Theory Dec. (2013) 75:175–191
DOI 10.1007/s11238-012-9322-7

Individual vs. couple behavior: an experimental investigation of risk preferences

Mohammed Abdellaoui · Olivier l'Haridon · Corina Paraschiv

Published online: 15 July 2012

Abstract In this article, we elicit both individuals' and couples' preferences assuming prospect theory (PT) as a general theoretical framework for decision under risk. Our experimental method, based on certainty equivalents, allows to infer measurements of utility and probability weighting at the individual level and at the couple level. Our main results are twofold. First, risk attitude for couples is compatible with PT and incorporates deviations from expected utility similar to those found in individual decision making. Second, couples' attitudes towards risk are found to be consistent with a mix of individual attitudes, women being more influent on couples' preferences at low probability levels.

Keywords Couples · Group decisions · Risk · Uncertainty · Prospect theory · Utility · Gender

1 Introduction

Many economic and financial decisions like consumption, saving, investment, or labor supply are made at the family level and often imply the two adult members of the

M. Abdellaoui
HEC-Paris & GREGHEC-CNRS, 1 rue de la libération, 78351 Jouy en Josas Cedex, France
e-mail: abdellaoui@hec.fr

O. l'Haridon (✉)
Crem-Université de Rennes 1 & GREGHEC, Rennes, France
e-mail: lharidon@greg-hec.com

C. Paraschiv
Université Paris Descartes & GREGHEC, Jouy en Josas Cedex, France
e-mail: paraschiv@hec.fr

family household. These decisions cover a wide range of choices from basic needs to the consumption of durable goods, careers, housing, fertility, investment in children's education, retirement, and health plans. The literature on within-household decisions has been devoted mainly to riskless consumption decisions. However, most of the financial decisions made by couples involve risk. Examples of such decisions include the choice of either a common saving or bank account or a private one, choice of a monthly saving rate and choice of a retirement plan. Recent literature suggests that an understanding of such risky decisions for couples might be of great interest, especially from a behavioral perspective (Bateman and Munro 2005; Chambers and Echenique 2010; De Palma et al. 2011; Mazzocco 2004). This article aims at investigating this issue and provides an experimental test of the nature of decision making within the couple.

We complete previous studies on couple decision making in several respects. First, we interview couples' members both individually and jointly. The joint interview allows to investigate the behavioral interactions between the couple and its individual members and, in particular, women and men bargaining power in collective decision. Second, we propose a more precise comparison between individuals and couples than in previous studies because we estimate the components of prospect theory (PT). This insight allows us to compare couples and men/women not only in terms of utility but also in terms of probability weighting. We assume that the couple's members are characterized by individual preferences and that the couple acts as a specific decision maker whose preferences could be measured independently of its members' preferences.

The main results are the following: probabilistic risk attitudes of individuals and couples show similar judgement biases and are compatible with an inverse S-shaped probability weighting function. Moreover, couples' probabilistic risk attitude is found to be a mix of the partners' attitudes: it lays within the boundaries of each partners' individual attitude: couples' preferences are closer to the women's preferences at low-probability levels and closer to the men's at high-probability levels.

The remainder of the article is organized as follows. Section 2 gives a brief literature review. In Sect. 3, we review PT and the associated measurement method. The design and the experiment are described in Sect. 4. Results are presented in Sect. 5. Section 6 concludes the article.

2 Related literature

In the usual approach to consumer behavior it is often assumed that households behave as if they were single decision-making units maximizing a well-behaved utility function. It is now well established that this standard representation of preferences fails to provide a satisfactory representation of household preferences. First, a large body of empirical evidence shows that the so-called unitary model of household consumption, based on the assumption of a single set of preferences, is rejected by the data (Lundberg et al. 1997; Browning and Chiappori 1998; Cherchye et al. 2009). Second, assuming a single preference relation for the household is not consistent with the usual requirements of methodological individualism. As multi-adult households consist of several members, their behavior results from bargaining processes and should be investigated on the basis of individual preferences through a collective decision model (Chiappori 1988; Browning and Chiappori 1998). The collective approach assumes that each member of the household has specific preferences and that the household behavior results from a bargaining process leading to a Pareto efficient allocation of resources. The bargaining process attaches weights to each member of the household in the intra-household bargaining process. These weights can be interpreted as bargaining power. The utility of the household is then defined as the weighted average of the utilities of its members. In this riskless setting, the weights are endogenous and depend on individual characteristics. Kebede et al. (2011) in a large-scale experiment showed that significant amounts of potential surplus are not realized and efficiency not attain in household decision making. These results cast doubts on both the unitary and the collective model.

With few exceptions (Mazzocco 2004), the literature refers to the unitary model for decisions under risk. There might be, however, significant heterogeneity in individual attitudes towards risk within a couple and joint decisions might emerge from various bargaining processes. Using data from the *German Socio-Economic Panel*, Dohmen et al. (2011) found positive correlation of risk-attitudes within couples. On the basis of data from the *US Health and Retirement Study*, Mazzocco (2004) showed that about 50 % of couples reported that the wife's risk preferences differed from the husband's. Moreover, Mazzocco (2004) found that average savings of couples was a U-shaped function of the husband's risk-aversion. In the experimental economics literature, Bateman and Munro (2005) were the first to design an experimental test for couple decisions under risk. Their results were twofold. First, they showed that couples followed a similar pattern of violations of expected utility when compared to the usual results in individual choice experiments. Munro and Popov (2009) also found that couples exhibit an endowment effect in more or less the same way as the non-couples in a between-subject experiment. Second, Bateman and Munro showed that individual and collective expressions of value for decision under risk were different. Indeed, couples appeared to be more risk-averse when facing tasks together than when the partners faced the same decision making tasks alone. A similar result can also be found in the recent literature on group decision making (Baker et al. 2008; Shupp and Williams 2008 for high-risk situations). On the contrary, Chambers and Echenique (2010) showed that when aggregating individual welfare with certainty equivalents, a couple of risk-averse individuals behave in the aggregate in a fashion that is less risk-averse than its individual members. De Palma et al. (2011) also built an experiment that aimed to establish the relationship between couples' and individuals' behavior toward risk. They paid special attention to the dynamics of the decision process. They found that the balance of power was modified during the bargaining process and that women gained more and more power over the course of the decision-making process as time passed. Carlsson et al. (2009) measured the link between wealth and spouses' similarity in risk attitudes in an experiment performed in rural China. They observed a positive link between household wealth and similarity of partners' preference toward risk. They also found the balance of power to be in favor of men. However, when women contributed relatively more to the household income, couples' risk attitude was more likely to reflect women's preferences.

3 Theoretical background and elicitation method

To compare individuals' and couples' preferences under risk and test the difference between them, we use PT (Tversky and Kahneman 1992), currently the most widely used descriptive theory of decision under risk (Starmer 2000; Wakker 2010). In what follows, we first present PT. Then, we present the elicitation method used to estimate individuals' and couples' preferences under risk.

3.1 Decision making under risk and PT

We consider a decision maker, either an individual or a couple, who has to make a choice between two risky prospects. As we only use prospects with at most two distinct outcomes, we restrict the framework to such prospects. $(x, p; y)$ denotes a prospect that results in outcome x with probability p and in outcome y with probability $1 - p$. We assume that $x \geqslant y \geqslant 0$. Outcomes are monetary amounts and higher numbers are always preferred. If $x = y$, the prospect is said to be riskless, otherwise it is said to be risky. Under PT, preferences over prospects are represented by a real-value utility function, $u(\cdot)$, defined over monetary outcomes as changes with respect to a reference point, and by a probability weighting function $w(\cdot)$. The utility function is strictly increasing from $\mathbb{R}$ to $\mathbb{R}$ and satisfies $u(0) = 0$. As we consider only gains in this article, then PT corresponds to rank-dependent utility (Quiggin 1982). Also, we take 0 as the reference point.

