Married by the Rev. Bayes? An experiment testing joint reasoning within married couples.*

Alistair Munro[1,]

National Graduate Institute for Policy Studies (GRIPS), Roppongi 7-22-1, Minato-ku, Tokyo, Japan, 106-8677.

Abstract

Although there are many views on how married couples process information, no attempt has been made to examine this process experimentally or to compare it to how individuals and pairs of unrelated individuals behave in similar circumstances. In this lab-in-field experiment, held in eastern Uganda, participants get rewards for correctly choosing the identity of an unknown bag after learning some information about its contents. In some sections, some individuals work alone while other pairs share the information and make a joint decision. In some sections, individual partners see different information sets and the only communication allowed between partners is a recommendation. I find that faced with the basic task, married pairs perform consistently better than individuals. However, married couples are significantly outperformed by the teams consisting of unrelated partners. When players are separated from their partners, and can only communicate in a restricted way, performance falls below that of individuals. Married couples are not better at these tasks than unmarried pairs, but they are no worse. Wives are more likely to defer to their husbands than vice versa and are less sensitive to context in their deference. More generally, subjects place less weight than they should on their partner's information and more weight on their own observations. Taken at face value the results suggest that there is no obvious information-processing gain from making decisions with spouses.

*This research was supported by JSPS Kakenhi grant 17H02498. Thanks are due to Arjan Verschoor for his reliable advice, to the members of the Field Lab in Mbale for their data gathering and especially to Joshua Balungira for his intelligent and careful leadership of the fieldwork. THIS WORK IS PRELIMINARY. COMMENTS ARE WELCOME, BUT PLEASE DO NOT CITE.

Keywords: Bayes Theorem, intra-household, Uganda, group decision, information,

1. Introduction.

Understanding how decision-makers process new information is crucial to the construction of reliable models of economic behaviour. For individual decision-makers, both theory and evidence are plentiful. In contrast, for decisions made by groups the evidence - and the theory - is thinner on the ground, despite the fact that decision-making groups of one kind or another are at the heart of all human societies. The widespread acknowledgement of this contrast has led to an increase in economic research into group decisions (e.g. Blinder and Morgan (2005), Charness and Sutter (2012)), though most of the focus has been on groups assembled for the purposes of the experiment and not pre-existing decision units. Meanwhile globally, the majority of adults live with one or more other adults, making decisions jointly and separately that affect the household as a whole. Despite this fact, the information processing household is neglected.

As a response to the lacuna, this paper presents a lab-in-the-field experiment that tests basic hypotheses about how households process information. To be specific, the research centres on three things: first, whether married couples make better (in the sense of Bayes) judgements of an uncertain situation, compared to individuals. Second, in the same context whether couples are better than two-person teams consisting of people who are not married to one another. Third, I test the degree to which partners take advice from each other in situations which vary the information held by each person. I do this both for married couples and for pairs where the partners are from different households. The participants for the experiment are married, heterosexual couples drawn from the eastern region of Uganda, on the slopes of Mt. Elgon. Though the sample region is particular, the couples living there face many of the same challenges and experiences as households throughout the world.

To conduct these tests, I use a very simple urn-type game[1] in which players must make a judgement whether an unknown bag of balls is of type

[1]Such games are commonly used in tests of individual decision-making and experiments on social networks. E.g. Goeree and Yariv (2015), Anderson and Holt (2008)

A or of type B, based on draws without replacement of some of the balls from the bag. Players are financially rewarded for choosing correctly and monetary incentives are aligned between the partners. For some of the draws it is logically possible to deduce the bag; in other cases it is only possible to decide which bag is the more likely. In some of the parts of the experiment where players are separated from their partners, but can pass on advice, one person in each pair has more knowledge about the contents of the bag, and this fact is known to both players. By manipulating information sets, I test the degree to which individuals follow more and less informed opinions.

To preview the results, I find that faced with the basic judgement task, married pairs perform consistently better than individuals. However, married couples are significantly outperformed by the teams consisting of unrelated partners. When players are separated from their partners, and can only communicate in a restricted way, performance falls below that of individuals. Married couples are not better at these task than unmarried pairs, but they are no worse. Generally, subjects place less weight than they should on their partner's information and more weight on their own information. Wives are more likely to defer to their husbands than vice versa. While some of these results are modified by age, education or mathematical understanding, factors such as trust in others, a belief in the value of teamwork or attitudes to deference play little role in explaining patterns in the data.

1.1. Motives for researching this issue.

As already mentioned, the major motives for this experiment come from the absence of experimental evidence within economics on how partners within a household make decisions and in particular, how information is aggregated. There is of course, an enormous list of works of fiction and pop psychology[2] about how couples do or do not communicate. Outside economics, there is also an extensive academic literature on the subject, (Sillars and Scott (1983), Navran (1967)), which provides many insights, largely on the basis of interviews and recordings of decisions (e.g. Spiro (1983)).

Historically, economists have generally skirted around the process by which households make decisions, preferring instead to focus on models that

[2]For example, https://www.dailymail.co.uk/news/article-2873728/Proof-men-really-selective-hearing-Study-finds-average-bloke-switches-just-six-minutes-chatting-half-pays-attention-15-minutes-talking-sport-mates.html is based on a 2014 survey by a betting company in the U.K.

use observable data on allocations (e.g. Chiappori and Ekeland (2009)). Some notable exceptions include the theoretical work of Katz (1997) as well as the empirical investigation of Bernard et al. (2018) which uses a survey instrument and manipulation of the context of hypothetical decisions to examine patterns of influence. Meanwhile, Ben-Porath (1980) proposes that there is an efficiency premium for the family in its enhanced ability to trust information shared within its confines.

Insights into intra-household decisions have also been provided by the steady growth in experimental work on households (e.g. Peters et al. (2004), Bateman and Munro (2005); Ashraf (2009)) over the last fifteen years or so (surveyed recently in Munro (2018)), which has revealed patterns of allocation and investment that are starkly at odds with standard models of the households based on assumptions of Pareto efficiency. Playing simple public-good games, couples around the world have been reluctant to invest in the common pool, Verschoor et al. (2019), routinely hide payoffs from partners (Castilla and Walker (2013); Hoel (2015)) and show patterns of behaviour in trust and dictator games Kebede et al. (2014) that clearly reject models of household efficiency or individual altruism. While various designs have been used to examine household decisions, the focus has been on the efficiency of the product of the decision-making rather than the efficiency of the process. As such, the designs used typically have a trade-off between individual and household payoffs. Here, I take a step in a different direction, focussing on how decisions are made in situations where, at least in cash terms, there is no conflict between private and household returns.

Some other motives for this experiment are also provided by the recent discussion about group decision-making amongst economists working in disparate fields. Experimental subjects often struggle with Bayesian inference, when the optimal decision conflicts with simpler intuition Grether (1980). Generally though, (e.g Blinder and Morgan (2005), Kocher and Sutter (2004) or Charness et al. (2007)), the thrust of the evidence is that groups make better decisions than individuals in the sense of coming closer to the predictions of rational choice theory, including Bayes' Theorem, particularly when there are opportunities for learning (Charness and Sutter (2012)). There are also suggestions, from non-economics experiments (e.g. De Dreu and Weingart (2003); Reinig et al. (2015)) that behaviour within established groups is often affected by feelings of relationship maintenance in addition to the straightforward desire to maximize group performance. Emotions, long-run relationship concerns and so on, might well interfere with making logically correct

decisions if, for example, one partner loses face.

Some further light on this possibility comes from the growing body of evidence from experiments and elsewhere on how individuals respond to others in their social network when making their own decisions Anderson and Holt (2008), Gale and Kariv (2003). Much of the evidence suggests that individuals are not simple Bayesians and may be significantly affected by their social connection with the person who supplies information Berger et al. (2018). For example, in information cascade experiments, subjects are exposed to their own draw from an urn of unknown composition as well as partial information about the judgements made by previous participants. Typically, Çelen and Kariv (2004) individuals place weight on the decisions made by their peers, with the effect larger when members of the chain share a closer identity (Berger et al. (2018)). However, Huck and Oechssler (2000) and Kraemer et al. (2006) provide evidence that individuals often overweight their own information, particularly in situations where information is contradictory.

Teams typically share a common payoff, but they also provide a situation in which members can exchange advice and information. Studies by Birnbaum and Mellers (1983); Fishchler and Sorkin (2006) amongst others examine situations where advice of varying quality is received before a prediction is made by one party acting alone. In general, more weight is placed on the advice when it is labelled as coming from a source with greater expertise. See also Yaniv (2004) who asked subjects to answer general knowledge questions and provide uncertainty bounds before and after they received advice about the correct answer. Liberman et al. (2012) investigate dyads: student subjects are asked to make predictions about the percentage of the population sharing particular normative views (e.g. acceptability of gay marriage) before and after receiving the predictions from a partner. People are influenced by partners, but put more weight on their own initial guess. Mercier et al. (2012) argue that weighting of advice from others is similar across cultures. Meanwhile, Tost et al. (2012) manipulate feelings of power through a vivid writing priming task and then examine receptivity to receiving advice in a subsequent task that involved estimating the weight of people based on photographs. They find that higher power individuals were less willing to take advice (put less weight on it) compared to lower power individuals. In an urn guessing game, Sharvit and Sproten (2012) find that men are more likely than women to turn to advice. This result, which they view as being against general gender stereotypes as well as prior questionnaire-based evid-

Table 1: Optimal Decisions for Pooled Knowledge.

Draw	Prob $(A \mid Draw)$	Optimal guess
Two yellow, two blue	1	A
Three yellow, one blue	$\frac{2/5}{2/5+4/5}$=1/3	B
Four yellow	0	B

ence, is put down the incentivized nature of the task and the possibility that the willingness to seek advice may be domain-sensitive.

While these experiments from various, scattered fields of research provide a fascinating background, it really is not apparent that the lessons learnt from them can be automatically applied to intra-household decisions, where the individuals concerned are typically tied to each other by shared payoffs and a lengthy relationship histories. And while, other-regarding preferences seem to be a feature of many economics domains, the strength of feelings between family members is probably stronger and more complex than almost any other realm of human interaction. Thus an experiment specifically on households is required.

2. Basic Design

The experiment is in six parts, but all parts have a common component.[3] There are two possible bags. Each is known to contain 5 balls, some of which are yellow and some of which are blue. Specifically, it is equally likely that the bag contains 3 yellow and two blue balls (bag A) or that it contains 4 yellow balls and a single blue ball (bag B). Some balls are drawn from the bag without replacement and the player or team has to guess whether the bag is A or B. A correct guess is rewarded with money. When playing as part of a team, both team members get the reward. In the basic version of the task, which is used for the first two parts of the experiment, subjects see four balls. In other words only one ball is hidden.

When the decison maker knows the colour of all four balls, then the optimal decision (from a Bayesian perspective) is as shown in Table 1.

The calculation for the middle case arises from considering the probability of the colour of the remaining ball. If the bag is A, then there is a 2/5 chance

[3]See AsPredicted.org for the pre-registration documentation.

that a single ball is blue, while if the bag is B there is a 4/5 chance that a single ball is yellow.

This task was chosen after some prior tests with individuals. Although it is trivial for someone with training in probability theory, for many individuals the optimal strategy is not obvious at first, even for the cases where the nature of the bag can be deduced with certainty. For the non-probability theorist, the case where only one blue ball is seen is potentially challenging: even a participant who consistently chooses with B bag will be wrong 33 % of the time. I therefore label the one blue ball case 'difficult'.

One of the key decisions was whether to use a relatively abstract task such as the one eventually picked, or to frame the problem in more concrete terms. For instance, the 'bags' could be bags of seed and the blue balls could be rotten seed. Initial discussions with fieldworkers suggested that the framed task was no easier to understand than an unframed task. The other reason an abstract task was picked was to avoid obvious gender associations. Concrete problems might be linked to decision more familiar to one gender or the other and this might be an issue, especially in parts 3 to 6 (see below). Of course, it is still important to understand whether households solve framed or more realistic decision problems differently, but for a first experiment on the topic, it was judged more straightforward to begin with an abstract task that was similar to problems used in laboratory experiments with individuals and groups. [4]

In parts 1 and 2 of the experiment, subjects play this task repeatedly, either on their own or as part of a team of two people. The team is the same throughout the experiment and in this section, when making joint choices the partners sit side by side and can discuss their joint choice.

