesser thinks this is so. "Guesser Is Married" indicates that the guesser is married; this is always one in the couples sample, but not in the strangers sample. "Comprehension index" is the fraction of comprehension questions the guesser answered correctly, normalized by subtracting the mean and dividing by the standard deviation of the entire sample.

Table A.VIII: Couples' Weights on Husbands' and Wives' Info in Joint Guesses

	Discussion Round						Draw-Sharing Round					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Husband's Info	0.45***	0.35***	0.45***	0.40***	0.48***	0.34***	0.62***	0.49***	0.60***	0.66***	0.60***	0.49***
	(0.04)	(0.08)	(0.05)	(0.05)	(0.06)	(0.10)	(0.06)	(0.10)	(0.09)	(0.07)	(0.08)	(0.12)
Wife's Info	0.50***	0.61***	0.58***	0.57***	0.52***	0.80***	0.54***	0.65***	0.54***	0.54***	0.47***	0.58***
	(0.04)	(0.06)	(0.06)	(0.05)	(0.06)	(0.07)	(0.06)	(0.07)	(0.07)	(0.07)	(0.10)	(0.11)
Husband's Info X		0.16*				0.13		0.20*				0.26**
Husband Better		(0.09)				(0.09)		(0.12)				(0.12)
Wife's Info X		-0.25***				-0.24***		-0.22**				-0.24**
Husband Better		(0.09)				(0.09)		(0.11)				(0.11)
Husband's Info X			0.03			0.02			0.03			0.08
Husband Says He's Better			(0.09)			(0.09)			(0.12)			(0.13)
Wife's Info X			-0.18*			-0.18**			0.02			-0.01
Husband Says He's Better			(0.09)			(0.09)			(0.13)			(0.13)
Husband's Info X				0.17**		0.14				-0.14		-0.21
Husband Says He's HHDM				(0.09)		(0.09)				(0.12)		(0.13)
Wife's Info X				-0.23**		-0.24**				0.01		0.04
Husband Says He's HHDM				(0.10)		(0.10)				(0.12)		(0.12)
Husband's Info X					-0.05	-0.04					0.04	-0.01
Surveyor Present					(0.09)	(0.08)					(0.12)	(0.12)
Wife's Info X					-0.05	-0.09					0.15	0.14
Surveyor Present					(0.09)	(0.08)					(0.12)	(0.12)
Observations	800	800	800	800	800	800	400	400	400	400	400	400
p-value: Interaction has no impact on relative information weights		0.00	0.17	0.01	0.97			0.02	0.93	0.47	0.58	

Notes: This table shows OLS regressions where the dependent variable is the joint guess in the first *Discussion* round. "Husband's Info" indicates the Bayesian risk-neutral optimal guess after observing only the husband's own set of draws. Similarly, "Wife's Info" indicates the Bayesian guess given only the wife's draws. "Husband better" is an indicator variable for whether the husband's first guesses (using only his first set of draws) have a higher expected earnings than his wife's. "Husband Says He's Better" indicates whether the husband's guess of his average earnings at the experimental task is higher than his guess of his wife's earnings. "He says He's HHDM" indicates whether the husband says that he is the primary household-decision maker. Standard errors are clustered at the couple level. *, **, and *** indicate significance at the p<0.10, 0.05, and 0.01 levels.

Table A.IX: Couples: Expected Earnings in Joint versus Private Guesses

	Joint Guess Compared to Private Guesses Made by:					
	Pooled	Husbands	Wives	Pooled	Husbands	Wives
	(1)	(2)	(3)	(4)	(5)	(6)
Private Guess	-2.17***	-2.70***	-1.72*	1.77	2.97	0.63
	(0.70)	(0.91)	(0.90)	(1.48)	(1.84)	(2.00)
# Husband's Draws				2.34***	2.49***	2.19***
				(0.42)	(0.44)	(0.44)
# Wife's Draws				2.76***	2.80***	2.65***
				(0.41)	(0.45)	(0.43)
Private Guess X # Husband's Draws				-0.27	-0.11	-0.45
				(0.25)	(0.30)	(0.34)
Private Guess X # Wife's Draws				-0.81***	-1.46***	-0.17
				(0.26)	(0.34)	(0.34)
Constant	121.32***	121.44***	120.74***	102.26***	101.69***	102.55***
	(2.32)	(2.55)	(2.55)	(3.42)	(3.73)	(3.70)
Observations	4400	2800	2800	4400	2800	2800