The decision maker evaluates each prospect separately and chooses the prospect that offers the highest value. The PT valuation of prospect $(x, p; y)$ is given by

$$w(p)u(x) + (1 - w(p))u(y) \tag{1}$$

Function w is strictly increasing from [0, 1] to [0, 1] and satisfies $w(0) = 0$ and $w(1) = 1$. In Eq. 1, $w(p)$ and $1 - w(p)$ are the decision weights attached to the larger and lower outcomes, respectively. Kahneman and Tversky (1979) assumed that the probability weighting function $w(\cdot)$ overweights small probabilities and underweights moderate and high probabilities, giving rise to an inverse S-shaped function. This assumption has been confirmed by the literature (Tversky and Kahneman 1992; Tversky and Fox 1995; Wu and Gonzalez 1996). Under expected utility, these decision weights are equal to p and $1 - p$ and the probability weighting function—if any—corresponds to the identity function. Tversky and Kahneman (1992) assumed that the utility function was concave for gains, which is confirmed by the empirical literature (Tversky and Kahneman 1992; Gonzalez and Wu 1996; Abdellaoui et al. 2007; Booij and van de Kuilen 2009).

3.2 Elicitation method

Our elicitation method of utility under risk is based on Abdellaoui et al. (2008) and consists of two stages. In the first stage, utility is elicited, then, in the second stage, we focus on the elicitation of probability weights. We start by selecting a probability p^*

and a series of k pairs of outcomes $\{(x_i, y_i) : i = 1, \ldots, k\}$ that are kept fixed throughout the elicitation of the utility function in the gain domain. Next, we elicit k certainty equivalents $z_1, \ldots, z_k$ for a series of risky prospects $(x_i, p^*; y_i)$ with $i = 1, \ldots, k$. The advantage of keeping the probability p^* fixed is that only one additional parameter, $\tau = w(p^*)$, has to be estimated besides the parameter(s) of the utility function. If we adopt a parametric specification for utility, the series of certainty equivalents allows us to estimate τ and the utility parameter under PT through nonlinear least squares $\|z - \hat{z}\|^2$ with

$$\hat{z}_i = u^{-1}[\tau \cdot (u(x_i) - u(y_i)) + u(y_i)] \tag{2}$$

$i = 1, \ldots, k$. Once utility has been determined in the first stage, the decision weights can be elicited in the second stage. To do so, we select an outcome x^* and a series of m probability levels $\{p_j : j = 1, \ldots, m\}$ that are kept fixed throughout the elicitation process. Then m certainty equivalents $z'_1, \ldots, z'_m$ are elicited for a series of risky prospects $(x^*, p_j; 0)$, $j = 1, \ldots, m$. Under PT and according to Eq. (1), the decision weights associated with a given probability p_j can be computed as:

$$w(p_j) = \frac{u(z'_j)}{u(x^*)} \tag{3}$$

$j = 1, \ldots, m$.

4 Experiment

4.1 Subjects

Subjects were people living in the city of Paris, France. 130 subjects took part in the experiment (65 couples). Each couple was paid €50 for its participation. In addition, we implemented a between-subject random-lottery incentive scheme. Before starting the experiment, participants (individuals and/or couples) were informed that they could be selected to play one of their choices for real and could win a maximum of €1,200 depending on their choices. Together with gender, we also collected information on age, number of children, and the length of the relationship. Couples were recruited through advertisements made in a number of public places in Paris: schools, associations, day-care centers, and social events. A couple was selected to participate in the study only if each of the partners was over the age of 25 and the couple had lived at least 1 year together. The minimum age condition was used to exclude from the study young student couples that were not yet financially independent from their parents. Moreover, 1 year of life in common was requested to insure that the two people had already had the opportunity to take financial decisions together.

The experiment was conducted in the form of computer-based individual interviews. A special software had been developed for the purpose of the experiment. Subjects were told that there were no right or wrong answers, and were allowed to take a break at any time during the session. The responses were systematically entered in the computer by the interviewer so that the subjects could focus on the choice questions. We always carried out the individual interviews before the couple interview. This design was meant to minimize the potential impact of couples' answers on individuals' answers. For individual interviews, the gender order (female/male or male/female) was random. The structure of the individual and joint session was the same. In the latter, couples were allowed to freely communicate and no time constraint was imposed to the decision-making process. This allowed couples to bargain over the decisions. According to the collective view of the household, the potential gains were proposed to the couple without any predetermined sharing rule. The experiment was part of a larger experiment that lasted on average one hour and a half, including 5–10 min for task explanation and practice questions.[1] The average duration of the experiment for women (15.4 min) was no significantly higher than the duration for the men (12.6 min) but was higher that duration for the couples (11.2 min). The average duration for the couples' may differ from the individuals' for one of two reasons. First, an adaptation effect may have arisen because each of the individuals in the couple had already provided a personal answer to the tasks they were subsequently asked as a couple. This effect might have resulted in reducing the duration of the experiment. Second, a bargaining effect might also exist as couple decisions which are collective decisions with an extensive use of communication (Ashraf 2009). This effect might have increased the duration of the experiment. The lower average duration observed for couples showed that the adaptation effect was higher than the bargaining effect for women only. This is partly consistent with Blinder and Morgan (2005) result which shows that groups are just as quick as individuals to reach decisions.

We discarded one couple from the experiment because the husband did not understand the tasks and gave incomplete answers. This left 128 subjects (64 couples) for the analysis.

4.2 Experimental methods

We used 11 certainty equivalence questions to elicit the utility function and the decision weights. Details are provided in the Appendix. All indifferences were elicited through a three-step iterative choice list procedure. In the first step, subjects had the choice between a fixed prospect and a variable amount of money. The latter was framed as linearly equally spaced outcomes between the minimum and the maximum amounts given in the fixed prospects. The second and third steps refined the choice at the point where subjects had switched in the previous list. Both sides of the switching point served as a bound for the next choice list. We divided the range into 11 categories for each of the steps.

To ensure consistency and incentive compatibility and to control for response error, we added a fourth step to the choice list procedure (Abdellaoui et al. 2011 for a

[1] The second part of the larger experiment consisted in decisions over time whose results are not reported here.

Table 1 Classification of subjects in terms of risk attitude

	Risk-averse	Risk-seeking	Mixed
Women	27	17	20
Men	25	22	17
Couples	23	24	17

similar procedure). This fourth step corresponded to the entire choice list that would have been generated by refining every possible switching point from the first list. Assuming monotonicity, the computer pre-filled the list based on the answers given during the previous steps. Then the list was presented to the subject for validation. The software allowed backtracking if participants did not wish to validate their previous series of choices. Figure 3 in the Appendix gives an example of the way in which the experimental questions were displayed.

The order in which the six prospects used to elicit utility were presented was randomized. However, we learned from the pilot sessions that the subjects found it easier to deal with an increasing order in probabilities (from $p = 0.05$ to $p = 0.95$) than with a randomized order. The order in which the five prospects used to elicit decision weights were presented was deterministic and always came after the questions used to determine utility.

5 Results

5.1 Risk attitudes

Table 6 in the Appendix shows the median responses to all the certainty equivalence questions that we asked during the experiment. Overall, 53 % of the choices made by women, 49 % of the choices made by men, and 48 % of the choices made by couples were consistent with risk-aversion. Women were more risk-averse than men and couples at $p = 0.019$.[2] De Palma et al. (2011) found a similar result. In contrast, we found no significant difference between men and couples risk-aversion. This contradicts Bateman and Munro (2005) results for decision making in couple but agrees with the findings by Budescum et al. (2010) in group decision making. To obtain a more precise picture of risk attitudes, we constructed an individual classification of subjects, reported in Table 1. We classified a subject as risk-averse (risk-seeking) if at least two-thirds of certainty equivalents were smaller (higher) than the expected value of the risky prospect. If the participant—individual or couple—was neither classified as risk-averse nor risk-seeking, then he/she/it was classified as mixed.