In parts 3 to 6 of the experiment, the partners of the team are separated. One team member sees four balls as before, while one sees a subset of the four balls. In each part, one player plays of the role of the recommender: he or she can only communicate 'A' or 'B' as advice to the other player. It is not possible to communicate the mix of balls seen. The other partner plays the role of 'chooser' and must choose on behalf of the team after receiving the advice and observing two or four balls, depending on the part.

[4] Blinder and Morgan (2005) for example uses an urn task and voting in groups of three. Goeree and Yariv (2011) in an experiment with groups nine use two bottles containing seven balls of two different colours.

Table 2 summarizes the design. In this table, 'individual' means playing the task alone; 'couple' refers to a married couple and 'pair' refers to two people who are not married to each other (but by design will be married to someone else in the experiment.) Pairs may be of the same gender or mixed and players are matched to the same partner throughout the experiment.[5] The difference between treatments 1 and 2 is solely in the order of the joint and individual choice. Blinder and Morgan (2005) is one example of an experiment that finds significant order effects, so I specifically control for the possibility.

We randomly label individuals as 'Red' or 'Green' within the team, with assignment done at the entrance to the venue. Then, 'Red4' means that the Red player sees four balls and acts as the recommender while Green is the chooser and sees the subset of two. Similarly 'Green2' means that green is the recommender and sees only two balls, while Red chooses and sees four. In each part, players play six tasks.

Table 2: *Design*

Part	Treatment 1	Treatment 2	Treatment 3
1	Individual	Couple	Pair
2	Couple	Individual	Individual
3	Red4	Red4	Red4
4	Red2	Red2	Red2
5	Green4	Green4	Green4
6	Green2	Green2	Green2

When the subjects are separated, one person is shown four of the balls, while the other is shown only two (from the same set). The fact that one set is a subset of the other is common knowledge. Define a Bayesian universe as one in which both partners use Bayes' theorem to make probabilistic judgements and there is common knowledge of this within the partnership. Given this, in a Bayesian universe the person who holds only two balls should defer to the

[5]There are no strict prohibitions about separation of men and women, but obviously there are some potential sensitivities about married people working in mixed pairs with someone who is not their spouse. After consulting with the local team about the possibility of marital violence, we went ahead after we found a label for 'pair' in the local language that had no sexual connotations and was unlikely to provoke spouses.

Table 3: Optimal Inference for Two Ball Case.

Draw	Prob $(A \mid Draw)$	Optimal message
Two blue	1	a
One yellow, one blue	$\frac{6/10}{6/10+4/10} = 6/10$	a
Two yellow	$\frac{3/10}{3/10+6/10} = 1/3$	b

judgement of the other person. Conversely, the person who sees four balls should ignore the advice of the partner.

Of course, in reality, players may depart from this ideal notion of rationality. They may for example suspect that the other person is not Bayesian. They themselves may also update probabilities in a way that is not compatible with Bayesian theorem. Thus, the player who holds four balls may not complete discount the opinions of the person with two balls, while the person who sees only two balls may not fully credit the other player will all the expertise.

Let us first consider the sub-case where the recommender or chooser knows the colour of four balls. The optimal message then follows the pattern of Table 1.

When the recommender sees only two balls, then the message should be irrelevant, but the optimal inference is as in Table 3. Lower case is used to indicate advice.

From this we see that when all four balls are yellow or when both players see two blue balls, then the inference made is aligned in the sense that separately both come to the same judgement. In other cases though, separate judgement may lead to contradictory inference. In particular, when the two ball player sees only one blue ball but the four ball person sees two blue balls, then separately they will come to opposite conclusions. While this is not relevant for the Bayesian universe, it may be important when interpreting the actions of players who depart from full rationality (or when there is an absence of common knowledge). Consider for example, a person who believes that their partner is liable to make mistakes. In particular, they believe their partner chooses either A or B to recommend randomly with probability q and chooses according to Bayes' theorem with probability 1-q. A player with four balls will still wish to ignore the recommender's advice. A player with two balls may now also wish to ignore the advice if q is large enough. For the player who sees two balls, the specific probabilities are given

Table 4: Optimal Decisions for Pooled Knowledge.

Draw	Advice	Prob $(A \mid Draw)$	Prob $(A \mid Draw, Advice)$	Critical q
Two blue	a	1	1	-
Two blue	b	1	1	-
One yellow, one blue	a	6/10	$\frac{3(1-q)+4q}{5(1-q)+4q}$	-
One yellow, one blue	b	6/10	$\frac{2q+3(1-q)}{5(1-q)+6q}$	1/3
Two yellow	a	1/3	$\frac{2(1-q)+q}{4(1-q)+q}$	4/5
Two yellow	b	1/3	$\frac{3(1-q)+4q}{9(1-q)+16q}$	-

in Table 4:

The final column of this Table shows the critical value of q below which it is optimal to ignore the advice given by the recommender. A '-' indicates that the value of q, makes no difference to the optimal inference. When two blue balls are seen, for example, it is always optimal to choose A. On the other hand, if the chooser makes decisions purely on the basis of the two balls in their draw, then he or she make choices that contradict the advice given, particularly when one ball is blue and the other is yellow. The de Groot model (DeGroot (1974)) is frequently used as an alternative positive model of choice. In this approach, individuals base their guess on a weighted average of the alternatives they hear (including their own opinion). If the weights placed on the views of partners are sufficiently large, then the choice made by a person can differ from their inital, private opinion of the optimum. Here, a de Groot model could allow a person who sees two blue balls to choose B if that is the opinion of their partner. A de Groot model could also allow a person who sees four balls, to defer to the views of their partner. Thus, a person seeing four yellow balls or two yellow and two blue, could choose A (respectively, B) if that is the advice received from the spouse. In this way, the theoretically-consistent set of choices differs between the de Groot model and the Bayes and imperfect Bayes models.

2.1. Hypothesis tests.

As the introduction suggests, there is no clear body of evidence or theory that provides a firm expectations about the behaviour of couples versus

Table 5: Choices consistent with different theories.

Draw	Advice	Bayes Partner	Non-Bayes Partner	de Groot
BB	a	A	A	A
BB	b	A	A	A,B
YB	a	A	A	A
YB	b	B	A,B	A,B
YY	a	A	A,B	A
YY	b	B	A,B	B
BBYY	a	A	A	A
BBYY	b	A	A	A,B
YYYB	a	B	B	A,B
YYYB	b	B	B	B
YYYY	a	B	B	A,B
YYYY	b	B	B	B
stuff				

individuals or unmarried pairs in a decision-making experiment. For that reason, I state the hypotheses in quite general terms.

Using this design I test the following hypotheses for the first two parts:

1. Pairs make better judgements than individuals. Specifically, the judgements of pairs are closer to the predictions of Bayes.
2. Married partners working together make better judgements than non-married pairs.
3. Performance improves.

For parts 3-6, I test:

1. Married partners make better judgement than non-married pairs.
2. Players overweight their own information, compared to their partners'
3. Players put more weight on their own information when they see four balls, compared to when they see two.
4. Women underweight their own information when playing with men. Specifically, when women see four balls and choose they are influenced by the recommendation of their partners.

5. Men overweight their own information. Specifically, when men see two balls and choose they ignore the advice of their partners.
6. Men overweight their own information more when playing with women who are not their spouses.

3. Location and implementation.

A pilot was carried out at the end of February 2019 with implementation of the main experiment in the following month in Simu sub-county in the south-east of Uganda. This district is on the densely populated and fertile slopes of the 4,300m Mabala (Mt. Elgon), an area with a mean population density of 284 people per km^2, an average farm size of 1.4-1.5 ha and rainfall of about 1186mm per annum (Iversen et al. (2011)). Livelihoods are mostly agricultural, but still complex: in amongst the crop growing there is livestock rearing, petty trading and services, pursued individually and jointly within the household. Soils are fertile and crops widely grown include maize, beans and groundnuts. Infrastructure is patchy: most villages away from the main roads lack electricity while sealed roads are uncommon and often poorly maintained.[6] The location was selected in part because of prior experiments on household behaviour (e.g. Iversen et al. (2011) in the area and also the good evidence of complexity and diversity in the relationships between spouses (Verschoor (2008)), Jackson (2013)).

Prior to the experiment, 16 villages were selected and a census of married couples between the ages of 20 and 70 took place. Out of the set of eligible households, 20 married couples[7] were selected at random and invited to take part in the experiment. For all these couples, a household survey was conducted at their homes. At the same time, for each village, 2-3 married couples were invited to be 'reserves' in case the original couples did not turn up.[8] On the day of each experiment, paid mobilisers were used to encourage the invited couples to turn up. Each day two sessions of a maximum

[6]Some further information about the region can be found in Iversen et al. (2011).

[7]To be clear, this means that even for the treatment where pairs were not spouses, all the participants in each session were married to someone else in the session.

[8]In general, the full complement of invited couples turned up, except for a few villages where weather affected the turnout. In a handful of cases, pairs were rejected after the initial survey, when it became clear that they were pretending to be a married couple in order to receive payment.

of twenty people were conducted, using classrooms in the village school, or community buildings. To avoid contamination, one treatment was run in the morning and a different treatment in the afternoon (according to a pre-arranged plan), and couples waiting for the second session were kept apart from couples exiting the first session.

At the entrance to the classroom, pairs were assigned numbers and individuals within each pair were handed red or green cards at random. Desks had been pre-assigned to specific pairs and the experimenters were careful to make sure that participants sat where they were told to. In the cases where the experiment started with joint decision-making, pairs were sat as a team. Where individual decision-making formed the first part of the experiment, subjects were sat individually with partners clearly separated from one another. For treatment 3, where individuals were paired with non-spouses, we also took care to ensure that spouses were not sat close to one another. After a brief introduction and check on identity, the five experiments began the instructions in the local language, using visual aids and posters on the blackboard to help the explanations.[9] Bags containing blue and yellow balls were used to explain the basic problem, but for the actual questions, each pair or individual had a sheet in front of them showing the information that was being revealed and with a place on the sheet for them to mark their guess (see Fig. 1).[10] For the joint questions, subjects were encouraged to discuss their answers quietly with their partners, but otherwise participants were asked to refrain from speaking to one another. Before the initial question, each player faced four tests of understanding[11] and once we were satisfied

[9]See appendix for the English language training script for parts 1, 2, 3 and 6 for treatment 1. The other parts and scripts for the other treatments are very similar and are therefore omitted.

[10]The answers were predetermined. For assurance sake, we kept a set of envelopes containing the answer to each question, with the question number clearly marked on the outside. If any subject challenged the research team, the envelope could be opened and the answer shown. The solutions for the questions were generated in two stages. In each groups of six questions, I deliberately created three tasks where the solution was deducible. For the other three tasks, where only one ball out of the four displayed was blue, a random number generator was used to select the correct bag. In this way, the probability that the true bag was A or B, conditional on the information displayed, matched the underlying theory.

[11]Some further test questions were asked later in the experiment when we switched formats.

that everyone in the room understood the procedures, the incentivized questions began. For each correct answer we paid USh500 per person. For the individual questions this means 500 shillings; for questions where the team provides a single prediction, each player receives 500 shillings for a correct answer. It implies that a person who got all 36 questions correct would earn several days wages for 2-3 hours of work. Subjects are told this information at the start of the experiment. They are also informed that they will be paid separately and individually. Of course, in only one part of the experiment will earnings potentially vary across the individuals within the team.

In parts 1 and 2 after each question has been completed, the enumerating team circulate through the room and mark the answer sheets as correct or not. Then the next question is done until all six questions are complete. Prior to part 2, the teams are reseated: for treatment 1 this means the married couples are brought together; for the other treatments the partners are separated and required to work individually for the next part.