Notes: This tables compares expected earnings generated by couples' joint guesses and private guesses. The table shows OLS estimates of the following two equations:

$$\textit{Expected Earnings}_{irt} = \alpha + \mu_r + \mu_c + \beta \cdot \textit{Joint}_{irt} + \epsilon_{irt} \quad (8)$$

$$\begin{aligned} \textit{Expected Earnings}_{irt} = \alpha + \mu_r + \mu_c + \beta \cdot \textit{Joint}_{irt} + \gamma_1 \textit{Husband's Draws}_{irt} + \gamma_2 \textit{Wife's Draws}_{irt} \quad (9) \\ + \gamma_3 \cdot \textit{Joint}_{irt} \cdot \# \textit{ Husband's Draws}_{irt} + \gamma_4 \cdot \textit{Joint}_{irt} \cdot \# \textit{ Wife's Draws}_{irt} + \epsilon_{irt} \end{aligned}$$

where $\textit{Expected Earnings}_{irt}$ is expected earnings from i's guess in round r and treatment t. We include round-fixed effects μ_r and couple-fixed effects μ_c in both regressions. $\textit{Joint}_{irt}$ is an indicator for being a joint guess. Equation 9 additionally includes $\# \textit{ Husband's Draws}_{irt}$ and $\# \textit{ Wife's Draws}_{irt}$, indicating the number of draws the husband and wife got, respectively, and interacting these with $\textit{Joint}_{irt}$.

- We restrict regressions to treatments with both private and joint guesses (that is, excluding the *Individual* round). We include all guesses in these rounds that occur after potentially learning about both the spouse's information (i.e., all guesses except first private guesses).
- Table shows OLS regressions. All regressions include joint guesses from the *Discussion* and *Draw-Sharing* rounds and individual guesses from the *Discussion* round (after the discussion) and *Draw-Sharing* round (both the guess immediately after being told of their partner's information but before the discussion, and the guess after the discussion).
- Cols 1 and 4 include joint guesses and private guesses by both husbands and wives. Columns 2 and 5 include only joint guesses and husbands' guesses. Columns 3 and 6 include only joint guesses and wives' guesses.
- "Joint Guess" is an indicator for couples' joint guesses post discussion. "# Husband/Wife's Draws" are the number of draws that husband/wife had in their private set of signals.
- All regressions include order fixed effects and couple fixed effects. Standard errors are clustered at the group level. *, **, and *** indicate significance at the p<0.10, 0.05, and 0.01 levels.

Table A.X: Reduced Form: Including Interactions with First Information

	Couples			Strangers		
	Pooled	Husbands	Wives	Pooled	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
First Info	0.42***	0.47***	0.37***	0.45***	0.44***	0.45***
	(0.04)	(0.06)	(0.07)	(0.06)	(0.09)	(0.09)
Second Info	0.54***	0.55***	0.54***	0.49***	0.54***	0.46***
	(0.05)	(0.08)	(0.07)	(0.07)	(0.10)	(0.10)
First Info X Only Accessible via Discussion	0.02	0.07	-0.00	0.12	0.13	0.13
	(0.05)	(0.08)	(0.08)	(0.07)	(0.12)	(0.11)
First Info X Draw-Sharing (pre-discussion)	0.07	0.12	0.04	0.12	0.07	0.19
	(0.08)	(0.11)	(0.12)	(0.09)	(0.15)	(0.13)
First Info X Draw-Sharing (post-discussion)	0.01	0.13	-0.08	-0.08	-0.11	-0.03
	(0.08)	(0.11)	(0.12)	(0.09)	(0.16)	(0.14)
Second Info X Only Accessible via Discussion	-0.20***	-0.33***	-0.04	-0.28***	-0.39***	-0.18
	(0.06)	(0.08)	(0.09)	(0.08)	(0.12)	(0.11)
Second Info X Draw-Sharing (pre-discussion)	-0.30***	-0.56***	-0.02	-0.44***	-0.52***	-0.38**
	(0.08)	(0.12)	(0.12)	(0.11)	(0.15)	(0.16)
Second Info X Draw-Sharing (post-discussion)	-0.13	-0.38***	0.13	-0.23**	-0.28*	-0.21
	(0.08)	(0.12)	(0.12)	(0.10)	(0.15)	(0.14)
Constant	10.29***	10.13***	10.39***	10.71***	10.78***	10.64***
	(0.12)	(0.15)	(0.15)	(0.14)	(0.19)	(0.21)
Observations	4000	2000	2000	3000	1500	1500
Includes Info X Order FEs	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows reduced-form results, estimating equation 1 by OLS.