The proportion of risk-averse women (42.2 %) was higher than the proportion of risk-seeking women ($p = 0.03$). Although a majority of men were classified as risk-averse (39 %), this proportion was not significantly higher than the 34.4 % of risk-seeking men ($p = 0.29$). Among couples, 35.9 % were classified as risk-averse and 37.4 %

[2] The binomial test was used to test for differences between proportions.

as risk-seeking. The difference in proportion was not significant here. Risk-averse couples were composed mainly of risk-averse individuals (9 out of 23) and risk-seeking couples were composed mainly of risk-seeking individuals (8 out of 25). Couples classified as mixed were composed mainly of risk-averse women and risk-seeking men (5 out of 17).

To better account for the relation between couples' and individuals' risk attitudes, we computed individual relative risk premia (risk premium for prospect i is defined as $(\mathrm{ev}_i - z_i)/(\mathrm{ev}_i)$, where ev_i denotes the expected value). Risk premia allowed us to measure the strength of risk-aversion (risk-seeking): a higher positive (negative) value of the risk premium indicates a higher level of risk-aversion (risk-seeking). Correlations between men and women risk premia were low and not significant (mean correlation coefficient: $\bar{\rho} = 0.10$, absence of significance at 5 % for 10 correlation coefficients over 11). We found neither positive assortative matching: risk-averse (seeking) women are not specifically matched with risk-seeking (averse) men nor negative assortative matching: risk-averse (seeking) women are not specifically matched with risk-averse (seeking) men.[3] By contrast, correlations of risk attitudes between both men and couple and women and couple were high ($\bar{\rho} = 0.415$ for women, $\bar{\rho} = 0.56$ for men, all single correlation coefficients significant at 1 %, except one). Men risk attitudes appeared to be more correlated with couples' risk attitude than women risk attitudes. Eliciting PT allows a more structured view of the composition of risk attitudes within the couple: results are given in the next subsection.

Consistent with previous experimental findings in individual decision making under risk (Starmer 2000 for review), risk-aversion varied systematically with the probability used to elicit the certainty equivalents. We found risk-seeking over low-probability gains, and risk-aversion over high-probability gains. These behaviors are characteristic of the fourfold pattern of risk attitude, which was, according to its authors, one of the most distinctive implications of PT (Tversky and Kahneman 1992, p. 306). Figure 1 shows the percentage of risk-averse participants as a function of the probability used in the elicitation.

Risk-seeking was clearly dominant for probability $p = 0.05$ and risk-aversion was dominant for probabilities $p = 0.75$ and $p = 0.95$. For intermediary probabilities $p = 0.25$ and $p = 0.50$, evidence was less clear cut. For probability $p = 0.25$, we found risk-seeking for couples ($p = 0.02$) but neither for men ($p = 0.21$) nor for women ($p = 0.13$). For probability $p = 0.50$, we found no evidence for risk-aversion.

5.2 Prospect theory

5.2.1 Utility

We used a power parametric specification for utility: $u(x) = x^\alpha$. For gains, the power function is concave if $\alpha < 1$, linear if $\alpha = 1$, and convex if $\alpha > 1$. Table 2 reports the utility estimates for women, men, and couples, together with the number of individuals with a concave or a convex utility function.

[3] Carlsson et al. (2009) found similar results in a rather different context of choice under risk.

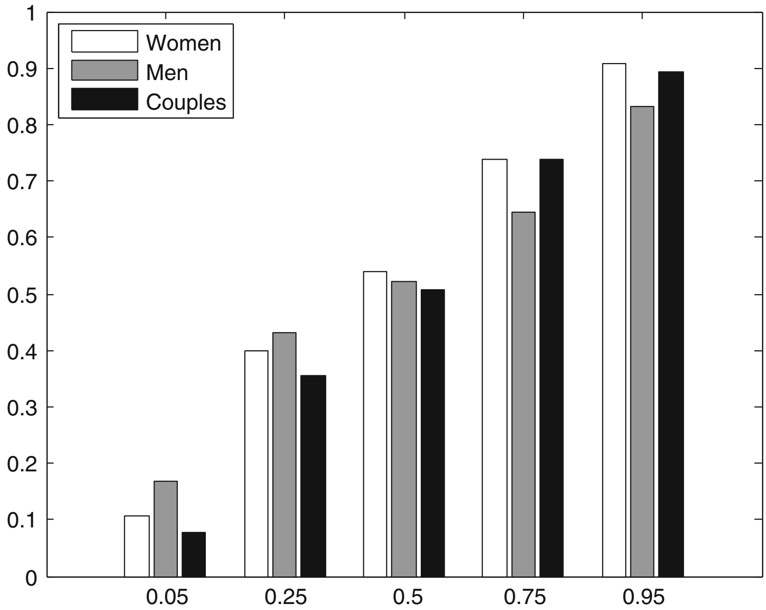


Fig. 1 Risk-aversion (% of risk-averse subjects) as a function of probabilities

Table 2 Median estimates for utility under risk

	Power (α)		
	Women	Men	Couples
Median	0.88	0.81	0.96
IQR	0.74–1.36	0.62–1.05	0.83–1.17
$\#(\alpha \leqslant 1)$	39	46	36
$\#(\alpha > 1)$	25	18	28

IQR interquartile range

For couples, utility was linear: the median power coefficient of 0.96 did not significantly differ from 1 [$t(63) = 0.33$, $p = 0.74$]. For men, utility was concave: the median power coefficient was equal to 0.81 (Wilcoxon, $p < 0.01$, $t(63) = 2.11$, $p = 0.04$). For women, although the median power coefficient was lower than 1, concavity was not significant (Wilcoxon, $p = 0.94$, $t(63) = 0.84 p = 0.40$). The large spread of elicited utility parameters among women explains this result. On the other hand, we found more concave utility functions than convex utility functions for both women and men ($p < 0.01$). For couples, we could not reject equality of counts of concave and convex utility functions ($p = 0.16$).

We found high correlations both between couples' utility parameters and men's utility parameters ($\rho = 0.68$, significant at $p < 0.01$) and between couple's utility parameters and women's utility parameters ($\rho = 0.65$, marginally significant at $p = 0.09$). To evaluate the balance of power within the couple, we regressed couples' utility parameter on individual's utility parameters, assuming the linear constraint that individual weights sum to one. The linear constraint has two complementary

Table 3 Women's bargaining weights in couple decisions

Model	1	2
Women's weight	0.401 (0.136)	0.404 (0.143)
Length of the relationship	−0.003 (0.006)	0.003 (0.004)
Children	−0.469** (0.174)	0.037 (0.131)
2 Children	0.169* (0.067)	−0.028 (0.039)
Age gap	−0.001 (0.011)	0.043 (0.014)
Observations	64	64×5
R^2	0.240	0.389

Standard errors in parentheses (* significant at 5 %; ** significant at 1 %). For women's bargaining weights, significance is measured towards equal weighting (one-half). Dependent variables are utility (model 1) and decision weights (model 2). Model 2 includes clustering for multiple answers by the same individual

interpretations. First, if the weights lie between 0 and 1, they can be interpreted as the balance of power in the couple. Second, if weights are outside the [0, 1] range, this means that the couples' utility are not well determined by a balance of power between the individuals. A potential drawback of the regression procedure is the potentially ill-defined nature of individual's weights in the decision-making process. Indeed, the regression is compatible with a collective model of the household but also with a dictatorship of one altruistic individual putting some weight on the preference of the other. To that respect, a better definition of the individual weights could be obtained by undertaking within-couple heterogeneity in individual characteristics. In what follows, we assume that individual weights depend on the partners' age gap, the length of their relationship, and the number of children, if any. Table 3 shows the results of the OLS regression in column 1 (model 1). The base weight associated with women's power over the couple decision equalled 0.404 (not significantly different from equal weighting). We found children to have a nonlinear effect on women weights: while the presence of one or two children decreased women's bargaining weights, having more than two children significantly increased women balance of power within the couple. We found no significant effect of neither the age gap nor the length or the relationship on the balance of power for utility.