For part 3 onward the subjects are separated from their partners. Green subjects are in one room, with red subjects in the other room. When the subjects in a room are the choosers, they are asked to wait while the other people in the other classroom make their recommendations. They are told what the other people are doing and it is made clear that they are being paid as a team and who has what information. Subjects in the recommenders room are also told about the choosers task. After this instruction they are asked to make the recommendations for the six questions they face. Once all recommendations have been done, the experimenters take the recommendation sheets to the other room and present them to the partners who then make the six choices. Once a part has ended we introduce the next part until the experiment is over at which point total performance is tallied and individuals are paid. While they are waiting the subjects complete a very short questionnaire about their view of the experiment and their belief in their own abilities compared to their partner.

The basic idea of the experiment was understood very well by subjects. In the initial check questions,[12] 93% got all four correct at the first attempt. Subjects who did not get the check questions correct were taken individually through the instructions again until they understood. Most of the effort in

[12]How many yellow balls are in bag A; how many yellow balls are in bag B, if you choose the correct bag, how do you win, if you choose the wrong bag, how much do you win?

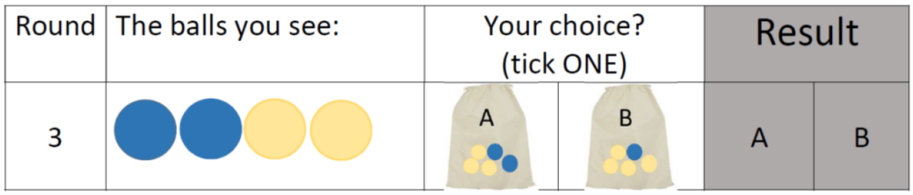


Figure 1: Example Question

revising in and after the pilot was in smoothing the flow of information and paper so as to shorten the duration of the experiment which typically took around two hours to complete.

4. Results.

In Table 6 I present descriptive statistics for the sample. In all, 564 individuals took part in the experiment. The numbers are not equal across all treatments due to the pattern of no-shows varying by village, so we have 172 participants for the third treatment (non-spouses), 180 for the second treatment (joint decisions first) and 212 for the first treatment (individual decisions first). All households are engaged in farming with nearly all growing the staple maize as well the cash crop, coffee and several other crops besides.

We ask some basic arithmetic questions, such as what is 46-19, and men score more highly for this. Men also have slightly more experience of formal education. In addition to features of the household and its farming practices, we also ask various attitude questions, a few of which are summarized in Table 6. Both men and women are more likely than not to agree that wives should aways obey husbands, but men are more likely to agree with the statement. Men are less likely than women to state that they are the primary financial decision-maker, but both sexes agree that men are generally in charge of household finances. Reflecting this, men generally feel they have permission to spend more with consulting their spouses, compared to women. Levels of trust in strangers are generally low for both sexes. Most respondents report that their spouse is the person they would most likely consult with before buying a new mobile. For other decisions, the pattern is similar except for the decision about using a new seed, where men are more likely to consult other people in the same village with similar experiences. We ask some

Table 6: Descriptives

Variable	Female	Male
Age	35.7	42.8
Highest Education (mean level)	2.2	2.5
Land holdings (acres, mean)	2.2	2.2
Household size (mean)	5.9	5.9
Basic arithmetic test (out of 4, mean)	2.1	3.1
Household grows maize (%)	92.5	92.5
Household grows coffee (%)	84.5	84.5
Women should always obey (% completely agree)	63.8	73.1
Husband alone is primary financial decision-maker (%)	84.0	69.9
Whether people can be trusted (% you can't be too careful)	30.5	34.4
With whom would you discuss new mobile (% 'spouse')	69.9	75.9
Often change opinion to please others (% completely agree)	52.8	25.9
Prefer to work in team (% completely agree)	75.2	84.0
Largest sum you could spend alone (USh, median)	10000	5000

Note: for Education level, 0= No education, 1 = Some primary education, 2 = Some secondary, 3 = Tertiary; Land holdings, household size and crop questions are only asked at household level. All other questions asked individually.

questions about agreeability. Women are generally more likely to report that they would change an opinion to suit others. Both genders typically report preferring to work in teams, although in a different question, men (27.1%) are much more likely than women (14.5%) to report they completely agree they prefer to do a job alone.

4.1. Experimental Results.

I focus initially on the performance of individuals versus couples and pairs in Parts 1 and 2 of the experiment. In later subsections I consider Parts 3 to 6.

Table 7 summarises the success rate of the various categories of subjects in the first two parts of the experiment.

Result 1. *Pairs, whether married or not, generally make better judgements than individuals.*

The results in Table 7 consistently support this claim, in both Part 1 and Part 2. Generally, individuals perform worse than the spouse pairs, who in turn perform worse than the non-married pairs. The difference in performance is true overall, but also for the three questions where the correct answer is deducible. When we compare male and female individuals we find

Table 7: Decisions together: Individuals versus couples versus pairs.

Comparison	N1	N2	Mean 1	Mean 2	t-stat
Part 1					
Individual v. Couple	212	90	3.854	4.178	-2.30**
Individual v. Pair	212	86	3.854	4.511	-4.65***
Couple v. Pair	90	86	4.178	4.511	-2.25**
Couple v. Mixed Pair	90	42	4.178	4.595	-2.30**
Couple v. Male Pair	90	22	4.178	4.500	-1.30
Couple v. Female Pair	90	22	4.178	4.432	-0.80
Couple v. Female Individual	90	106	4.178	3.632	-3.35***
Couple v. Male Individual	90	106	4.178	4.075	-0.70
Female Ind. v. Male Ind.	106	106	3.632	4.075	-2.80***
Part 1 - Questions where Bag Deducible					
Individual v. Couple	212	90	2.373	2.655 8	-2.95***
Individual v. Pair	212	86	2.373	2.861	-5.30***
Couple v. Pair	90	86	2.655	2.861	-2.75***
Part 2					
Couple v. Individual	106	180	4.17	3.955	1.55
Couple v. Individual	106	172	4.17	3.837	2.80***
Couple v. Individual	106	352	4.17	3.900	2.30**
Male v. Female	176	176	4.000	3.796	1.75*
Part 2 - Questions where Bag Deducible					
Couple v. Individual	106	180	2.745	2.506	2.90***
Couple v. Individual	106	172	2.745	2.456	3.45***
Couple v. Individual (all)	106	352	2.745	2.486	3.35***
Male v. Female	176	176	2.563	2.409	1.90*

*, **, *** significant at 10, 5, 1% level, two-tailed t-test.

Table 8: Decisions together: Individuals versus couples versus pairs.

Comparison	N2	Mean 1	Mean 2	t-stat
Within pairings, between part 1 and 2				
Ind 1 vs. Couple 2	212	3.854	4.175	-3.10***
Couple 1 vs. Ind 2	180	4.178	3.955	2.00**
Pair 1 vs. Ind. 2	172	4.506	3.837	6.55***
Couple 1 vs. Male Ind. 2	90	4.178	4.045	0.85
Couple 1 vs. Female Ind. 2	90	4.178	3.866	1.95*
Pair 1 vs. Male Ind. 2	86	4.506	3.954	4.05***
Pair 1 vs. Female Ind. 2	86	4.506	3.721	5.20***
*, **, *** significant at 10, 5, 1% level, two-tailed t-test				

that men do significantly better than women in both Part 1 ($p < 0.01$) and Part 2 ($p < 0.10$). Female individuals do significantly worse than couples and pairs in both Part 1 and Part 2. Male individuals do significantly worse than pairs, but though the mean level of success for men is lower than than for couples the difference is not significant when Part 1 and Part 2 are taken separately. What is perhaps also remarkable is that individuals in Part 2 perform worse on average than when they worked as a couple in Part 1 ($p < 0.01$).

Result 2. *Married partners working together make worse judgements than non-married pairs.*

Table 7 shows that on average non-married couples do better than married couples. The performances by non-spouse pairs is significantly higher both overall and when we focus only on the three questions where the correct answer can be deduced. In the latter case, the difference is significant at the 1% level in both Part 1 and Part 2.

Result 3. *Results improve within a part, for all subject types, but do not improve between parts.*

4.2. Decisions with separation.

These results come from Parts 3-6 of the experiment. First we consider whether results improve.

Table 9: Choosers see four balls: Parts 4 and 6 versus Parts 1 and 2.

Comparison	Sample size	Treatments	Mean 1	Mean 2	t-stat
Individuals					
Part 4 v. Part 1	106	1	4.37	3.80	4.06***
Part 6 v. Part 1	106	1	3.88	3.77	0.62
Part 4 v. Part 2	176	2 and 3	3.95	4.21	-2.07**
Part 6 v. Part 2	176	2 and 3	3.93	4.18	-2.05**
Couples					
Part 4 v. Part 2	106	1	4.37	4.54	-1.26
Part 6 v. Part 2	106	1	3.88	4.52	-3.86***
Part 4 v. Part 1	176	2 and 3	3.95	4.18	-1.74*
Part 6 v. Part 1	176	2 and 3	3.93	4.17	-1.98**
*, **, *** significant at 10, 5, 1% level; two-tailed test					

Result 4. *The fraction of rational decisions falls in parts 3-6 compared to the jointly made decisions in parts 1 and 2 or compared to decisions made by individuals with some experience. However, rational choices are in parts 3-6 are more common than when inexperienced subjects make individual choices.*

One way to emphasize the decline in performance in parts 3 - 6 is to compare the rate at which individuals or teams make Bayes-rational choices in the problems where this is theoretically straightforward. In each part there are three such choices. I combine the data from parts 4 and 6, where the chooser sees four balls and the data for parts 3 and 5, where it is the recommender who sees the four balls. For both spouse pairs and non-spouse pairs, the ranking of performance is the same: joint choice yields the highest mean rate of choosing the Bayes-rational option (0.901 for spouses, 0.952 for non-spouse), then individuals (0.811 and 0.821 respectively) then the parts where the chooser sees four balls (0.758 and 0.754) and finally the parts where only two balls are seen by the chooser (0.644 and 0.554). For each type of team, the differences in each step are statistically significant. For example, for joint choice versus individual choice the relevant z-scores for a signrank test on the number of Bayes-rational choices are 5.671 and 5.844 (both p-values < 0.001), while for individual choice versus the four ball case, z-scores are 4.522 and 4.099 ($p < 0.001$) and for four balls versus two balls, the z-scores are 7.742 and 7.507 ($p < 0.001$).

Does this general fall-off in performance arise through the interaction of the two players in each team or is it more general? To answer this question I compare the optimality of the predictions in the first two parts of the game

Table 10: Recommenders see four balls: Parts 3 and 5 versus Parts 1 and 2.

Comparison	Sample size	Treatments	Mean 1	Mean 2	t-stat
Individuals					
Part 3 v. Part 1	106	1	4.01	3.77	1.41
Part 5 v. Part 1	106	1	4.00	3.80	1.31
Part 3 v. Part 2	176	2 and 3	4.05	4.18	-1.10
Part 5 v. Part 2	176	2 and 3	4.06	4.21	-1.21
Couples					
Part 3 v. Part 2	106	1	4.01	4.54	-3.05***
Part 5 v. Part 2	106	1	4.00	4.52	-3.15***
Part 3 v. Part 1	176	2 and 3	4.05	4.18	-1.08
Part 5 v. Part 1	176	2 and 3	4.06	4.17	-1.07
Within subjects t-test; *, **, *** significant at 10, 5, 1% level;					

to the recommendations made by players in parts 3 and 5. Recall that the recommender sees four balls in each of these two parts.