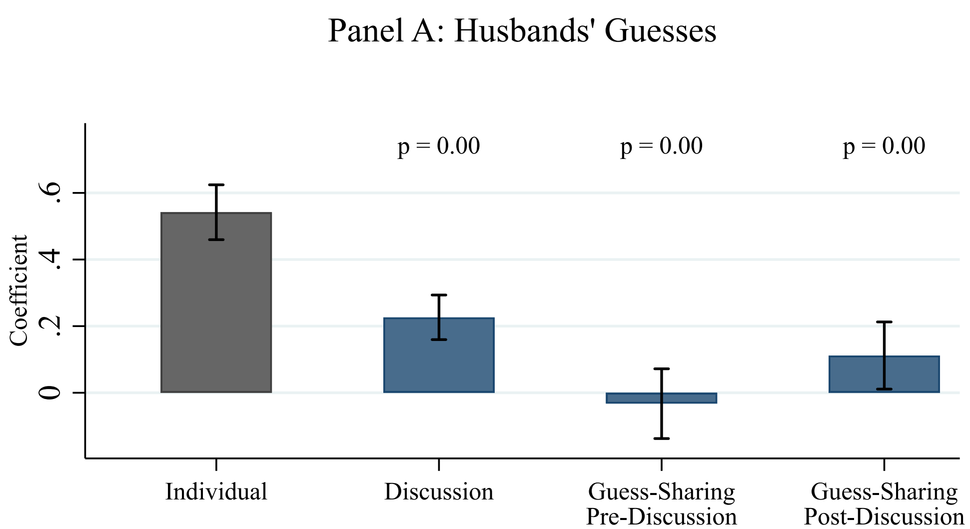
- Columns 1 to 3 include the sample of married couples, and columns 4 to 6 include the sample of pairs of strangers.
- The dependent variable is participants' private guess. "First Info" indicates the net number of red draws (i.e., red draws minus white draws) in the first set of signals. Similarly, "Second Info" indicates the net number of red draws in the second set of draws. "Discussion" is an indicator that equals one for the final private guess in the *Discussion* round, when the second set of draws were drawn by the participant's spouse and then (potentially) communicated to them through discussion.
- "Draw-Sharing (pre-discussion)" indicates the second private guess in the *Draw-Sharing* round, after the participant was directly told their spouse's information but before discussing it with their spouse. "Draw-Sharing (post-discussion)" indicates the third private guess in the *Draw-Sharing* round, after the discussion.
- All regressions include order fixed effects interacted with "First Info" and "Second Info." Standard errors are clustered at the couple or group level. *, **, and *** indicate significance at the $p < 0.10$, 0.05, and 0.01 levels.

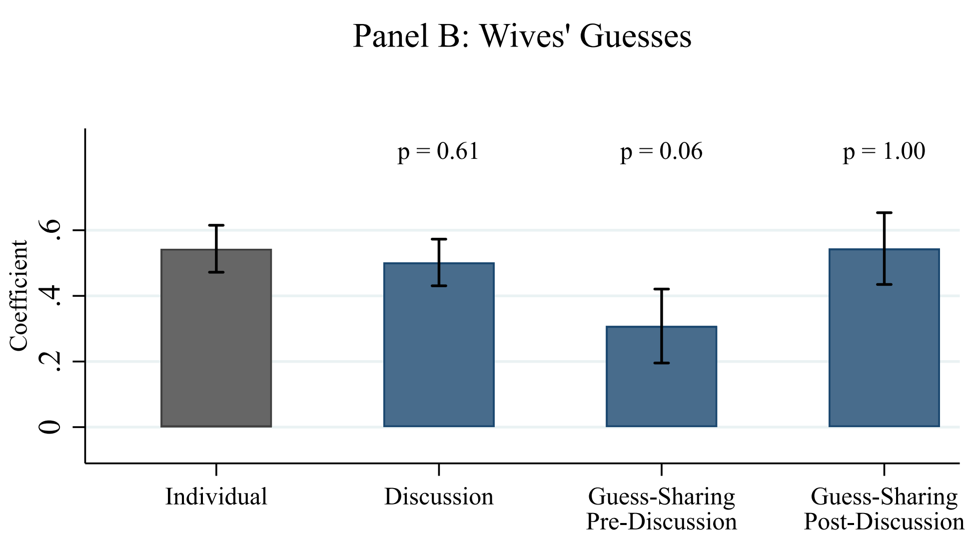
A.1 Guess-Sharing Round

The *Guess-sharing* round is identical to the *Draw-sharing* round, except that instead of sharing with each person the number of balls of each color their spouse drew, the surveyor shares their spouse's guess and the total number of draws (1, 5 or 9) on which that guess was based. Figure A.IV shows estimates for the *Individual*, *Discussion*, and *Guess-Sharing* rounds. The results look similar to those for the *Draw-Sharing* round. Husbands strongly discount their wife's information relative to their own in both the pre-discussion and post-discussion guesses. Wives appear to put less weight on husband's information in the pre-discussion *Guess-Sharing* round. This could be explained by differential processing of own compared to others' information, but also by other (potentially rational) reasons, such as mistrust of husbands' guesses or the increased computational difficulty of backing out what the husband's information must have been given his guess. Post-discussion wives put equal weight on their own and their husband's information. Table A.XI shows the corresponding reduced-form and structural estimates, which confirm the visual impressions from Figure A.IV.[33]

[33] Note that the structural estimates assume that participants are able to back out from their partner's guess what their information must have been. Less weight on partner's information could therefore reflect not just intrinsic discounting others info but also the extent to which that this is a difficult problem for participants to solve (or one they do not attempt to solve).