5.2.2 Probability weighting

The second component of behavior toward risk under PT is probabilistic risk attitude, which is captured through probability weighting (Wu and Gonzalez 1996 for discussion). Table 4 gives the median decision weights along with the corresponding interquartile ranges. The median results are consistent with an inverse S-shaped probability weighting function for women, men, and couples: overweighting of small probabilities, underweighting of large probabilities. These inverse S-shaped functions have, however, different shapes for women, men, and couples: men were generally more optimistic than both their partner and their couple: men overestimated more small probabilities and underestimated less large probabilities.

Table 4 Median values for decision weights

p	Women			Men			Couples		
	Median	IQR	$t(63)$	Median	IQR	$t(63)$	Median	IQR	$t(63)$
0.05	0.16	0.05–0.23	7.44**	0.16	0.09–0.29	8.09**	0.17	0.10–0.243	10.79**
0.25	0.27	0.16–0.36	1.24^{ns}	0.32	0.23–0.43	4.02**	0.28	0.22–0.39	3.27**
0.50	0.45	0.32–0.54	−2.97**	0.53	0.38–0.65	0.37^{ns}	0.46	0.40–0.57	−1.27^{ns}
0.75	0.61	0.52–0.71	−6.93**	0.70	0.58–0.81	−3.44**	0.66	0.55–0.74	−5.79**
0.95	0.81	0.68–0.87	−9.46**	0.86	0.74–0.93	−6.33**	0.84	0.74–0.90	−8.69**

ns not significant, *IQR* interquartile range
$t(63)$: two-tailed t test, H_0: $w(p) = p$ * significant at 5%; ** significant at 1%

Table 4 reports t tests supporting the view that probability 0.05 was overweighted and that probabilities 0.75 and 0.95 were underweighted for women, men, and couples. For women only, Table 4 also shows that probability 0.25 was not transformed and that probability 0.5 was underweighted. As regards men and couples, probability 0.5 was not transformed and probability 0.25 was overweighted. Paired t tests for binary comparisons (Table 7 in the Appendix) show that both men and women's probabilistic risk attitudes differed at all probability levels, except 0.95. This suggests that partners appeared to be similarly sensitive to the certainty effect (i.e., the strong impact of the shift from probability 0.95 to certainty), but not to the possibility effect (i.e., the strong impact of the shift from impossibility to probability 0.05). For the latter, men exhibited a higher overestimation of probabilities. Finally, partners did not behave similarly towards intermediate probabilities. A pooled regression for all probability levels, clustering for multiple responses, shows that women had a weight of 0.40 on decisions and men had a weight of 0.60 (Table 3, model 2). We found no significant effect of neither children, the age gap nor the length of the relationship on the balance of power for probabilistic risk attitude. Binary comparisons of probabilistic risk attitudes (Table 7 in the Appendix) show that couples' attitudes are significantly closer to men's risk attitudes at large probability levels and significantly closer to women's risk attitudes at low probability levels. Figure 2 presents elicited probability weighting functions based on the median responses for women, men, and couples. We used Prelec (1998) two-parameter compound invariance specification of the weighting function: $w(p) = \exp(-\delta(-\ln p)^{\gamma})$, where δ controls for optimism/pessimism and γ controls for insensitivity to changes in probabilities. Figure 2, plotted on median data, shows that probability weighting for couples lies within the boundaries of individuals' probability weighting. The only difference in parameters we found was between women's and couples' insensitivity to probability, with women being less sensitive [$t(63) = 1.85$, $p = 0.03$, one-tailed paired test].

6 Conclusion

The first finding of this article is that correlations between risk attitudes within couples are rather weak. We observed no significant relation between men and women's

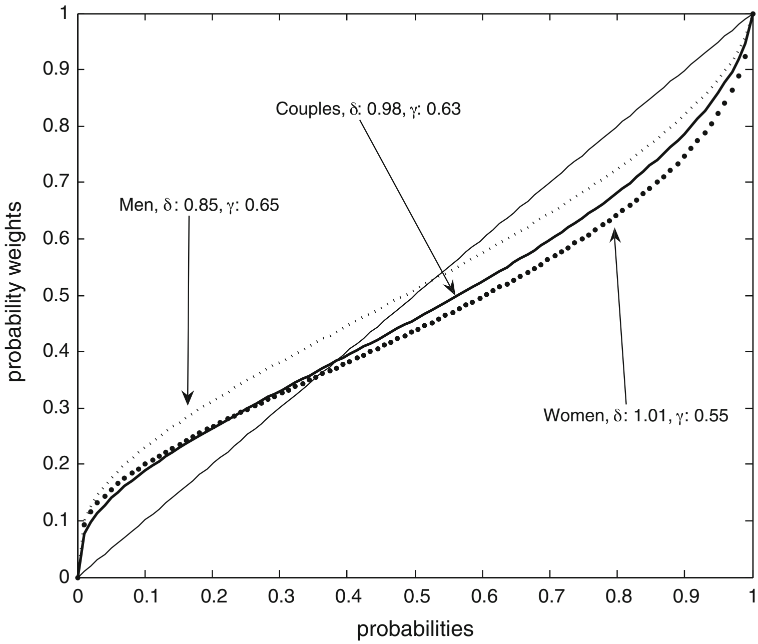


Fig. 2 Probability weighting functions, median parameter values

risk-aversion. Hence, we found no support for assortative matching within the couple. Our results give support neither to Becker's theory of marriage which, assuming one-dimensional types of the partners, complementary in producing surplus, induces assortative matching (Becker 1973) nor to Chiappori and Reny (2006) risk-sharing model, where the more risk-averse men are matched with the less risk-averse women. Gierlinger and Laczó (2011) show that the existence of limits to commitment can produce positive assortative matching in the risk-sharing model. Following these steps, the unobserved heterogeneity in partners' limits to commitment in risk-sharing is a plausible explanation for the large heterogeneity in assortative matching observed in our experiment.

Second, our study shows that the probability weighting function for couples is inverse S-shaped. This result shows that this feature, standard in individual behavior, can be extended to couple behavior. Observing women's behavior seems to provide a better predictor of couple decisions at low-probability levels, while observing men's behavior provides the best predictor of couple decisions at high-probability levels. We also found that men were more optimistic than their partner and more optimistic than their couple in choice under risk. Such gender effect in optimism is consistent with existing experimental evidence on gender preferences (Croson and Gneezy 2009) as well as with the findings on overconfidence in financial decision making (Dellavigna 2009). Gender effects have also been observed in risky decisions by Fehr-Duda et al. (2006) who found that women underestimate large probabilities more than men do. An explanation could be that men are more overconfident than their partner about their relative performance in individual risky tasks (Niederle and Vesterlund 2007) but that decisions in couples tend to correct.[4] Contrary to recent articles (Bruhin et al. 2011; Harrison and Rutström 2009), we did not find that women deviate more strongly from linear probability weighting than do men.

Third, correlations between men and women's overconfidence were close to zero in both cases. We found, however, that the couples' risk attitude was closer to the men's risk attitude at high-probability levels and closer to the women's risk attitude at low-probability levels. One interpretation of this result is that joint decision counterbalances probability weighting: men's overconfidence at low-probability levels and women's underconfidence at high-probability levels are both lowered in couple decisions. Accordingly, decision making in couples seems to reduce the distance between decision weights and objective probabilities. However, we did observe an inverse S-shaped probability weighting function for couples. This suggests that decisions resulting from the couple deliberative process do not completely eliminate the usual discrepancies between the descriptive model (PT) and the normative model (expected utility). Empirical evidence regarding comparisons between individual and collective decision making under risk supports this view. First, Sutter (2007) shows that loss aversion affects group decision making, even though it is lower than individual loss aversion. Second, Charness et al. (2007) find that deviations from the courses of action prescribed by normative models of decision making under risk, namely violations of first-order stochastic dominance in their experiment, are lower in groups but do not disappear, except when the group increases in size.