Comparing Tables 9 and 10, it seems that the major difference is that the recommendation accuracy in Parts 3 and 5 is slightly higher than the frequency of rational choice in parts 4 and 6. As a result, the gap in performance between parts 3-6 on the one hand and jointly made choice in parts 1 or 2 on the other hand, is generally smaller. Nevertheless, it is still the case that the separated individuals are less likely to recommend the rational option, compared to people make joint choices in part 2 especially. Overall, I summarize this as:

Result 5. *The fraction of rational recommendations in parts 3 and 5 is generally lower than the fraction of rational choices made jointly or by experienced individuals.*

In this discussion, I merge Treatments 2 and 3 on the grounds that subjects in these sessions played as individuals after initially making joint choices, whereas subjects in Treatment 1 played as individuals in the first part before moving on to make joint choices. There is a question whether the behaviour of the spouse pairs and the non-spouse pairs differs in parts 3-6. Table 11 summarizes the evidence.

Which leads to the conclusion that:

Result 6. *Pairs and couples perform equally well when separated.*

It seems therefore that there is no advantage to being married when the partners are separated.

Table 11: Decisions apart: Couples versus pairs in Parts 3-6.

Comparison	N1	N2	Mean 1	Mean 2	t-stat
Couple v. Pair, Part 3	196	86	3.408	3.209	1.10
Couple v. Pair, Part 4	196	86	4.107	3.872	1.40
Couple v. Pair, Part 5	196	86	3.709	3.511	1.15
Couple v. Pair, Part 6	196	86	3.699	3.872	-1.10
Couple v. Pair, Parts 3-6	196	86	14.924	14.465	1.2
*, **, *** significant at 10, 5, 1% level					

Table 12: Fractions Following Recommendations

Sees Four Balls				Sees Two Balls			
Blue Balls	Recommended Bag	Fraction follows	Obs.	Blue Balls	Recommended Bag	Fraction follows	Obs.
0	A	0.257	315	0	A	0.631	655
0	B*	0.823	249	0	B	0.638	1037
1	A	0.486	870	1	A	0.705	1100
1	B*	0.652	822	1	B	0.605	592
2	A*	0.772	544				
2	B	0.272	584				

Note: "Fraction follows" is the fraction of decision-makers who choose according to the recommendation. * indicates cases where Bayesian should choose the same as the recommendation.

4.3. Following and ignoring advice.

When taken together, Tables 9 and 10, suggest that the advice given in parts 3-6 is more rational than the subsequent actions. However, the tentative conclusion is based on a comparison of the choices made in parts 4 and 6, with the advice given in parts 3 and 5. I now dig deeper to examine the conditions under which subjects ignore the advice given to them. Table 12 shows that a significant fraction of people fail to choose optimally in parts 3 to 6. When faced with four yellow balls, 82.3% of choices are for Bag B, when that bag has been recommended, but only 74.3% when the partner has recommended the other bag. To put this in some contrast, faced with four yellow balls, 96.2% of treatment 3 subjects, 87.8% of treatment 2 subjects and 78.5% of treatment 1 subjects choose Bag B in part one, while in part 2, the respective figures are 85.4%, 86.7% and 94.3%. In other words, when their partner gives advice consistent with the correct answer, they opt for the correct bag in roughly the same proportion of cases as when working individually in parts 1 and 2. However, if the partner offers advice that conflicts with the correct inference,

then the rate of choosing correctly is lower by nearly 8% (p=0.022, t-test). Interestingly, the gap in choice behaviour according to recommendation is smaller but still significant at the 5% level when two balls are blue (p=0.048, t-test). At the same time, the overall rate of choosing the correct answer is lower, although this difference is not statistically significant from the mean rate of success when all the balls seen are yellow (p=0.139, t-test).[13]

4.4. Responding to Advice

Figures 2 and 3 show how the fraction of A choices varies according to the gender of the chooser, the balls they see, and the advice they receive from their partner. The top rows show the data from treatments 1 and 2 where individuals are playing with spouses. The bottom row summarizes the data when the partners are non-spouses. When choosing individuals see four balls, they tend to choose B when there are no blue balls, but still in approximately 20% of cases, bag A is nominated. The advice received is not particularly influential, except that when women are advised by their spouse to choose A rather than B, there is some effect that is absent for men.

At the other extreme, when choosing men see two blue balls in the four presented to them, their choices do not vary with the advice from their spouses. However, as with the zero blue balls case, women are more likely to choose A if their partner has recommended it. When chooser see only one blue, the responses of both sexes vary with advice. For non-spouse pairs the pattern is somewhat similar, except that when two blue balls are observed, women's response does not vary with the recommendation received whereas men are more likely to choose A if A is recommended.

Figure 3 summarizes the variation in choice of A across gender and advice received when the chooser sees only two balls. In this situation there is greater variation in the choice of A across advice. Compared to the four ball case, subjects are not particularly responsive to the number of blue balls they see.

And Figure 3 shows the pattern of responses when the chooser only sees two balls.

In all categories except one, men and women are sensitive to recommendation, in the sense that they are less likely to choose B when A is recommended. The exception is men in the non-spouses case.

[13]In a laboratory experiment on social networks with two urns Goeree and Yariv (2015) also find evidence of conformity to advice that lowers payoffs.

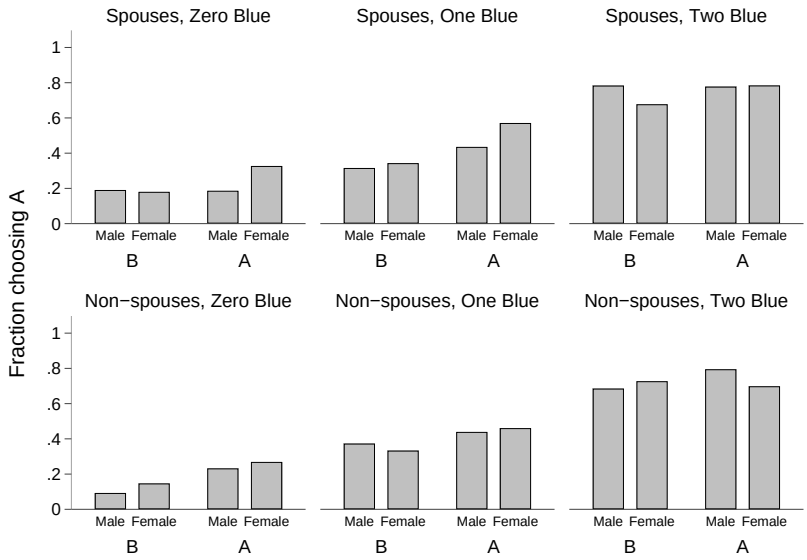


Figure 2: Choosing A After Recommendations by Number of Blue balls seen by Chooser.

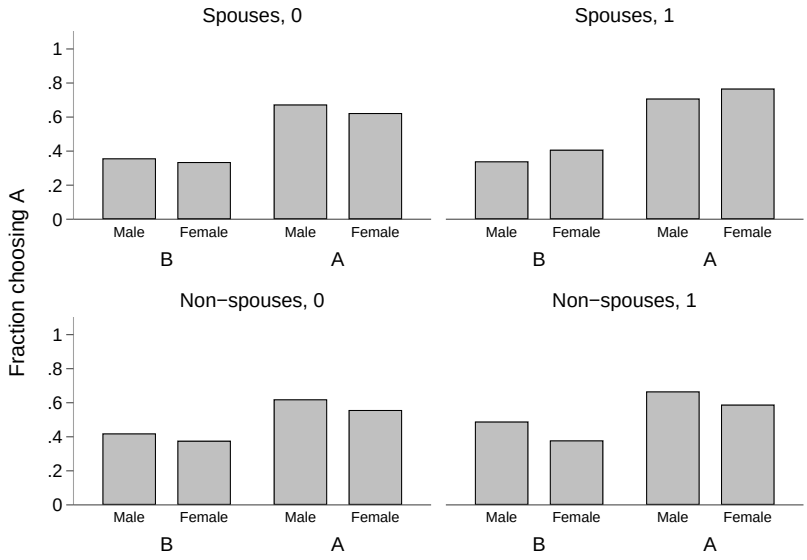


Figure 3: Choosing A After Recommendations When Two Balls Seen

Table 13: Fraction choosing A given balls seen and advice: Individuals versus couples versus pairs.

Blue balls	Gender	N1	N2	Mean 1	Mean 2	z value
0	Male	89	107	0.187	0.191	-0.073
0	Female	113	83	0.327	0.181	2.299**
1	Male	319	269	0.436	0.316	2.98***
1	Female	294	294	0.571	0.344	5.546***
2	Male	207	185	0.777	0.784	-0.134
2	Female	181	211	0.785	0.678	2.237**

1 = "A advised"; 2 = "B advised" *, **, *** significant at 10, 5, 1% level two sided test of proportions.

Tables 14 and 13 provides formal tests of the differences observable in Figure 2. In each case we are comparing the fraction choosing A given advice to choose A, versus the proportion who choose A given that B is recommended. It can be seen that wives are generally sensitive to advice, in the sense that the proportion who choose A is always significantly higher when their husbands have advised them to choose A. For men, the pattern is less clear: in the cases where the number of blue balls means that the identity of the bag can be deduced with probability 1, men's choices are not sensitive to the recommendation received.[14] However, when only one blue ball is seen, in common with their spouses, men who are advised to choose B are less likely to choose A, compared to the case where A has been recommended.

Result 7. *In the case where the chooser is the more informed partner, wives are more sensitive to advice compared to their husbands.*

For the non-spouses, obviously the power of the tests is weaker given the smaller sample sizes. With one exception, though, the proportion who choose A given A is recommended is higher than when B is suggested. In two cases, this difference is statistically significant at either the 10 or 5% level. For women who see two blue balls, being advised to choose A produces a lower proportion choosing A, compared to the effect of receiving a recommendation for B, but this difference is not significant.

[14]It's perhaps worth pointing out that in neither case are men making the rational choice with probability 1.

Table 14: Fraction choosing A, given balls seen and advice received, by gender

Blue balls	Gender	N1	N2	Mean 1	Mean 2	z value
Spouses.						
0	Male	89	107	0.187	0.191	-0.073
0	Female	113	83	0.327	0.181	2.299**
1	Male	319	269	0.436	0.316	2.98***
1	Female	294	294	0.571	0.344	5.546***
2	Male	207	185	0.777	0.784	-0.134
2	Female	181	211	0.785	0.678	2.237**
Non-Spouses.						
0	Male	43	43	0.232	0.093	1.753*
0	Female	52	34	0.269	0.147	1.335
1	Male	116	142	0.439	0.373	1.082
1	Female	141	137	0.461	0.333	2.081**
2	Male	83	89	0.795	0.685	1.637
2	Female	71	99	0.699	0.727	0.411

1 = "A advised"; 2 = "B advised" *, **, *** significant at 10, 5, 1% level two sided test of proportions.

In everyday life, not listening to partners has a pejorative meaning, but here it is not so obvious that individuals should be swayed by the recommendations of their spouses, especially in the cases where the identity of the bag should be deducible. We do not exactly know how the subjects would choose in the absence of the advice, but I note that the success rate of individuals in treatment 2 was lower in the parts of the experiment where they received advice, compared to part 2 where they made individual choices without advice. For treatment 1 individuals there is no statistically significant difference between part 1 and parts 3 and 5.

In Table 15, the same data is used to compare male and female rates of choosing A, in response to different advice and information. Comparing genders, it can be see that in general wives are more likely to choose in line with advice received, compared to their husbands, but that there are no statistically significant differences between men and women when non-spouses play the game.

Result 8. *Amongst non-spouses, neither men or women are particular sensitive to the advice given by their partners.*

When only two balls are seen, all categories of participants are sensitive to advice (see Table 16). Non-spouses are less sensitive, in the sense that

Table 15: Fraction choosing A given balls seen and advice received from spouse: females versus males.