Figure A.IV: Couples' Weights on Own vs. Others' Info in Guess-Sharing Rounds





Notes: This figure shows the weights husbands (*Panel A*) and wives (*Panel B*) in our couples sample put on different pieces of information. We estimate coefficients using a version of equation 1 and then display the sum of $\beta_2 + \beta_{3t}$ for each of the following four types of private guesses: (a) *Individual*, where participants collect all information on their own; (b) *Discussion*, in which participants collect the first set of information on their own and the second set is only accessible via discussion; (c) *Guess-Sharing*, where participants receive the second set of information directly from the experimenter; (d) *Guess-Sharing + Discussion*, in which participants receive the second set of information directly *and* have the chance to discuss it with their partner. Bands show ± one standard error.

Table A.XI: Couples' Weights in Guess-Sharing Rounds

	Reduced-Form Coefficients			Structural Parameters		
	Pooled	Husbands	Wives	Pooled	Husbands	Wives
	(1)	(2)	(3)	(4)	(5)	(6)
β_1: First Info	0.43***	0.50***	0.37***	0.72***	0.74***	0.61***
	(0.03)	(0.05)	(0.05)	(0.14)	(0.15)	(0.19)
β_2: Second Info	0.53***	0.54***	0.54***	1.20***	1.15***	1.13***
	(0.05)	(0.08)	(0.07)	(0.24)	(0.26)	(0.38)
$\beta_{3,1}$: Second Info X	-0.20***	-0.32***	-0.04	-0.64***	-0.90***	-0.27
Discussion	(0.06)	(0.08)	(0.08)	(0.19)	(0.23)	(0.33)
$\beta_{3,2}$: Second Info X	-0.42***	-0.57***	-0.24*	-0.97***	-1.24***	-0.54*
Guess-Sharing (pre-discussion)	(0.09)	(0.12)	(0.12)	(0.23)	(0.27)	(0.40)
$\beta_{3,3}$: Second Info X	-0.23***	-0.43***	0.00	-0.56***	-0.88***	-0.16
Guess-Sharing (post-discussion)	(0.09)	(0.12)	(0.12)	(0.22)	(0.28)	(0.40)
α: Constant/	10.29***	10.14***	10.39***	61.37***	55.70***	64.34***
Structural Noise Parameter	(0.12)	(0.15)	(0.15)	(4.51)	(5.00)	(5.27)
Observations	4000	2000	2000	4000	2000	2000
p-value: $\beta_{3,1}$ equal across genders		0.02			0.14	
p-value: $\beta_{3,1}$ equal across genders		0.04			0.21	
p-value: $\beta_{3,1}$ equal across genders		0.01			0.19	
Includes Info X Order FEs	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows reduced-form and structural estimates for the couple sample's *Discussion* and *Guess-Sharing* rounds.

Reduced-form coefficients: Columns 1 to 3 show reduced-form results, estimating equation 1 (though replacing the *Draw-Sharing* round with the *Guess-Sharing* round) by OLS. The dependent variable is participants' private guess. "First Info" indicates the net number of red draws (i.e., red draws minus white draws) in the first set of signals. Similarly, "Second Info" indicates the net number of red draws in the second set of draws. "Only Accessible via Spouse" is an indicator that equals one for the final private guess in the *Discussion* round, when the second set of draws were drawn by the participant's spouse and then (potentially) communicated to her through discussion. "Guess-Sharing (pre-discussion)" indicates the second private guess in the *Guess-Sharing* round, after the participant was directly told their spouse's previous guess (as well as the number of draws that guess was based on) but before discussing it with their spouse. "Guess-Sharing (post-discussion)" indicates the third private guess in the *Guess-Sharing* round, after the discussion. All regressions include order fixed effects interacted with "First Info" and "Second Info." Standard errors are clustered at the couple level. *, **, and *** indicate significance at the $p<0.10$, 0.05, and 0.01 levels.

Structural parameters: Columns 4 to 6 show estimates of the structural model described in Section 4.3. "First Info" and "Second Info" indicate the weights placed on the first and second set of signals in the agents' quasi-Bayesian updating rule. The interaction terms indicate the difference in weight on the second signal in *Discussion* and *Guess-Sharing* rounds (see notes for reduced-form estimates for more detail). Bootstrapped standard errors (clustered at the couple level) in parentheses. *, **, and *** indicate significance at the $p<0.10$, 0.05, and 0.01 levels.

A.2 Third Experiment

After the completion of the experimental rounds with couples and strangers, we conducted a third experimental round with an additional sample of participants who had not previously participated in our study. The goal of this experimental round was to test the limits of our findings by considering robustness to three different features: (i) the salience of participants' partners' draws, (ii) the possibility for anchoring on their own draws first, and (iii) the size of the incentives for good guessing.

A.2.1 Treatments

This experiment consisted of six rounds, played in a completely randomized order, corresponding to six different treatment conditions. Two of these were a *Discussion* round and a *Draw-sharing* round, exactly as in our first experiment, included to provide a baseline and comparison with our previous sample. The other four rounds were variations of the *Draw-sharing* round and are described below.