Of course, our experiment has several limits. First, the stakes used in the experiment (up to €1, 200) are not high enough to capture some of the most important financial decisions made by couples. Second, we used only positive amounts of money in our analysis of couples' decision making. For a more complete understanding of couple decisions, it would be necessary to perform a distinct analysis of couples' behavior for losses, as well as for mixed prospects involving both gains and losses. This would provide a useful basis for the comparison of individuals' and couples' degrees of loss aversion. As the current experiment was already relatively long, we felt that a longer session including losses and mixed prospects would have been detrimental to the quality of the experimental data. Third, our analysis was restricted to couples' financial decisions involving monetary outcomes. It would be interesting to integrate our findings in a more general context taking into account the division of roles within the couple and the nature of altruism within the family.

7 Appendix

7.1 Stimuli

We used six certainty equivalence questions to elicit the utility function for gains. The prospects for which we determined the certainty equivalents are displayed

[4] Similarly, couples' decision would tend to correct women's underconfidence in performance at high-probability levels if any.

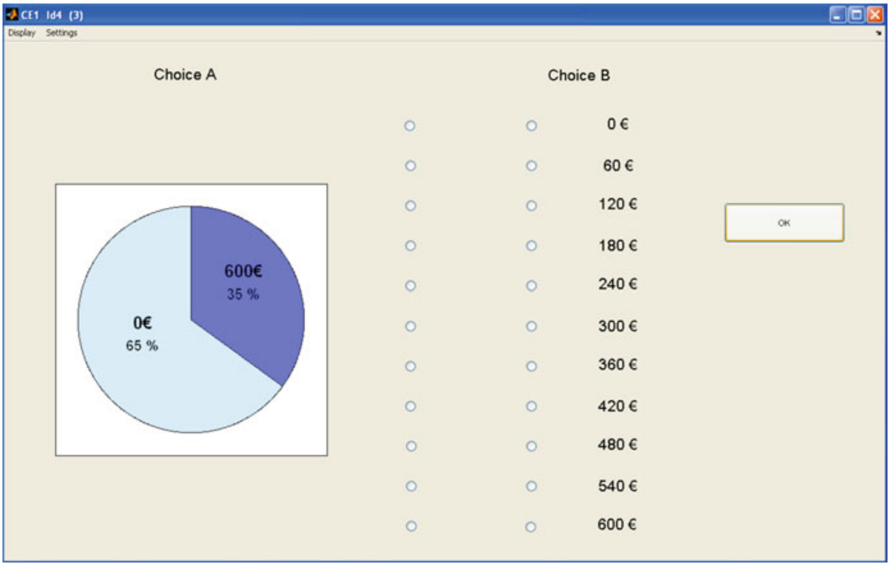


Fig. 3 A typical display used in the experiment

Table 5 Risky prospects used to elicit utility and weighting functions

	Index i										
	Utility						Decision weights				
	1	2	3	4	5	6	7	8	9	10	11
x_i	600	900	900	1,200	900	1,200	1,200	1,200	1,200	1,200	1,200
p_i	0.35	0.35	0.35	0.35	0.35	0.35	0.05	0.25	0.50	0.75	0.95
y_i	0	0	300	300	600	600	0	0	0	0	0

in Table 5 ($i = 1, \ldots, 6$). The measurements were not chained and not vulnerable to error propagation. The probability p^* used to elicit utility was $p^* = 0.35$.

For the elicitation of decision weights, we used five certainty equivalence questions. We elicited decision weights for $p = 0.05, 0.25, 0.50, 0.75, 0.95$ using $x^* = 1200$ (Table 5, $i = 7, \ldots, 11$).

7.2 Certainty equivalents

See Tables 6, 7.

Table 6 Certainty equivalents for utility and decision weights

Prospect index *i*	EV	Women		Men		Couple	
		Median	IQR	Median	IQR	Median	IQR
1	210	206	150–281	201	150–300	221	180–300
2	315	327	250–400	300	200–400	310	278–400
3	510	501	450–572	506	450–600	501	480–600
4	615	600	500–680	620	550–750	600	550–700
5	705	721	701–750	750	701–775	750	701–776
6	810	801	765–900	851	801–900	845	801–900
7	60	150	101–251	104	71–275	200	101–300
8	300	300	240–401	300	200–445	322	240–465
9	600	575	401–655	563	401–750	581	480–655
10	900	750	600–900	800	661–900	800	655–880
11	1,140	1,001	802–1,001	1,001	873–1,100	1,001	900–1,080

EV expected value, *IQR* interquartile range

Table 7 Binary comparisons of decision weights

Probability	Women vs. men	Women vs. couples	Men vs. couples
p	$t(63)$	$t(63)$	$t(63)$
0.05	-1.93^{*}	-0.45^{ns}	2.03^{*}
0.25	-2.22^{*}	-1.09^{ns}	1.87^{*}
0.50	-2.20^{*}	-1.66^{ns}	1.24^{ns}
0.75	-2.27^{*}	-1.86^{*}	1.12^{ns}
0.95	-1.53^{ns}	-2.00^{*}	-0.28^{ns}

ns not significant
$t(63)$: Student's t, one-tailed paired test * significant at 5%; ** significant at 1%

References

Abdellaoui, M., Bleichrodt, H., & Paraschiv, C. (2007). Loss aversion under prospect theory: A parameter-free measurement. *Management Science, 53*(10), 1659–1674.

Abdellaoui, M., Bleichrodt, H., & l'Haridon, O. (2008). A tractable method to measure utility and loss aversion under prospect theory. *Journal of Risk and Uncertainty, 36*(3), 245–266.

Abdellaoui, M., Baillon, A., Placido, L., & Wakker P. (2011). The rich domain of uncertainty. *American Economic Review, 101*, 695–723.

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

Baker, R., Laury, S., & Williams, A. (2008). Comparing small-group and individual behavior in lottery-choice experiments. *Southern Economic Journal, 75*(2), 367–382.

Bateman, I., & Munro, A. (2005). An experiment on risky choice amongst households. *Economic Journal, 115*(502), 176–189.

Becker, G. (1973). A theory of marriage: Part I. *Journal of Political Economy, 81*(4), 46–813.

Blinder, A., & Morgan, J. (2005). Are two heads better than one? Monetary policy by committee. *Journal of Money, Credit and Banking, 37*(5), 789–811.

Booij, A., & van de Kuilen, G. (2009). A parameter free analysis of the utility of money for the general population under prospect theory. *Journal of Economic Psychology, 30*, 651–666.

Browning, M., & Chiappori, P. A. (1998). Efficient intra-household allocations: A general characterization and empirical tests. *Econometrica, 66*(6), 1241–1278.