Blue balls (Rational choice)	Advice	N Male	N Female	Mean Male	Mean Female	z value
Spouses.						
0 (B)	A	113	107	0.187	0.327	2.378**
0 (B)	B	89	83	0.181	0.191	0.173
1 (B)	A	319	294	0.436	0.571	3.357***
1 (B)	B	269	294	0.316	0.344	0.694
2 (A)	A	207	181	0.777	0.785	0.160
2 (A)	B	185	211	0.784	0.678	-2.364**
Non Spouses.						
0 (B)	A	43	52	0.232	0.269	-0.409
0 (B)	B	43	34	0.093	0.147	-0.733
1 (B)	A	116	141	0.439	0.461	-0.342
1 (B)	B	142	137	0.373	0.333	0.668
2 (A)	A	83	71	0.795	0.699	1.390
2 (A)	B	89	89	0.685	0.727	-0.6302

*, **, *** significant at 10, 5, 1% level two sided test of proportions.

their probability of choosing A is less sensitive to differences in the advice, compared to corresponding figures for spouses. However, for all groups, the rate at which the advice is ignored is around 30-40% . Since the overall rate of giving rational advice in these cases is 68%,[15] while the rate at which the choosing individuals make the rational choice is only 58%, then ignoring the advice is costly for many households. When both the balls seen are yellow, the probability that the bag contains two blue balls is 1/3, whereas when one ball is blue the probability is 3/5. Two yellow balls are therefore slightly more informative than one blue ball, but there is no clear evidence that individuals are less sensitive to advice when no blue balls are present.

Result 9. *When only two balls are seen, all categories of participants are sensitive to advice.*

[15]This may seem low, but actually 55% of subjects always give the correct advice when the bag identity is deducible.

Table 16: Fraction choosing A, given balls seen and advice received, by gender

Blue balls	Gender	N1	N2	Mean 1	Mean 2	z value
Spouses.						
0	Male	233	355	0.674	0.358	-7.502***
0	Female	219	378	0.624	0.336	-6.739***
1	Male	388	200	0.709	0.340	-8.593***
1	Female	387	201	0.767	0.408	-8.638***
Non-Spouses.						
0	Male	108	150	0.620	0.420	-3.176***
0	Female	104	154	0.558	0.377	-2.868***
1	Male	162	96	0.667	0.490	-2.807***
1	Female	163	95	0.589	0.379	-3.255***

1 = "A advised"; 2 = "B advised" *, **, *** significant at 10, 5, 1% level two sided test of proportions.

4.5. Regression analysis.

In this section, I extend the analysis, allowing for controls. First, I consider the correlates of making rational choices in parts 1 and 2. Table 17 confirms that when individuals make choices in parts 1 and 2, rational choice is more likely when the decision-maker sees zero or two blue balls, compared to the situation where one is encountered. In fact, there is some evidence that seeing two blue balls is harder to process successfully compared to seeing none, but this feature of the data is only clear-cut in part 1. In individual decisions, women do worse than men in part 1, but in part 2 there is no difference between the men and women. Older and less educated individuals make fewer rational choices, as do those who achieve a lower score in a four item arithmetic test included in the pre-experiment survey.

The correlates of rational choices and rational advice in parts 3 to 6 are analyzed using a logit model in Table 18. For equations (1) to (3), the rationality of the choice is the dependent variable, while for the other three equations the rationality of the advice is on the left-hand side. Within each group of three equations, the first column includes all the questions. The second equation includes the questions where, faced with four balls, perfect deduction of the bag is possible. The third equation in each group uses data only from those questions where even with four balls, the identity of the bag is uncertain.

Recall that in parts 4 and 6, the choosing agent sees four balls, while the

Table 17: *Rational choice in parts 1 and 2*

	(1) Part 1	(2) Part 2	(3) Individuals	(4) All
One Blue	-1.500***	-1.619***	-1.522***	-1.974***
	(0.165)	(0.164)	(0.115)	(0.091)
Two Blue	0.106	-0.369**	-0.188	-0.297***
	(0.226)	(0.179)	(0.135)	(0.112)
Female	-1.456***	0.165	-0.662**	-0.677***
	(0.403)	(0.428)	(0.305)	(0.237)
Age	-0.023***	-0.019***	-0.020***	-0.018***
	(0.008)	(0.007)	(0.005)	(0.004)
Female × Age	0.030***	-0.007	0.007	0.006
	(0.010)	(0.010)	(0.007)	(0.005)
Maths score	0.128**	0.135**	0.125***	0.089***
	(0.057)	(0.054)	(0.040)	(0.030)
Highest level of education	0.267**	0.199*	0.224***	0.152**
	(0.123)	(0.120)	(0.085)	(0.062)
Treatment 3		0.410**		0.247
		(0.193)		(0.158)
Part 2			0.557***	0.499***
			(0.181)	(0.128)
Treatment 2			-0.400*	0.042
			(0.205)	(0.157)
Part 2 × Female			0.090	0.463**
			(0.224)	(0.187)
Treatment 2 × Female			0.397	0.512***
			(0.258)	(0.197)
Part 2 × Treatment 2				-0.343
				(0.217)
Part 2 × Treatment 3				-0.245
				(0.199)
Treatment 3 × Female				0.747***
				(0.213)
Part 2 × Treatment 2 × Female				-0.473
				(0.296)
Part 2 × Treatment 3 × Female				-1.095***
				(0.296)
Constant	1.768***	2.210***	1.795***	2.021***
	(0.525)	(0.524)	(0.377)	(0.291)
Observations	1272	2112	3384	6768

Notes. 'Part 1' sample consists of treatment 1, while 'part 2' is treatments 2 and 3. 'Individuals' puts together all individual choices, while 'all' includes the choices made by teams, in which case there are two data points for every decision - one for each partner. ***, **, * indicate two-sided significance levels at 1, 5, and 10 %, respectively; robust standard errors (in parentheses); dependent variable takes value 1 if agent makes Bayes optimal choice, 0 otherwise. Fixed effects for village omitted.

advising agent sees only two. Consequently, it is perhaps no surprise that

Table 18: *Rational choice and advice, parts 3-6*

	Choice			Advice		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	Easy	Difficult	All	Easy	Difficult
One Blue × Sees four=1	-0.900***			-0.963***		
	(0.114)			(0.114)		
No blues for two	0.191**		0.192**	-0.168**		-0.168**
	(0.074)		(0.075)	(0.069)		(0.069)
Part 4	0.872***	1.127***	0.086	-1.546***	-1.437***	-0.577***
	(0.129)	(0.147)	(0.117)	(0.140)	(0.148)	(0.117)
Part 5	-0.098	0.420***	-0.577***	-0.011	0.020	-0.032
	(0.077)	(0.116)	(0.111)	(0.087)	(0.130)	(0.117)
Part 6	0.731***	1.038***	-0.097	-1.605***	-1.716***	-0.418***
	(0.126)	(0.141)	(0.113)	(0.135)	(0.142)	(0.113)
Mixed	-0.172*	-0.344***	-0.029	-0.209**	-0.231*	-0.193
	(0.094)	(0.125)	(0.147)	(0.093)	(0.118)	(0.125)
Two women	0.010	-0.171	0.163	-0.298***	-0.282*	-0.317**
	(0.111)	(0.176)	(0.152)	(0.110)	(0.163)	(0.150)
Two men	0.122	0.030	0.205	0.033	-0.027	0.083
	(0.122)	(0.161)	(0.182)	(0.132)	(0.173)	(0.159)
Female	0.334*	0.269	0.401*	0.132	0.245	0.040
	(0.175)	(0.274)	(0.242)	(0.182)	(0.241)	(0.242)
Age	0.003	0.000	0.005	0.003	0.003	0.002
	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
Female × Age	-0.008**	-0.006	-0.010	0.001	-0.002	0.003
	(0.004)	(0.006)	(0.006)	(0.004)	(0.006)	(0.006)
Individual deduction	0.376***	0.475***	0.299*	0.221*	0.259*	0.188
	(0.121)	(0.156)	(0.176)	(0.119)	(0.152)	(0.153)
Partner's deduction	0.389***	0.620***	0.193	0.041	0.062	0.025
	(0.110)	(0.161)	(0.162)	(0.121)	(0.151)	(0.159)
Education	0.192***	0.201**	0.190**	0.217***	0.285***	0.165**
	(0.059)	(0.088)	(0.076)	(0.053)	(0.077)	(0.070)
Women should obey	0.025	0.073	-0.012	0.039	0.014	0.060
	(0.042)	(0.058)	(0.060)	(0.046)	(0.063)	(0.054)
Trust	-0.011	-0.004	-0.018	-0.005	-0.028	0.014
	(0.022)	(0.032)	(0.031)	(0.022)	(0.030)	(0.028)
Conform	0.029	-0.014	0.067*	-0.032	-0.041	-0.026
	(0.030)	(0.043)	(0.040)	(0.029)	(0.037)	(0.038)
Dislike teamwork	0.077*	0.107*	0.057	0.020	-0.011	0.046
	(0.044)	(0.064)	(0.058)	(0.047)	(0.062)	(0.058)
Constant	-1.019***	-1.660***	-0.835**	0.451	0.418	-0.492
	(0.312)	(0.421)	(0.420)	(0.281)	(0.393)	(0.359)
Observations	6768	3384	3384	6768	3384	3384

Notes. ***, **, * indicate two-sided significance levels at 1, 5, and 10 %, respectively; robust standard errors (in parentheses); dependent variable takes value 1 if agent makes Bayes optimal choice or advice, 0 otherwise. "Women should obey" takes highest value (4) if person strongly disagrees; "Trust" takes highest value (5) if most people can be trusted; "Conform" takes highest value (4) if person strongly disagrees that they often change their opinion to please others; "Dislike teamwork" takes highest value (4) when person strongly disagrees that they prefer working in teams to alone.

the rate of choosing rationally is higher in parts 4 and 6, compared to parts

3 and 5.

Compared to pairs consisting of spouses, mixed pairs perform worse, notably in the easier questions, where the rate of choosing rationally and advising correctly is lower, compared to other types of pairs. Two women groups also do worse than spouse pairs when it comes to advising, though not in making the final choices. Unlike parts 1 to 2, age does not reduce performance, except that there is some evidence that women's performance decreases with age, relative to men. However, overall conditional on other factors, women generally make more rational choices and offer more rational advice. Meanwhile, higher levels of education are associated with more rational choice and advice, across both easy and difficult questions. In general, there is no clear assocation between the answers to the attitude and belief questions in the survey and the probability of giving a rational answer or advice. The exception is the 'prefer working in a team' question where subjects who declare that they prefer to work alone are more likely to make a rational choice. There is no similar pattern in the advice data.

Result 10. *More highly educated people are more likely to give good advice. Players in mixed, non-spouse pairs are less likely to give Bayes-rational advice.*

Table 19 shows several logit models that provide further evidence on the correlates of following the advice given. In all cases, the dependent variable takes the value 1 if the chooser follows the advice and 0 otherwise. The first equation combines all the available data from parts 3-6, while the two other equations separate the analysis into cases where the chooser sees two balls and the parts where the chooser sees four balls. Looking at the first equation, in parts 4 and 6, the chooser faces four balls and the advisor sees only two, so it is not surprising that in these parts of the experiment the choosing agent is less likely to follow the advice given. Meanwhile, players in mixed teams are less likely to follow advice compare to other groups.

Table 19 shows that women in general are more likely to follow advice received, compared to men. Yet, the picture is more complicated because age affects men and women differently. Table 20 shows marginal effects of gender using the 'Two balls' and 'Four balls' equations, at various ages, across the parts of the experiment.