***Draw-by-Draw* round**. This round aims to increase the salience of partner's draws. It proceeds identically to the *Draw-sharing* round, except that the experimenter shares their partner's draws with each participant ball-by-ball. For instance, the experimenter would say 'Your partner first drew a red ball, then a white ball, ...' and so on.

***In-person* round**. In this round, we increase the salience of partner's draws further. In particular, while each participant is making their set of draws, their partner remains in the booth and can observe their draws. (As a result, there is no need for the experimenter to share draws afterwards). Both participants guess, in private, after observing each set of draws; there is no discussion or joint guess. This design also controls for order effects by construction, since the participant who draws second learns their partner's draws before their own.

***No-First-Guess* round**. In this round, to seek to mitigate any potential anchoring on own information, participants do not make a guess directly after making their own set of draws. Otherwise, the round proceeds identically to the *Draw-sharing* round.

***Reverse-Order* round**. This round has one participant learn their partner's draw first, with the aim of ensuring that any tendency to anchor on the first piece of information would work against under-weighting their partner's information. To begin with, one person makes a set of draws from the urn. The other participant is then informed of their partner's draws and makes a guess, before making their own set of draws and another guess.[34] We only include this second participant, who learned their partner's draws first, in the analysis. Finally, the pair hold a discussion, and make post-discussion guesses as usual.

***Higher-Stakes* treatment.** We increased the incentives for accurate guessing in a randomly-chosen 3 out of 6 rounds. The maximum amount couples could earn and their loss in earnings per ball away from the truth were both increased by 50%, to Rs.15 and Rs. 45 respectively.

A.2.2 Results

Our follow-up experiment was interrupted mid-data collection by the COVID-19 pandemic, and (at time of writing) has not been safe to resume since. We have data for 146 pairs of individuals—68 mixed-gender pairs, 67 same-gender male pairs, and 11 same-gender female pairs—out of a planned sample size of 400 pairs (200 mixed-gender, 100 male and 100 female). To increase statistical power, we pool results across both genders and all types of pairs.

Comparison with main strangers sample. Figure A.V shows the coefficients from estimating equation 1 on our follow-up sample. The upper panel first checks that we obtain similar results in the baseline *Discussion* and *Draw-Sharing* rounds as we do in our previous strangers sample. As in the main strangers sample, participants weight their teammate's information significantly less than their own in the *Discussion* round and the *Draw-Sharing* round pre-discussion. In contrast, the post-discussion weights on their teammate's information is lower but not statistically different than the weight on their own information. Relative to the strangers results shown in Panels C and D of Figure 4, participants put slightly less weight on their partner's information in the *Discussion* round, slightly more in the *Draw-Sharing* round pre-discussion, and a very similar weight in the *Draw-Sharing* round post-discussion.

As we did not include an *Individual* round, we cannot directly test for order effects within rounds (i.e. that people weight own information more because it came first) in the same way as our main experiment. However, the similar results below from our *In-*

[34] In the same manner as the original *Draw-sharing* round, e.g. 'four red, five white'.

Person and *Reverse Order* rounds, in which the order was reversed for some, suggest that as with the strangers sample, this overweighting of own information is unlikely to be due to order effects.

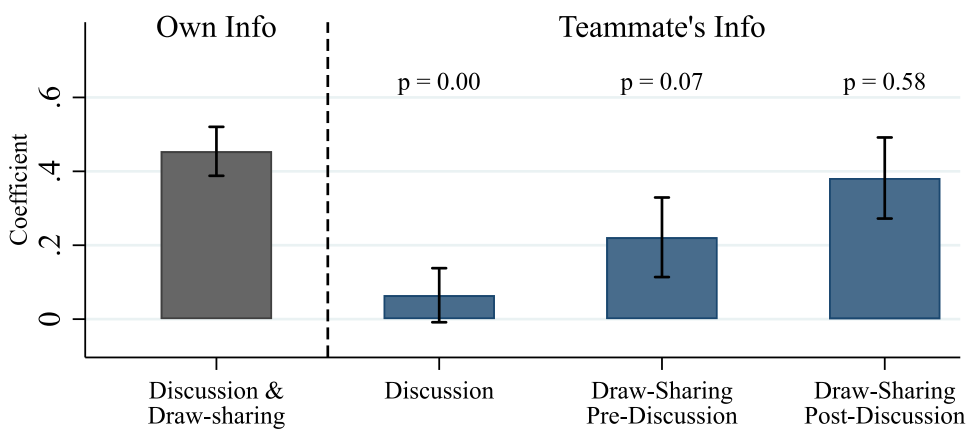
Effects of new treatments. The lower panel of Figure A.V shows the weights participants place on each set of signals in the *Draw-Sharing* round and the new treatments (rounds) introduced in our follow-up design. We focus on the pre-discussion guess in all rounds, as this is where the relative underweighting of partner's information emerges in the *Draw-Sharing* round. The new rounds produce only modest increases in the weight attached to partner's information, and in all cases it remains underweighted relative to own information. Compared to a *Draw-Sharing* round weight of 0.29, participants place weights on partner's information of 0.32 in the *Draw-by-Draw* round, 0.44 in the *In-Person* round, 0.37 in the *No-First-Guess* round and 0.44 in the *Reverse-Order* round. All of these are substantially less than the average weight put on own information (0.58), and the difference between own and partner's information is significant except for the *Reverse-Order* round (which was the most underpowered).