Bruhin, A., Fehr-Duda, H., & Epper, T. (2011). Risk and rationality: Uncovering heterogeneity in probability distortion. *Econometrica, 78*, 1375–1412.
Budescum, D., Diecidue, E., & Keck, S. (2010). *Ambiguity in group decision making*. Working Paper. INSEAD.
Carlsson, F., Martinsson, P., Qin, P., Matthias, S. (2009). *Household decision making and the influence of spouses? Income, education, and communist party membership: A field experiment in rural china*. Working Papers 2009-09. University of Innsbruck.
Chambers, C., & Echenique F. (2010). *When does aggregation reduce risk aversion?* Working Paper. California Institute of Technology.
Charness, G., Karni, E., & Levin, D. (2011). 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*(2), 129–148.
Cherchye, L., De Rock, B., & Vermeulen, F. (2009). Opening the black box of intrahousehold decision making: Theory and nonparametric empirical tests of general collective consumption models. *Journal of Political Economy, 117*(6), 1074–1104.
Chiappori, P.-A. (1988). Nash-bargained households decisions: A comment. *International Economic Review, 29*(4), 96–791.
Chiappori, P.-A., Reny, P. (2006). *Matching to share risk*. Working Paper. University of Chicago.
Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature, 47*(2), 74–448.
De Palma, A., Picard, N., & Ziegelmeyer, A. (2011). Individual and couple decision behavior under risk: Evidence on the dynamics of power balance. *Theory and Decision, 70*, 45–64.
Dellavigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic Literature, 47*, 315–372.
Dohmen, T., Falk, A., Huffman, D., & Sunde U. (2011). The intergenerational transmission of risk and trust attitudes. *Review of Economic Studies* doi:10.1093/restud/rdr027.
Fehr-Duda, H., de Gennaro, M., & Schubert, R. (2006). Gender, financial risk, and probability weights. *Theory and Decision, 60*(2), 283–313.
Gierlinger, J., & Laczó, S. (2011.) Matching and self-enforcing insurance. mimeo.
Gonzalez, R., & Wu, G. (1996). On the shape of the probability weighting function. *Cognitive Psychology, 38*, 129–166.
Harrison, G., & Rutström, E. (2009). Expected utility theory and prospect theory: One wedding and a decent funeral. *Experimental Economics, 12*(2), 133–158.
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263–291.
Kebede, B., Tarazona, M., Munro, A., & Verschoor, A. (2011). *Intra-household efficiency: An experimental study from ethiopia*. Technical Report.
Lundberg, S., Pollak, R., & Wales, T. (1997). Do husbands and wives pool their resources? Evidence from UK child benefit. *The Journal of Human Resources, 32*(3), 463–480.
Mazzocco, M. (2004). Saving, risk sharing, and preferences for risk. *American Economic Review, 94*(4), 1169–1182.
Munro, A., & Popov, D. (June 2009). *A missing link in behavioural economics? A portmanteau experiment on the relevance of individual decision anomalies for households*. Royal Holloway, University of London: Discussion Papers in Economics 09/10.
Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics, 122*(3), 1067–1101.
Prelec, D. (1998). The probability weighting function. *Econometrica, 66*, 497–528.
Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization, 3*(4), 323–343.
Shupp, R., & Williams, A. (2008). Risk preference differentials of small groups and individuals. *Economic Journal, 118*(525), 258–283.
Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature, 28*, 332–382.
Sutter, M. (2007). Are teams prone to myopic loss aversion? An experimental study on individual versus team investment behavior. *Economics Letters, 97*(2), 128–132.
Tversky, A., & Fox, C. (1995). Weighting risk and uncertainty. *Psychological Review, 102*, 269–283.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*, 297–323.

Wakker, P. (2010). *Prospect theory: For risk and ambiguity*. Cambridge: Cambridge University Press.

Wu, G., & Gonzalez, R. (1996). Curvature of the probability weighting function. *Management Science, 42*, 1676–1690.

Journal of Public Economics 129 (2015) 1–13

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

How individual preferences are aggregated in groups: An experimental study☆

Attila Ambrus [a,*], Ben Greiner [b], Parag A. Pathak [c,d]

[a] Duke University, Department of Economics, Durham, NC 27708, United States
[b] University of New South Wales, School of Economics, Sydney, NSW 2052, Australia
[c] MIT, Department of Economics, Cambridge, MA 02139, United States
[d] NBER, United States

ARTICLE INFO

Article history:
Received 17 January 2015
Received in revised form 20 May 2015
Accepted 26 May 2015
Available online 6 June 2015

JEL classification:
C72
C92
H41

Keywords:
Group decision-making
Role of deliberation
Social influence

ABSTRACT

This paper experimentally investigates how individual preferences, through unrestricted deliberation, are aggregated into a group decision in two contexts: reciprocating gifts and choosing between lotteries. In both contexts, we find that median group members have a significant impact on the group decision, but the median is not the only influential group member. Non-median members closer to the median tend to have more influence than other members. By investigating the same individual's influence in different groups, we find evidence for relative position in the group having a direct effect on influence. These results are consistent with predictions from spatial models of dynamic bargaining, for members with intermediate levels of patience. We also find that group deliberation involves bargaining and compromise as well as persuasion: preferences tend to shift towards the choice of the individual's previous group, especially for those with extreme individual preferences.

1. Introduction

Many important decisions, in various contexts, are made by groups, such as committees, governing bodies, juries, business partners, teams, and families. Group decisions are typically preceded by deliberation among members, who enter the process with varying opinions and preferences. The expansion of democratic institutions and rapid progress in communication technology further highlight the prevalence of group decisions – in politics and business, among other facets of society – and the importance of investigating the process of such decisions (see the related discussion in Charness and Sutter, 2012).

This paper presents an experimental investigation of group decision-making in two settings that are stylized versions of important real-world decision problems: (i) choosing how much to reciprocate as the second mover in a sequential gift-exchange game (Brandts and Charness, 2004; Fehr et al., 1993), and (ii) choosing between (comparatively) safe and risky lotteries, using a version of the risk-preference elicitation questionnaire of Holt and Laury (2002). Gift-exchange games are often used as a stylized framework for employment relationships with incomplete labor contracts, in which the employee performance is not always enforceable (for example, see Brown et al., 2004; Charness, 2004; Charness et al., 2012; Fehr and Gächter, 1998; Fehr et al., 1993), while the lottery choice can be considered a simplified version of financial portfolio or investment decisions. For both of the tasks above, there is no clear normative criterion for evaluating the quality of decisions.[1] In the gift-exchange game, a group's chosen reciprocation level (conditional on the first-mover's gift) should depend on members' social preferences, while lottery choices should depend on members' risk preferences. Hence, in our experiments the main focus is how different preferences shape the group decision, through bargaining and/or persuasion.

☆ We thank the Warburg foundation and the UNSW Business School for financial support and Niels Joaquin, Peter Landry, and Cara Nickolaus for valuable research assistance. Eric Budish, Georgy Egorov, Lars Ehlers, and Mihai Manea provided assistance in some of the experimental sessions. Gary Charness, Ignacio Esponda, Denzil Fiebig, Drew Fudenberg, Stephanie Heger, Stephen Leider, Muriel Niederle, Patrick Schneider, Georg Weizsäcker and seminar participants at the IAS in Princeton, George Mason University, University of Technology Sydney, the 2013 ESA conference in Santa Cruz, the 2014 APESA conference in Auckland, and the 2014 Design and Behavior workshop in Dallas provided helpful comments.
* Corresponding author.
E-mail addresses: aa231@duke.edu (A. Ambrus), bgreiner@unsw.edu.au (B. Greiner), ppathak@mit.edu (P.A. Pathak).