Older men are more likely to follow advice from their partners. However, there is no such trend for women. If anything, older women are less likely to follow advice. Thus, in our sample, the gap between men and women's

Table 19: *Following advice*

	(1) All	(2) Two balls	(3) Four balls
Part 4	-0.693***		
	(0.146)		
Part 5	-0.177	-0.171	
	(0.129)	(0.128)	
Part 6	-0.738***		-0.054
	(0.134)		(0.121)
Female	0.198	0.193	0.698**
	(0.226)	(0.303)	(0.289)
Part 4 × Female	0.488***		
	(0.180)		
Part 5 × Female	0.560***	0.534***	
	(0.186)	(0.186)	
Part 6 × Female	0.248		-0.208
	(0.175)		(0.174)
One Blue × Sees four=1	0.218***		0.219***
	(0.082)		(0.082)
Mixed	-0.281***	-0.388***	-0.180
	(0.099)	(0.143)	(0.135)
Two women	-0.222	-0.132	-0.307
	(0.136)	(0.185)	(0.188)
Two men	-0.101	-0.234	0.021
	(0.109)	(0.166)	(0.184)
Age	0.008***	0.008*	0.008*
	(0.003)	(0.005)	(0.005)
Female × Age	-0.010**	-0.012*	-0.009
	(0.005)	(0.007)	(0.007)
Individual deduction	0.146	0.073	0.213
	(0.128)	(0.179)	(0.192)
Partner's deduction	-0.034	0.219	-0.268
	(0.128)	(0.172)	(0.175)
Highest level of education	0.041	0.103	-0.015
	(0.058)	(0.089)	(0.075)
Women should obey	0.016	-0.020	0.047
	(0.043)	(0.070)	(0.059)
Trust	-0.033	-0.068**	-0.002
	(0.023)	(0.035)	(0.032)
Conform	0.007	0.071	-0.050
	(0.032)	(0.046)	(0.043)
Dislike teamwork	0.036	0.065	0.011
	(0.045)	(0.068)	(0.063)
Constant	0.137	-0.156	-0.291
	(0.339)	(0.457)	(0.482)
Observations	6768	3384	3384

Notes. ***, **, * indicate two-sided significance levels at 1, 5, and 10 %, respectively; robust standard errors (in parentheses); dependent variable takes value 1 if agent follows the advice, 0 otherwise. No. of blue balls is the number seen when chooser sees four balls.

Table 20: *Following advice - Marginal Effects for Gender.*

	(1) Part 3	(2) Part 4	(3) Part 5	(4) Part 6
1.At aged 20	0.130***	0.079*	-0.011	0.109**
	(0.045)	(0.047)	(0.045)	(0.044)
2.At aged 30	0.108***	0.057	-0.039	0.082**
	(0.037)	(0.037)	(0.037)	(0.037)
3. At aged 40	0.086**	0.035	-0.066**	0.055
	(0.034)	(0.033)	(0.033)	(0.035)
4. At aged 50	0.064	0.014	-0.093**	0.028
	(0.039)	(0.037)	(0.037)	(0.039)
5. At aged 60	0.043	-0.008	-0.120***	0.001
	(0.050)	(0.047)	(0.046)	(0.047)
Observations	3384	3384	3384	3384

Notes. Table shows difference in probability of following advice for women compared to men. ***, **, * indicate two-sided significance levels at 1, 5, and 10 %, respectively; robust standard errors (in parentheses); evaluated at means of all other variables.

behaviour is largest for the youngest people. The pattern is most closely followed for parts 3, 4 and 6. For part 5, the age trend for women is the same, but because young men and women are equally likely to follow advice, the net effect is that older women are less likely to follow the advice received, compared to older men. Is the pattern of following adaptive for either sex, in the sense of raising payoffs? In general, it is optimal to follow the advice of the partner in Parts 3 and 5 and to ignore it in Parts 4 and 6. Relative to men, therefore, women's behaviour is not adaptive in part 5, but adaptive in parts 3, 4 and 6.

Result 11. *Women are more likely to follow the advice received than men.*

Result 12. *Non-spouses are more likely to ignore the advice of their partners, compared to spouse pairs. People with more education are more likely to play in line with Bayes theorem.*

4.6. Player Types.

In Table 5 the behaviour consistent with different types of players was outlined. One particular implication of that table is that players who conform to de Groot should follow a Bayes-rational recommendation when they see four balls. In fact, when faced with four balls, 50.71% of subjects never go

against the Bayes-rational choice when it aligns with the recommendation they receive. In addition 30.7% make only one choice against Bayes-rational advice and only one person makes six choices (the maximum) that all go against the Bayes-rational advice that has been received.[16] The dependent variable in the final column of Table 21 is the number of times (from 0 to 6) that a choosing person faced with four balls, goes against Bayes-rational advice.[17] As can be seen, in general, very little is predictive, but being female is associated with fewer occasions in which good advice is ignored and the same is true for people woh make more rational choices in the individual decisions part of the experiment. Older women are however, likely to make more contrary decisions. Measures of trust, attitudes to female obedience and conformity are not associated with ignoring good advice.

The other three columns of regression results use the same right hand side variables, but the dependent variable is the number of times the choosing person who sees four balls conforms to advice which is not Bayes-rational. For the first equation the dependent variable is the total number of times; for the second equation, only those questions where the identity of the bag can be deduced are included, while the third occasion uses only the questions where the chooser sees only one blue ball. As with the final equation, very few right hand variables show a statistical link to the dependent variable. Women are more likely to show conform to the advice and so are older people in general. There is some evidence that more educated people conform less when offered bad advice, particularly when they face a more clear-cut inference problem. The same is true for people who scored more highly in the individual decisions part of the experiment.

To bring together behaviour in the various parts of the experiment, I create two types of classification as follows. Neither classification is intended to be definitive. Rather they are ways of summarizing behaviour that cast light on the results examined above. Consider first the pattern of following. As indicated above, in general it is optimal to follow advice when the chooser sees only two balls and to discount the advice when the person sees four balls. Nevertheless, when faced with four balls the received advice might be good, in the sense that it coincides with the rational choice by a Bayesian. In such

[16]These figures need to be interpreted in light of the fact that only 16% of choosers receive six Bayes-rational choices and the mean number is 4.03.

[17]I report results for OLS. Results for a Tobit model and for ordered logit show the same pattern of statistically significant effects.

Table 21: *Player Types - De Groot.*

	(1) DeGroot?	(2) DeGroot. Clear	(3) DeGroot. Difficult	(4) Ignore good advice
Female	0.776***	0.259*	0.517**	-0.645***
	(0.209)	(0.140)	(0.193)	(0.193)
Age	0.011**	0.005*	0.006	-0.006
	(0.005)	(0.003)	(0.004)	(0.004)
Female × Age	-0.008	-0.002	-0.005	0.011**
	(0.006)	(0.005)	(0.005)	(0.005)
Partner's age	-0.004	-0.004*	0.000	0.004
	(0.003)	(0.002)	(0.003)	(0.003)
Education	-0.118	-0.111**	-0.006	-0.097
	(0.077)	(0.047)	(0.050)	(0.097)
Individual deduction	-0.433	-0.319*	-0.114	-0.592*
	(0.253)	(0.167)	(0.134)	(0.300)
Partner's deduction	-0.026	-0.043	0.017	-0.084
	(0.203)	(0.104)	(0.169)	(0.187)
Mixed	0.117	0.130	-0.013	0.007
	(0.122)	(0.095)	(0.066)	(0.111)
Two women	-0.284	-0.122	-0.161	0.042
	(0.207)	(0.092)	(0.171)	(0.144)
Two men	-0.042	0.026	-0.069	-0.057
	(0.187)	(0.065)	(0.170)	(0.122)
Land (acres)	-0.022	0.009	-0.031	-0.016
	(0.026)	(0.008)	(0.023)	(0.026)
Women should obey	-0.033	-0.056*	0.023	0.009
	(0.051)	(0.027)	(0.044)	(0.053)
Trust	0.011	0.003	0.008	-0.021
	(0.034)	(0.015)	(0.027)	(0.036)
Conform	-0.048	0.008	-0.056*	-0.004
	(0.045)	(0.022)	(0.029)	(0.039)
Dislike teamwork	0.025	0.013	0.011	-0.002
	(0.054)	(0.047)	(0.042)	(0.071)
Constant	1.569***	0.893***	0.676**	1.850***
	(0.434)	(0.253)	(0.255)	(0.536)
R^2	0.073	0.073	0.044	0.049
Observations	562	562	562	562

Notes. ***, **, * indicate two-sided significance levels at 1, 5, and 10 %, respectively; robust standard errors (in parentheses).

circumstances it is impossible to judge where the chooser is simply following blindly or making a judgement for him or herself. However, when the advice is bad - then such a judgement can be made. So, for the case where four balls are seen, I ignore cases where following advice would be optimal. I classify players into four types (see Table 22) based on the pattern of following for two and four balls. For example a player is classified as 'Bayesian' if the person follows advice in over three cases when they see two balls, and does not follow any bad advice when they see four balls. In contrast, a person is

classified as an 'Ignorer' if they ignore bad advice in the four ball case and also depart from the advice received in more than three cases when they face two balls.

For the second classification, the two dimensions are about giving rational advice when the player sees four balls and following advice when the player sees only two balls. When, for example, a player gives rational advice on more than four occasions, but follows advice on three or fewer occasions, he or she is labelled as 'Good advice' and so on.

Table 22: *Classifying Player Types.*

Following advice

		Follows when two balls	
		> 3	3 or less
Follows bad advice when four balls	0	Bayes	Ignorer
	> 0	Follower	Reverse

Team playing

		Follows when two balls	
		> 3	3 or less
Follows bad advice when four balls	0	Bayes	Ignorer
	> 0	Follower	Reverse

Given these classification I then use multinomial logit to see the factors associated with each player type. Table 23 provides marginal effects for certain key variables. The left hand side of the table provides marginal effects for the first classification system and the right hand side refers to the second taxonomy. For the following case, being female raises the probability of being in the follower category and significantly lowers the probability of being classified as a Bayesian. A higher maths score is not significantly associated with membership of any particular category, but education is linked to a higher probability of being a Bayesian. Finally, being assigned to the non-spouse treatment increases the probability of playing the game in the opposite way to a Bayesian and decreases the probability of being a follower.

Women are more likely to be classified as providing good advice and listening, but at the same time, less likely to be in the listens (but provides poor advice) category compared to men. A higher maths score is associated with a higher chance of providing poor advice and not-listening; at the same time, more educated people are less likely to be in the poor category and more likely to be in the 'listens, good advice' group. Treatment 3 players are

less likely to be placed in this category, but at the same time they are more likely to be in the giving good advice (poor at respecting advice) category.

Result 13. *Play does not conform to a simple de Groot model in which choosers put some weight on Bayesian predictions and some weight on advice received.*

5. Conclusions.

In this paper I provide the first experimental evidence in economics on how married couples process information in a shared decision environment. The experiment is designed around a task in which participants are asked to predict the identity of a bag, based on evidence about its contents. I find that individuals fare less well than people playing in dyads and that spouses are less effective than unrelated partners in making joint decisions. In terms of the hypotheses set out in Section 3, I find that hypothesis 1 is supported - 'two heads are better than one', but hypothesis 2 is rejected - day to day intimacy does not seem to aid profitable decision-making. Hypothesis 3 is partially supported. The accuracy of predictions increases within a round and people who play as individuals in Part 2 predict more accurately than individual predictors in Part 1. However, these Part 2 individuals do not predict as well as when they previously played in a pair, suggesting that some aspects of the game are not fully absorbed by all participants in teams.