Overall, these results suggest that neither the relative salience of partner's draws, nor anchoring on one's own draws with a first guess, can explain our findings. The *Reverse-Order* round indicates that even creating the possibility of anchoring on one's partner's draws first does not overturn the underweighting of partner's information. Perhaps most striking of all, even in the *In-Person* round, in which participants directly watch each other make the draws, they still put 24% less weight (p=0.03) on their partner's information than their own.

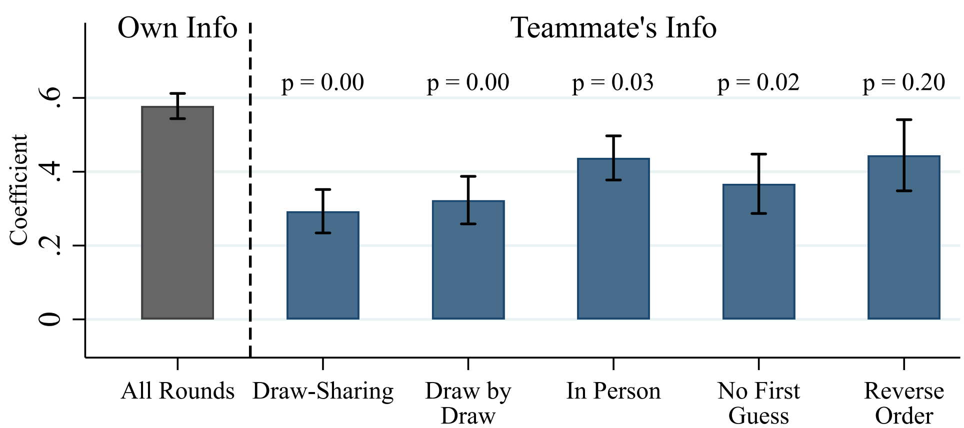
Effect of higher incentives. Figure A.VI shows participants' weight on signals separately by whether they faced the usual stakes for accurate guessing (Rs. 30 per ball away) or higher stakes (Rs. 45 per ball away). The figure shows that the stakes have little effect: when the stakes are high, participants do not appear to place a higher weight on their own or their partner's information in pre-discussion guesses.

Figure A.V: Weights on Partner's Information in the Follow-Up Experiment

Panel A: Discussion vs. Draw-Sharing



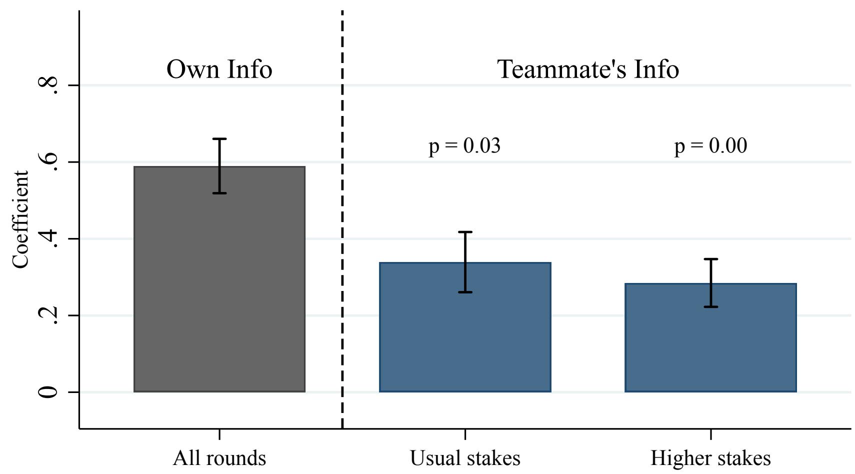
Panel B: Variations of Draw-Sharing (Pre-Discussion)



Notes: This figure shows the weights that men and women together in our follow-up sample put on their partner's information in different rounds, with the average weight on their own information across those rounds for comparison. We estimate coefficients from equation (1). The lighter bars show the weights on the guesser's own information and the darker bars show the weight on their partner's information. Bands show $\pm$ one standard error.

- **Panel A** compares the *Discussion* round to the pre- and post-discussion guesses in the *Draw-Sharing* round. These rounds had the same structure as the corresponding rounds in our main experiments.
- **Panel B** compares variations on the *Draw-Sharing* round, always looking at the pre-discussion guess (but after participants had learned their partner's information). The variations are: (a) *Sequential*, identical to the draw-sharing round except that the experimenter shares partner's information ball-by-ball; (b) *In Person*, in which each participant directly observes their partner make their draws, as well as making their own; (c) *No First Guess*, identical to the *Draw-sharing* round except that participants do not make a guess immediately after drawing their own set of balls; (d) *Reverse Order*, in which one participant learns their partner's draws before making their own.