[1] Such tasks are dubbed " non-intellective" by Laughlin (1980) and Laughlin and Ellis (1986). For recent experimental investigations of group decision-making with intellective tasks, see Blinder and Morgan (2005) and Cooper and Kagel (2005) and Kocher and Sutter (2005). Glaeser and Sunstein (2009) provide a related theoretical analysis.

http://dx.doi.org/10.1016/j.jpubeco.2015.05.008
0047-2727/

Experimental investigation of group decisions has long been a central research area in social psychology, and has recently attracted more attention in experimental economics.[2] A novel feature of our design is that before deliberation, we solicited each member's opinion on what she thought the group's choice should be. It was randomly determined whether the eventual group choice or one of the initial individual opinions was implemented, making the solicited initial opinions payoff-relevant. In either case, the implemented outcome applied to all members with respect to payoffs. Hence, the solicited opinions can be interpreted as the outcome for the group that the individual would have chosen before deliberation, as a dictator. Another distinguishing aspect of our experiment is that groups consist of five individuals, unlike most existing studies, which investigate three-person groups.[3] Five-person groups allow us to compare the influence of the extreme group members to the non-median members who are not at the extremes.[4]

Our central empirical investigation regresses the group decision on the ordered individual decisions by the group members.[5] This regression provides a detailed picture of how a member's influence on the group decision depends on her relative position within the group. In contrast, most of the existing literature focuses on comparing aggregate statistics of group and individual decisions.[6]

Conceptually, our empirical methodology is motivated by the influential work of Davis (1973), who defines social decision schemes as mappings between individual preferences and the group decision.[7]

To provide a more formal conceptual framework, we consider two types of dynamic spatial bargaining models considered in the literature. Both approaches feature multi-period games, such that in each period members consider and vote on a proposed action. The first approach, by Banks and Duggan (2000), assumes that a proposal is endogenously selected by a proposer, the identity of whom is determined randomly and independently across periods. We show that the model generates similar predictions both for the case of simple majority rule and unanimity rule. In particular, the expected group decision is a convex combination of individual opinions, and depending on the level of patience, it can span the range between the mean individual opinion (in the case of low levels of patience) and the median individual opinion (in the case of high levels of patience). In general, the model predicts that relative position within the group matters in how much influence the individual has on the group decision, and in particular members closer to the median member have more influence than extreme group members. An alternative modeling approach for group decision making over spatial policy outcomes is proposed by Compte and Jehiel (2010). They assume that proposals to be voted on emerge according to an exogenous process.[8] The main prediction from this model is that if some player can influence the expected group choice (which is when members are not too impatient) then it is either only the median member (in case the group adopts a simple majority voting rule) or exactly two members. In case of the group adopting a unanimity rule, the latter two members are the ones with extreme ideal points. For supermajority voting rules other than unanimity, the influential members can be closer to the median.

Our empirical findings are as follows. First, we find that the coefficient of the constant is insignificant, and we cannot reject the hypothesis that the sum of the coefficients of members' individual decisions is one. This is consistent with the group decision being a convex combination of the members' decisions. A constant significantly different from zero would indicate a level shift in group decisions, suggesting that the group decision situation itself sways members' preferences in a particular direction, independently of initial opinions. Second, the median group member always has a significant effect on the group choice. However, some (but not all) of the other group members also have an impact on the group choice. In the gift-exchange context, the members immediately above and below the median have a significant impact, but the members at the extremes do not. In the lottery choice context, besides the median, the second least risk-averse and the most risk-averse group members seem to be influential. Overall, while there is a tendency for groups to ignore extreme individual opinions, the most risk-averse member has some influence on the group decision, possibly because the arguments that can be brought up to support risk-averse choices are particularly persuasive.[9] In both settings we can reject the "mean hypothesis" that all members' opinions matter equally, and the "median hypothesis" that only the median member's opinion matters,[10] even though our results confirm that the median member has a significant influence.

The empirical results are broadly consistent with the predictions of the spatial bargaining models summarized above, for cases when members' patience is from an intermediate range. In the Banks and Duggan (2000) model this is the case when the acceptance set is likely to include the ideal points of members next to the median, but less likely to include the ideal points of extreme members. The Compte and Jehiel (2010) model can also explain the observed outcomes if members are not too impatient and groups tend to adopt a supermajority rule but not unanimity rule. In this case the theoretical prediction is that there are exactly two influential members, but their identities depend on the

[2] The investigation of risk attitudes of groups versus individuals started with Stoner (1961). See also Teger and Pruitt (1967), Burnstein et al. (1973), and Brown (1974). Recent papers in economics include Shupp and Williams (2008), Baker et al. (2008) and Masclet et al. (2009). Groups' attitudes towards cooperation and reciprocity were first examined in the context of prisoner's dilemma games: see Pylyshyn et al. (1966), Wolosin et al. (1975), Lindskold et al. (1977), Rabbie (1982), Insko et al. (1990), and Schopler and Insko (1992). Wildschut et al. (2003) provide a meta-analysis of the subject, while Charness et al. (2007) is a more recent contribution in economics. Other treatments investigate centipede games (Bornstein et al., 2004), ultimatum games (Bornstein and Yaniv, 1998; Robert and Carnevale, 1997 and dictator games (Cason and Mui, 1997; Luhan, Kocher and Sutter, 2009). Closest to our work is Kocher and Sutter (2007), who investigate gift exchange games similar to ours.

[3] Among the papers closest to our experimental design, Cason and Mui (1997) use two-person groups, while Luhan et al. (2009) use three-person groups.

[4] Besedes et al. (2014) also feature a design (in the context of an intellective task) in which individual opinions for what the group decision should be are solicited, and afterwards it is randomly determined which individual decision applies to all group members. However, in this treatment team members do not deliberate and make a group decision. In our experiment we observe both what the individuals would choose for the group before deliberation, and the group decision that the same individuals agree upon after deliberation.

[5] In the gift-exchange games, ordering is based on the extent of reciprocation of the first mover's gift. In the lottery choice problem, ordering is based on the frequency of choosing the safer (low-spread) lottery over the high-spread lottery in a list of lotteries with increasing odds of the higher outcome. In the main text we report results from OLS specifications, as the interpretation of regression coefficients is clearer in this case. In the Supplementary Appendix we also provide Tobit specifications and show that all our results qualitatively remain the same.

[6] For example, Teger and Pruitt (1967) and Myers and Arenson (1972) focus solely on comparing mean individual and mean group decisions. We are aware of five papers that examine the relationship between individual preferences and the group decision: Fiorina and Plott (1978), Corfman and Harlam (1998), Arora and Allenby (1999), Zhang and Casari (2012), and Casari et al. (2012). In the first three of the above papers preferences are exogenously imposed by the experimenter, essentially constructing pure bargaining situations. Zhang and Casari report on experiments in a lottery choice context, conducted in parallel to ours, in which subjects offer proposals to each other until an agreement is reached, where members' initial proposals are interpreted as their individual preferences. Casari et al. consider a very similar design in the context of an intellective task, bidding for a company takeover. However, the proposals in these experiments are suggestions to other group members, and might reflect strategic considerations to influence the subsequent group discussion, and hence cannot be interpreted as bids members would choose if they were dictators for the group. In contrast, the opinions solicited in our experiment before group discussions are not revealed to other members.

[7] For a detailed discussion on how various social decision schemes affect the ways in which the distributions of group and individual choices might differ, see Kerr et al. (1996).

[8] They also examine cases when members can exert costly effort to influence the proposal process.

[9] The persuasive argument theory (Brown, 1974; Burnstein et al., 1973), which originated in social psychology, posits that deliberation drives group decisions in a particular direction because arguments in that direction are more persuasive. A related explanation is that people with certain preferences tend to be more persuasive than others (for example, more selfish individuals are also more aggressive in deliberation).

[10] The latter would hold theoretically under a simple majority voting rule provided preferences are single-peaked (see Moulin, 1980).

exact locations of ideal points (besides the voting rule adopted). Another possibility generating similar predictions in this context is when some groups adopt a simple majority rule, while others a supermajority rule.

The above empirical analysis does not distinguish between the direct effect of relative positions within a group on the group choice and the effect of unobserved characteristics of individuals which may be correlated with their relative positions. We investigate this issue utilizing the experimental design feature that each subject participates in multiple groups and decisions. We compare three econometric models explaining the absolute difference between an individual's decision and the group decision. In the first one, the independent variables indicate the relative position of the individual in the group. In the second model we only allow subject-specific fixed effects to explain the difference between group and individual. Finally, in the third model we allow for both types of explanatory variables. We find that the first model has better explanatory power than the second one, and the combined third model improves explanatory power only by a modest amount relative to the first model. Controlling for individual fixed effects does not change significance and magnitude of the coefficients of relative position. A robust finding from these specifications is that being at either of the extreme relative positions significantly increases the difference between the group decision and the individual's decision, relative to the median member's difference to the group. At the same time, being at a non-median position next to the median does not increase the above difference significantly, relative to the median.