When the pairs are split up and placed in different rooms, partners are asked to pass recommendations to the other person in the team, who then chooses. In some cases the chooser knows more than the recommender and in other cases it is the other way round. In these rounds I find no support for the view that being married to the other person in a team gives a decision-making advantage. Thus hypothesis 1 of Parts 3-6 is not accepted. I do find evidence in support of the hypotheses (2 and 3) that people overweight their own evidence and that they do put more weight on evidence from four balls compared to that from two balls. Whether the recommender has less or more knowledge than the chooser, I do find evidence that the recommendation regularly affects the final decision of the choosing person. Nevertheless, even when they understand that their partner knows more than them about the bag, players within the experiment routinely go against the advice they receive. In general, women are more sensitive to the advice received than men, meaning that hypothesis 4 is supported. Meanwhile, men are more sensitive

Table 23: *Player Types - Marginal Effects for Key Variables.*

	(1) Following Types		(2) Advising, Following
	Female		
Ignorer	-0.036	Poor	0.002
	(0.033)		(0.033)
Bayes	-0.072**	Good Advice	-0.058
	(0.033)		(0.051)
Reverser	-0.021	Listens	-0.066*
	(0.040)		(0.034)
Follower	0.129**	Listens, Good Advice	0.122**
	(0.051)		(0.057)
	Maths Score		
Ignorer	0.010	Poor	0.094**
	(0.060)		(0.042)
Bayes	0.063	Good Advice	-0.039
	(0.064)		(0.057)
Reverser	0.038	Listens	0.070
	(0.067)		(0.056)
Follower	-0.112	Listens, Good Advice	-0.125
	(0.087)		(0.084)
	Education		
Ignorer	-0.016	Poor	-0.056**
	(0.019)		(0.026)
Bayes	0.049**	Good Advice	-0.013
	(0.024)		(0.028)
Reverser	-0.048	Listens	-0.047
	(0.037)		(0.034)
Follower	0.015	Listens, Good Advice	0.116**
	(0.038)		(0.046)
	Non-spouses Treatment		
Ignorer	0.029	Poor	0.037
	(0.029)		(0.034)
Bayes	-0.009	Good Advice	0.122***
	(0.036)		(0.039)
Reverser	0.130***	Listens	-0.025
	(0.038)		(0.037)
Follower	-0.150***	Listens, Good Advice	-0.134**
	(0.053)		(0.068)
Observations	562		562

Notes. Table shows difference in probability of falling in each category compared to baseline category, except Education and Maths score where it is the derivative. The other variables used in the estimation are the same as for Table 19. ***, **, * indicate two-sided significance levels at 1, 5, and 10 %, respectively; robust standard errors (in parentheses); evaluated at means of all other variables.

to whether the partner is more or less informed about the bag. Hence it is difficult to claim that hypothesis 5 is supported. Generally, spouses are more sensitive to advice than non-spouses, but because this works to the disadvantage of the team when the chooser plays less, overall spouse and

non-spouse pairs have similar payoffs in parts 3=6. There is some evidence to support the view that men overweight their information when playing with non-spouses, but the real lesson of the non-spouse teams is that mixed sex groups perform significantly worse than their single sex counterparts.

I provide evidence that most players in this game do not conform to either a Bayesian or a simple de Groot model. With respect to the former, most players make the wrong inferences at some stage, even with problems where the solution is theoretically clear. Moreover, most often this is because players place sufficient weight on (erroneous) advice of their partners that the Bayes-rational choice is rejected. Such behaviour suggests a de Groot style model, but a simple version of de Groot can also be rejected, because the majority of players also reject the Bayes rational option in some cases, even when their partner has recommended it. Perhaps the most appropriate model for most players is one of unconfident Bayesians: people who are mistrustful of their own inference abilities and hence put weight on their partners, but then also make mistakes, especially when faced with the harder problems.

Stepping back from the details of the results, there are many points arising from the experiment to ponder. Why are the unmarried pairs better than married couples at making the Bayes rational decision when they are sat together but worse when they are in different rooms and must communicate in an abbreviated way? As I emphasised above, rational advice is somewhat worse for some non-married pairings (mixed sex and all women) but does not explain the gap in performance, which is confined to mixed sex groups. After all, same sex couples do as well if not better than married couples.[18]

What about the robustness of these results? One can only speculate about whether different groups of participants would behave in a different way. One factor that may affect the results here is the relatively neutral and abstract nature of the task. A decision that is gendered in some way (e.g. for a problem that is normally seen as the domain of one sex rather than the other), may evoke different patterns of respect for partners' information.

Such factors may also play a role in understanding the wider implications of the data. In common with many other agricultured-based societies, in the region where the experiment was played, spouses typically share some decisions and work together on particular crops (Iversen et al. (2011)), albeit

[18]Working with non-relatives of the opposite sex is quite rare and this may account for the lower performance of mixed-sex pairs.

often in different roles. Cooperation is therefore common,[19]. Ben-Porath (1980) argues that one efficiency advantage of intra-household production rather is that within the family individuals are more able to trust information. The experiment in this paper does not constitute a full test of Ben-Porath's wide-ranging conjectures, but the results do suggest that there is no marriage premium in decision-making that makes intra-household cooperation more efficient than cooperation outside the family.

References

Anderson, L.R., Holt, C.A., 2008. Information cascade experiments. Handbook of experimental economics results 1, 335–343.

Ashraf, N., 2009. Spousal control and intra-household decision making: An experimental study in the Philippines. The American Economic Review 99, 1245–1277.

Bateman, I.J., Munro, A., 2005. An experiment on risky choice amongst households. The Economic Journal 115, C176–C189.

Ben-Porath, Y., 1980. The f-connection: Families, friends, and firms and the organization of exchange. Population and development review , 1–30.

Berger, S., Feldhaus, C., Ockenfels, A., 2018. A shared identity promotes herding in an information cascade game. Journal of the Economic Science Association 4, 63–72.

Bernard, T., Doss, C., Hidrobo, M., Hoel, J., Kieran, C., 2018. Ask me why: Patterns of intrahousehold decision-making. IFPRI Discussion Paper 1784 .

Birnbaum, M.H., Mellers, B.A., 1983. Bayesian inference: Combining base rates with opinions of sources who vary in credibility. Journal of Personality and Social Psychology 45, 792.

Blinder, A.S., Morgan, J., 2005. Are two heads better than one? monetary policy by committee. Journal of Money, Credit and Banking 37, 789–811. URL: `http://www.jstor.org/stable/3839147`.

[19]So is spousal conflict and deception, particularly over family planning and personal spending (Verschoor (2008); Jackson (2013)).

Castilla, C., Walker, T., 2013. Is ignorance bliss? The effect of asymmetric information between spouses on intra-household allocation. American Economic Review Papers and Proceedings 103.

Çelen, B., Kariv, S., 2004. Distinguishing informational cascades from herd behavior in the laboratory. American Economic Review 94, 484–498.

Charness, G., Karni, E., Levin, D., 2007. Individual and group decision making under risk: An experimental study of Bayesian updating and violations of first-order stochastic dominance. Journal of Risk and Uncertainty 35, 129–148. URL: https://doi.org/10.1007/s11166-007-9020-y, doi:10.1007/s11166-007-9020-y.

Charness, G., Sutter, M., 2012. Groups make better self-interested decisions. Journal of Economic Perspectives 26, 157–76.

Chiappori, P.A., Ekeland, I., 2009. The microeconomics of efficient group behavior: Identification1. Econometrica 77, 763–799.

De Dreu, C.K., Weingart, L.R., 2003. Task versus relationship conflict, team performance, and team member satisfaction: a meta-analysis. Journal of applied Psychology 88, 741.

DeGroot, M.H., 1974. Reaching a consensus. Journal of the American Statistical Association 69, 118–121.

Fishchler, I., Sorkin, R.D., 2006. Assessing and improving team decision making. Florida University Gainesville, Dept. of Psychology .

Gale, D., Kariv, S., 2003. Bayesian learning in social networks. Games and Economic Behavior 45, 329–346.

Goeree, J.K., Yariv, L., 2011. An experimental study of collective deliberation. Econometrica 79, 893–921.

Goeree, J.K., Yariv, L., 2015. Conformity in the lab. Journal of the Economic Science Association 1, 15–28.

Grether, D.M., 1980. Bayes rule as a descriptive model: The representativeness heuristic. The Quarterly Journal of Economics 95, 537–557.

Hoel, J.B., 2015. Heterogeneous households: A within-subject test of asymmetric information between spouses in Kenya. Journal of economic behavior & organization 118, 123–135.

Huck, S., Oechssler, J., 2000. Informational cascades in the laboratory: Do they occur for the right reasons? Journal of Economic Psychology 21, 661–671.

Iversen, V., Jackson, C., Kebede, B., Munro, A., Verschoor, A., 2011. Do spouses realise cooperative gains? experimental evidence from rural Uganda. World Development 39, 569–578.

Jackson, C., 2013. Cooperative conflicts and gender relations: Experimental evidence from southeast uganda. Feminist Economics 19, 25–47.

Katz, E., 1997. The intra-household economics of voice and exit. Feminist economics 3, 25–46.

Kebede, B., Munro, A., Tarazona-Gomez, M., Verschoor, A., 2014. Intra-household efficiency: An experimental study from Ethiopia. Journal of African Economies 23, 105–150.

Kocher, M.G., Sutter, M., 2004. The decision maker matters: Individual versus group behaviour in experimental beauty-contest games. The Economic Journal 115, 200–223.

Kraemer, C., Nöth, M., Weber, M., 2006. Information aggregation with costly information and random ordering: Experimental evidence. Journal of Economic Behavior & Organization 59, 423–432.

Liberman, V., Minson, J.A., Bryan, C.J., Ross, L., 2012. Naive realism and capturing the wisdom of "dyads". Journal of Experimental Social Psychology 48, 507–512.

Mercier, H., Yama, H., Kawasaki, Y., Adachi, K., Van der Henst, J.B., 2012. Is the use of averaging in advice taking modulated by culture? Journal of Cognition and Culture 12, 1–16.

Munro, A., 2018. Intra-household experiments: A survey. Journal of Economic Surveys 32, 134–175.

Navran, L., 1967. Communication and adjustment in marriage. Family Process 6, 173–184.

Peters, H.E., Ünür, A.S., Clark, J., Schulze, W.D., 2004. Free-riding and the provision of public goods in the family: A laboratory experiment. International Economic Review 45, 283–299.

Reinig, B.A., Horowitz, I., Whittenburg, G.E., 2015. Choice shifts in small groups engaging in repeated intellective tasks. Small Group Research 46, 131–159. doi:10.1177/1046496414566117, arXiv:https://doi.org/10.1177/1046496414566117.

Sharvit, G., Sproten, A.N., 2012. To take or not to take an advice? a framework for advice taking and delegation in uncertain decisions. Developmental Neuroeconomics: Lifespan Changes in Economic Decision Making , 111.

Sillars, A.L., Scott, M.D., 1983. Interpersonal perception between intimates: An integrative review. Human Communication Research 10, 153–176.

Spiro, R.L., 1983. Persuasion in family decision-making. Journal of Consumer Research 9, 393–402.

Tost, L.P., Gino, F., Larrick, R.P., 2012. Power, competitiveness, and advice taking: Why the powerful don't listen. Organizational Behavior and Human Decision Processes 117, 53–65.

Verschoor, A., 2008. Gender relations and female labour supply in east uganda, in: Work, Female Empowerment and Economic Development. Routledge, pp. 159–188.

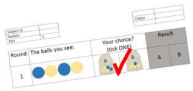
Verschoor, A., Kebede, B., Munro, A., Tarazona, M., 2019. Spousal control and efficiency of intra-household decision-making: Experiments among married couples in india, ethiopia and nigeria. The European Journal of Development Research 31, 1171–1196.

Yaniv, I., 2004. Receiving other people's advice: Influence and benefit. Organizational Behavior and Human Decision Processes 93, 1–13.

Appendix 1. Instructions. Not for publication

Treatment 1 Part 1 Individual decision Everyone sees 4 balls

[on the wall you should have the pictures of the two bags and the poster of how not to fill in the answer sheets



; make sure a second enumerator is near you to hold the bags and balls]

Welcome. Thank you for taking the time to come today. My name is …… and I am working with a university called GRIPS. This University is from Japan. We have invited you here, today, because we want to learn about how people in this area take decisions. You are going to be asked to take decisions that will earn you some money. The money that results from your decisions will be yours to keep; you can take it home with you.

Can I check that everyone has a card with number on it. Please keep this card with you for the whole experiment. In the room there's someone else with the same number as you. That person should be your spouse. Sometimes in this experiment you'll be working with them.

What you need to do will be explained fully in a few minutes. But first, there are three things I would like to explain to you clearly and you should consider them as very important.

> First of all, this is not our money. As I told you before, we work for a university and this money has been given to us by that university for this research.
>
> Participation is voluntary. You may still choose not to participate in the exercise.
>
> We also have to make clear that this is research about <u>your</u> decisions. Therefore I will not allow you to go out to talk with anyone else. This is very important. I'm afraid that if I you decide to talk with someone else, that will be the end of the study and what this means is that you will not be able to earn any money here today. The exception is that sometimes we will let you talk to your spouse. We'll tell you when that's okay. If you have any questions, please raise your hand and ask me. Could I also ask you to switch off your mobile phone if you have one?