Figure A.VI: Weights by Size of Incentives in the Follow-Up Experiment



Notes: This figure shows the weights that men and women together in our follow-up sample put on different pieces of information.

- We estimate coefficients from equation (1). We pool data across all rounds which had (a) the same incentives as our main experiment ("Usual Stakes") and (b) incentives increased by 50 % ("Higher Stakes").
- Each couple did three randomly-selected rounds with Usual Stakes and three with Higher Stakes.
- The regressions include controls for the round treatment condition (*Discussion*, *Draw-Sharing* etc.) as the randomization of stakes was not perfectly balanced by this variable. Bands show $\pm$ one standard error.

A.3 Comparing Structural to Non-Structural Results

In this section we consider whether the estimates of the structural model outlined in Section 4.3 are consistent with the main non-structural results presented elsewhere in Section 5. To do so, we simulate, using the estimates of the parameters of the model, what guesses participants would make given the signals they had. To eliminate unnecessary noise, instead of simulating just once for each guess of what the participant would choose (which is noisy), we calculate the expected guess. We then create versions of columns 4-6 of Tables 2 and Figure 3 using the simulated data. The question these analyses allow us to answer is, "Are the estimated biases from the model sufficient to explain the patterns found in the reduced-form and non-structural results?" If the model implied that the non-structural analyses would look very different than in fact they do, this would suggest that the model is not capturing something important about the biases we document.

Columns 1 to 3 of Table A.XII replicate the final three columns of Table 2, our main reduced-form results. Columns 4 to 6 show the same regressions but using the model-implied expected guesses as the dependent variable rather than participants' true guesses. Note that these variables, because they are expectations rather than single draws from the distribution of guesses, are mechanically much less noisy than the actual guesses. However, as Table A.XII shows, the *size* of the coefficients are extremely similar (i.e., comparing cols 1 to 4, 2 to 5, and 3 to 6). Our interpretation of these results is that the model estimates are sufficient to explain the pattern of results shown in the reduced-form analyses.

Figure A.VII compares the non-parametric results from Figure 3 with similar estimates using the model-simulated data. For clarity, there are four panels, representing the individual round and discussion round separately for husbands and wives. Each panel shows the estimates given the actual guesses that participants make (in gray) along with the model-simulated expected guesses (in blue). As expected, actual guesses are noisier, but slope of the curves are extremely similar within each panel, suggesting that the non-parametric and structural effect sizes are of comparable magnitude.

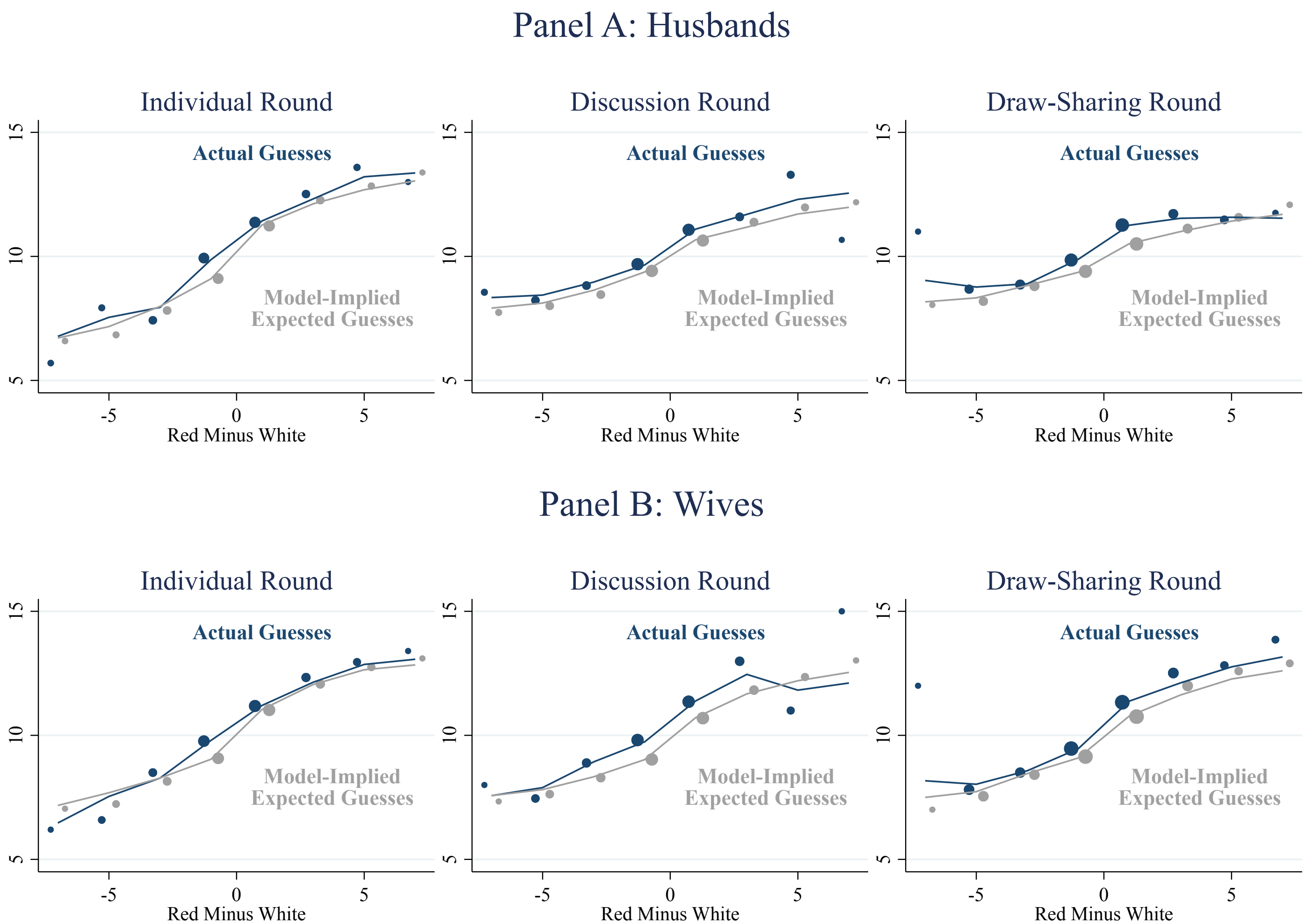
Table A.XII: Comparing Reduced-Form to Structural Results