Make sure that you listen carefully to me. You will be able to make a good amount of money here today, and it is important that you follow my instructions.

During today's programme, you will be asked to make a few decisions, which will be explained to you clearly. Every decision you make can make you some money if you get it right. So it will be important to think about every decision. In some cases you will be making a decision on your own, in which case any money you win will be yours. In some cases you will be making a decision with your spouse. When you are making a decision with your spouse, any money you make will be shared equally between you. But, you will be paid separately.

You can see on this poster two pictures of bags like the one I am holding (hold up a bag). Bag A has five balls in it of different colours. Three balls are yellow and two are blue. Bag B is a bit different: it also has five balls, but four are yellow and only one is blue. Bag A has more blue balls than bag B and fewer yellow balls.

If you decide correctly what type of bag this is, you will get 500sh. If you decide wrongly, you will get nothing. Each bag is equally likely.

To help you decide we are going to reveal four of the balls from the correct bag. These balls have been chosen at random from the bag so that there is only one ball left in the bag. We won't tell you the colour of the remaining ball.

Treatment 1 Part 1 Individual decision Everyone sees 4 balls

So that is the game: you see four balls and you decide whether the bag is A or B. If you choose correctly you win 500 shillings. If your choice is wrong you win nothing. You will have lots of questions like this. Each time you get a question right you will win money, so your money can add up through the experiment.

Any questions? [pause a while]

Before we start let's check that we have explained everything well.

Check question:

1. How many yellow balls are there in Bag A? How many blue balls?
2. How many yellow balls are there in Bag B?
3. If you choose the correct bag how much do you win?
4. If you choose the wrong bag how much do you win?

[**when check questions done**]
Now, in this part of the experiment we are going to play this game SIX times. Each time the bag may be different, so you should look carefully at the balls we show you. Please turn over the sheet in front of you. You will see a picture of a row of four balls**. [at this time the second enumerator should have the examples of wrong answers on the wall].**

When you have made a decision about which bag you think it is, you tick one of the bags. A or B. Make sure you do it right. If you tick A that means you think the bag is A. If you tick B that means you think the bag is B. [**point to the wrong way of answering on the wall and tell people why they are wrong]** Please think carefully before you tick. If you want to change your mind please raise your hand and an enumerator will come to help you

After you have ticked one bag, you have to raise your hand. An enumerator will come over and check whether you are right or wrong and tell you whether you have won some money. Any money will be paid at the end of the experiment. After the enumerator has checked your question you can move on to the next question.

Altogether you will play six times in this part. Any questions?

Let's play the first question. Please take your own time – it is not a competition.
[enumerator: check which is their final decision, quietly reveal the answer and tick whether they are correct. Remember to check your own copy so that you can tally the number of correct answers. Then tell them to move on to the next question. [**After question 6 has been answered by everyone, remove all answer sheets for each person from their desk. Make sure their identifying number is on all the bits of paper**]

Thank you. You have now completed Part 1 of the experiment. In the remaining parts of the experiment Bag A and Bag B will remain the same and you will have to make the same type of decisions, but each time we will change the rules slightly, so please pay attention.

Treatment 1. Part 2 Joint decision together Both see 4

[**Tidy up answer sheets from first section then get spouses to sit together. Make sure envelope for questions for Part 2 is on the table]**

In this section you have to make similar decisions to the ones you just made. The bags are the same as before. However, now you have to make a joint decision as a couple. How you make the decision is down to you as a couple, but still only one tick is allowed for each decision. If you make the correct decision you will each receive 500 shillings. If you make the wrong decision you will both get nothing. Any money you earn will be added to your personal winnings from Part 1 and paid to you at the end. By the way, it's okay to talk to your spouse in this part.

Any questions?

Now take your booklet out of the envelope. You will see questions like before. You have to look carefully at the colour of the four balls and choose which bag the balls came from. Remember, you have to agree on just one choice.

Okay please answer question 1. When you have ticked A or B, please put your pen down and raise a hand and an enumerator will come round.

[enumerators go round and mark each answer sheet as correct or not when they see a hand raised. Remember to mark your own tally sheet. When you've marked a question, turn over the page of booklet to the next question]

[at the end, after the sixth question has been answered by everyone]

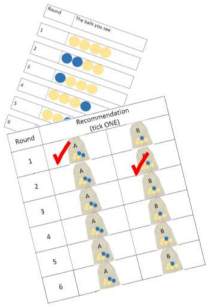
Thank you. That concludes part 2. Now for part 3 onwards we will split you up into two groups. Could all the people with RED cards go with FAITH to the other room. Please take your cards with you. People with GREEN cards, please stay here.

[if you feel refreshment is needed, this is the time when people can take a few minutes break]

Part 3 Red recommends, green decides Green sees 2, Red sees 4

[**from now onwards the instructions are the same for all treatments. Just use the word 'partner' for treatment 3, rather than 'spouse'**]

Part 3. RED Room



[**make sure everyone is seated and has their RED card. While one person puts people in numerical order, the other person makes sure the posters are on the wall or blackboard. Put the recommendation sheets on each desk.]**
[possible set up:

R1		R5		R8
R2		R6		R9
R3		R7		R10
R4				

In this section, you will be making a joint decision with your spouse

The bags are the same as before. If the correct decision is made you will each receive 500 shillings. If the wrong decision is made you will both get nothing. Any money you earn will be added to your personal winnings from Part 1 and part 2 and paid to you at the end.

Please turn over the answer sheet on your desk**.[check everyone has done this]**

Now, I am going to tell you the difference about this part. In this part, you will see 4 balls for each question. Your spouse in the other room will see two of the same balls. So you will see more balls than your spouse . And your spouse knows this. After you have studied the four balls, you have to make a recommendation to your spouse by circling the sheet. You can recommend bag A or bag B. You cannot recommend both bags and you must recommend one bag. You cannot say how many balls of each type you can see. Only A or B.

You will go through all six questions for this section and choose your recommendation for each one. How you make your decision is down to you, but remember you can only tick one recommendation per question.

After everyone has made their recommendation, the enumerators will take your recommendations to your spouse in the other room. Your spouse will make the final choice. It is their decision that will determine whether you each win 500 shillings or not.

For this part we will tell you how much you have won at the end of the experiment.

Any questions? [allow a little time]

Part 3 Red recommends, green decides Green sees 2, Red sees 4

Check questions:

1. How many balls will you see?
2. How many balls will your spouse in the other room see?
3. Who makes the final decision, you or your spouse in the other room?

Now please make your recommendation, starting with the first question

[allow some time and send enumerators through the room to check that everyone has answered all the questions. Make sure there is a tick on each question.

1. **Enumerators record the recommendations and take them to the other room**
2. **Be careful not to reveal other people's recommendations**
3. **When all six correct answers are revealed in the other room, enumerators bring back the spouse's decision and the result]**

[when the recommendation sheets are all done]

Thank you. We will now take the recommendations to the other room. Please wait a little. Please don't talk to your neighbour.

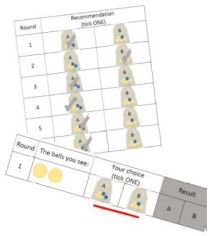
[make sure the recommendation sheets are in numerical order before you take them to the other room. This will help the other room. The one person who takes the sheets stays in the green room to help and only comes back at the end of the part]

[after the result have been returned]

Thank you. That concludes this part. We will now do part 4.

Part 3 Red recommends, green decides Green sees 2, Red sees 4

Part 3. GREEN Room



[**make sure everyone is seated and has their GREEN card. Layout this room in the same way as the room for RED. Make sure the right posters are put before you start the part.]**
In this section, you will still be making a joint decision with your spouse and the two bags are the same.

Now, I am going to tell you the differences about this part. In this part, you will see only **2 balls** for each question. Your spouse in the other room will see the **same two balls plus two more**. So in total, your spouse will see four balls. **You will see fewer balls than your spouse and your spouse knows this.**

The other difference is that your spouse will make a recommendation to you, but you will make the final decision about whether each bag is A or B. How will this work?

After your spouse has studied the four balls, they have to make a recommendation to you by ticking the sheet. They have to recommend one answer.

The enumerators will bring the recommendations from the other room. Then you will see your spouse [partner]'s recommendations and you have to make the final choice. It is your decision that will determine whether you each win 500 shillings or not.

Altogether we will do six questions like this. The enumerators will bring all the recommendations together and you will go through the six questions. We will tell your spouse how many questions you got right. However, they won't know which decision you made in each individual question.

Any questions?
Check questions:

1. How many balls will you see?
2. How many balls will your spouse in the other room see?
3. Who makes the final decision, you or your spouse in the other room?

Okay we will allow a little time for the people next door to make their recommendations. Please don't talk while we are waiting.

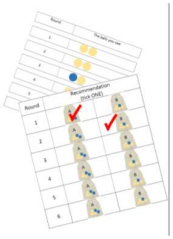
[when the recommendations come]
Please take your answer sheets out of the envelope.

Now please make your decisions. **Remember that your spouse sees 4 balls and you only see 2 of them. When you've finished the questions please put up your hand so that the enumerator can come and check.** [allow some time.]

[after the sixth question and the result, record the result and take it back to the spouse in the other room]
[When all have finished] Thank you. That concludes this part. We will now do part 4.

Part 6 Green recommends, red decides Green sees 2, Red sees 4

Part 4. RED room.



In this part, you will still be making a joint decision with your spouse and the two bags are the same.

What is the difference this time? In this part, you will see only **2** balls for each question. Your spouse in the other room will see the **same two balls plus two more**. So in total, your spouse will see four balls. You will see fewer balls than your spouse and your spouse will know this. Otherwise it's the same as the previous section. You recommend and they choose.

Any questions. [allow a little time]
Check questions (out loud, publicly – no need to record)

1. How many balls will you see?
2. How many balls will your spouse in the other room see?
3. Who makes the final decision, you or your spouse in the other room?

Now please make your recommendation, starting with the first question
[allow some time and send enumerators through the room to check that everyone has answered all the questions.

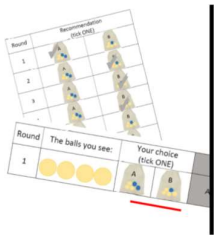
1. Enumerators record the recommendations
2. One enumerators takes them to the other room and stays there to help.
3. Be careful not to reveal other people's recommendations
4. When all six correct answers are revealed in the other room, the enumerator bring back the spouse decision and the result]

[when the recommendation sheets are all done]
Thank you. We will now take the recommendations to the other room. Please wait a little. Please don't talk to your neighbour.
[sort the sheets numerically, one enumerator stays in the other room to help]

[after the result have been returned]
Thank you. That concludes this part. We will now do part 5.

Part 6 Green recommends, red decides Green sees 2, Red sees 4

Part 4. GREEN Room



[put answer sheets on desks in the right order]
In this section, you will still be making a joint decision with your spouse and the two bags are the same.

What is the difference this time? In this part, you will see 4 balls for each question. Your spouse in the other room will see two of the same balls. So you will see more balls than your spouse [partner]. And your spouse knows this. Otherwise it's the same as the previous section. Your spouse recommends and then you choose.

Any questions?
Check questions (public, no need to record):

1. How many balls will you see?
2. How many balls will your spouse in the other room see?
3. Who makes the final decision, you or your spouse in the other room?

Okay we will allow a little time for the people next door to make their recommendations. Please don't talk while we are waiting.

[when the recommendations come]
Please take your answer sheets out of the envelope
Now please make your decisions. **Remember that you see 4 balls and your spouse only sees 2 of them. When you've finished the questions please put up your hand so that the enumerator can come and check.** [allow some time.]

[after the sixth question and the result, record the result and take it back to the spouse in the other room]
[When all have finished] Thank you. That concludes this part. We will now do part 5.

Parts 5 & 6 (omitted)