	Actual Guesses			**Model-Implied Expected Guesses**		
β_1: First Info	0.43***	0.50***	0.37***	0.40***	0.46***	0.35***
	(0.03)	(0.05)	(0.05)	(0.01)	(0.01)	(0.01)
β_2: Second Info	0.53***	0.54***	0.54***	0.52***	0.55***	0.52***
	(0.05)	(0.08)	(0.07)	(0.01)	(0.01)	(0.02)
$\beta_{3,1}$: Second Info X	-0.20***	-0.32***	-0.04	-0.18***	-0.30***	-0.06***
Discussion	(0.06)	(0.08)	(0.08)	(0.01)	(0.01)	(0.01)
$\beta_{3,2}$: Second Info X	-0.28***	-0.53***	-0.01	-0.25***	-0.44***	-0.05***
Draw-Sharing (pre-discussion)	(0.09)	(0.12)	(0.12)	(0.02)	(0.02)	(0.02)
$\beta_{3,3}$: Second Info X	-0.13	-0.35***	0.11	-0.10***	-0.25***	0.06***
Draw-Sharing (post-discussion)	(0.09)	(0.12)	(0.12)	(0.02)	(0.02)	(0.02)
α: Constant	10.29***	10.14***	10.39***	10.03***	10.01***	10.01***
	(0.12)	(0.15)	(0.15)	(0.02)	(0.02)	(0.02)
Observations	4000	2000	2000	4000	2000	2000
p-value: $\beta_{3,1}$ equal across genders		0.02			0.00	
p-value: $\beta_{3,2}$ equal across genders		0.00			0.00	
p-value: $\beta_{3,3}$ equal across genders		0.00			0.00	
Includes Info X Order FEs	Yes	Yes	Yes	Yes	Yes	Yes

Actual Guesses: Columns 1 to 3 show reduced-form results, estimating equation 1 by OLS. They are identical to those shown in columns 1 to 3 of Table 2. The dependent variable is participants' private guess. "First Info" indicates the net number of red draws (i.e., red draws minus white draws) in the first set of signals. Similarly, "Second Info" indicates the net number of red draws in the second set of draws. "Discussion" is an indicator that equals one for the final private guess in the *Discussion* round, when the second set of draws were drawn by the participant's spouse and then (potentially) communicated to her through discussion. "Draw-Sharing (pre-discussion)" indicates the second private guess in the *Draw-Sharing* round, after the participant was directly told their spouse's information but before discussing it with their spouse. "Draw-Sharing (post-discussion)" indicates the third private guess in the *Draw-Sharing* round, after the discussion. All regressions include order fixed effects interacted with "First Info" and "Second Info." Standard errors are clustered at the couple level. *, **, and *** indicate significance at the $p < 0.10$, 0.05, and 0.01 levels.

Model-Implied Expected Guesses: show the same regressions as in columns 1 to 3, but use the expected guesses (conditional on actual signals) implied by the structural estimates presented in columns 4 to 6 of Table 2.

Figure A.VII: Comparing Non-Parametric to Structural Results



Notes: This figure shows the average second private guess of husbands (Panel A) and wives (Panel B). The x-axis denotes the net number of red draws (i.e. red draws minus white draws) in the second signal of the round. Blue dots indicate average actual guesses, while gray dots indicate average expected guesses from the structural model described in Section 4.3 (using the estimated parameters in columns 5 and 6 of Table 2). Curves denote locally-weighted averages (lowess).