The above findings are also useful for interpreting aggregate level differences between individual and group choices. Consistent with earlier papers, we observe that groups on average reciprocate less than individuals in the gift-exchange setting — standardly referred to as the "selfish shift." Our data suggests that this shift is not attributable to subjects behaving differently in groups, [11] or to arguments to be selfish being especially persuasive, as in the persuasive argument theory. The influences of the group members immediately next to the median roughly cancel each other out, and on average the group choice is very close to the median member's choice. The selfish shift arises mainly because the distribution of individual preferences is skewed: in particular the median member's preferred reciprocation level is below that of the mean.[12] For lottery choices, groups are on average more risk-averse than individuals. Although the most risk-averse member influences the group choice, she is not the main driver of this "cautious shift", as her influence is roughly canceled out by that of the second least risk-averse member, and again on average the group choice is very close to the median. The cautious shift arises primarily because the median individual choice is more risk-averse than the mean one. This observation also suggests an explanation for why earlier experiments sometimes found cautious shifts, while others found risky shifts: namely, that for some types of lottery choice problems the median individual tends to be more risk-averse than the mean, while in other types the opposite holds.[13]

In both of our settings, the variance of group choices is smaller than that of individual decisions, mainly because extreme member preferences tend to be curbed by groups. This suggests that, in the types of decision contexts we examine, delegating the decision to a committee can reduce the variability of outcomes, and reduce the likelihood of extreme outcomes. [14]

We also find evidence of social influence in our experiments, in that group choices affect the subsequent individual choices of subjects. In particular, individuals tend to adjust their individual choices in the direction of prior group decisions. Our participants have no incentives to misstate individual preferences due to social pressure, since other group members only receive information about this choice at the very end of the experiment, and in a non-identifiable manner. This suggests that the group decision process does not just involve bargaining and compromise, but also persuasion, i.e. members trying to change each others' opinions.[15] We also find that the members who tend to change their opinions (in the direction of the previous group decision) are the extreme ones, hence deliberation leads to depolarization of opinions in the settings we examine. This finding suggests that social decision making, as in a deliberative democracy, could have an important role in preference formation, besides preference elicitation and aggregation.

An important feature of our experimental design is that group members can freely deliberate (face-to-face communication in a private room, with no experimenter present) and can select their own way to arrive at a group decision. This is motivated by the observation that for many real-world group decision problems, there is no externally-imposed decision mechanism (such as a voting rule), and there are no hard constraints on the amount of deliberation before the decision. This aspect is an important difference between our work and Goeree and Yariv (2011)'s recent experimental investigation of collective deliberation, in which different voting mechanisms (simple majority, two-thirds majority, and unanimity) are imposed.[16] In general, as there are some settings in which a decision rule is externally imposed and others in which there is no such constraint, we believe the investigation of both cases is warranted.

As shown recently by several papers (Charness and Jackson, 2009; Charness et al., 2007; Chen and Li, 2009; Sutter, 2009), individuals' decisions depend on whether their consequences only apply to them or to the entire group to which they belong.[17] Our main focus is not directly related to this effect, as we are interested in comparing initial individual opinions on what the group should choose to the group choice itself. Nevertheless, to compare this effect to the effects of preference aggregation, in one of our sessions we also solicited group members' choices before deliberation in a scenario in which the choice was only payoff-relevant for themselves (and all other group members received a constant payment independent of the choice). In this session

[11] An influential explanation in social psychology, the social comparison theory, argues that people behave fundamentally differently in group settings than individually (Levinger and Schneider, 1969). It posits that people are motivated both to perceive and to present themselves in a socially desirable way. To accomplish this, a person might react in a way that is closer to what he regards as the social norm than how he would act in isolation. There are several other explanations of why people make decisions differently in groups, that apply to particular types of choices. The identifiability explanation (Wallach et al., 1962, 1964) claims that people in group decisions act more selfishly because the other side's ability to assign personal responsibility is more limited. In the context of lottery choices, Eliaz et al. (2005) point out that subjects who are not expected utility maximizers exhibit a group shift, because the decision problem associated with the possibility of being pivotal in a group's lottery choice decision differs from individually deciding on the lottery choice if the probability of being pivotal is less than 1.

[12] The observation that the distribution of attitudes towards reciprocity is skewed towards the selfish direction is made, for example, in Ledyard (1995), Palfrey and Prisbey (1997), Brandts and Schram (2001), Fischbacher et al. (2001), and Ambrus and Pathak (2012).

[13] As Hong (1978) demonstrates, the cultural setting can also influence the direction of the shift.

[14] In different contexts, particularly those in which groups are asked to form a political opinion, deliberation can lead to extremization of opinions (Manin, 2005; Sunstein, 2000, 2002), although it can also lead to depolarization of opinions (Burnstein, 1982; Ferguson and Vidmar, 1971).

[15] We use the term *persuasion* in a broad sense: group decision-making changing a participant's privately stated individual choice in subsequent rounds; our design does not allow us to distinguish between different (psychological) channels of persuasion.

[16] The experiments of Goeree and Yariv also differ from ours with regards to the decision tasks, as well as many aspects of experimental design (including the nature of communication, which is impersonal in their experiments, via a computer network). Hence, our results are not directly comparable to theirs. For an earlier experimental investigation of group decision making with externally imposed voting rules, see Bower (1965). For a recent paper featuring both deliberation and externally imposed voting rule, see Elbittar and Gomberg (2012).

[17] This is related to the in-group versus out-group sentiments theory in social psychology (Kramer, 1991; Tajfel et al., 1971), which posits that subjects develop more other-regarding preferences towards their group members than towards subjects outside the group.

it was randomly selected whether the group choice, or one of the individual-for-group choices, or one of the individual-for-individual choices got implemented.[18] Consistent with the existing literature, we find that individuals reciprocate less when deciding for the group than for only themselves, and are less risk-averse. In both of the examined settings, the magnitudes of these differences are similar to the differences between average individual-for-group choices and average group choices resulting from preference aggregation.

2. Experimental design and procedures

To explore how individual opinions are aggregated in groups, our experiment utilizes non-intellective decision-making situations from the two main domains of economic experiments: strategic social interaction and non-strategic, individual decision making. We confront subjects with the choice of a second mover in a gift-exchange game, and with a list of binary lottery choice situations. As we elicit both individual and group choices from the same individuals over the same decision task for the group, our design allows us to observe the aggregation of individual choices to group decisions.

The first game featured in our experiments is structurally the same as the one in Brandts and Charness (2004), and following their terminology we refer to it as a gift-exchange game.[19] In our version of the game, a first mover and a second mover are each endowed with 10 tokens, which have monetary value. First, the first mover may send a gift of 0 to 10 tokens to the second mover. The amount is deducted from the first mover's account, but is tripled on the way before being awarded to the second mover. Then the second mover decides whether to send a gift of 0 to 10 tokens to the first mover under the same conditions: for each token sent, one token is deducted from the second mover's account, and triple the amount is added to the first mover's account. While the socially optimal behavior is to exchange maximal gifts, in the unique Nash equilibrium outcome neither player contributes any gift.

The typical experimental data on this game shows that first movers extend trust and there is a significant likelihood of reciprocation among second movers, yielding outcomes that are closer to the socially efficient one. Individuals differ both in their degrees of trust as well as in their pattern of reciprocation. In our experiment we concentrate on the latter, studying how individual reciprocal patterns to a diverse set of stimuli are aggregated to group behavior.

For the risk choice situation, we used a version of the risk preference elicitation questionnaire of Holt and Laury (2002). Subjects made ten choices between two lotteries, namely $p[\$11.50] \oplus (1-p)[\$0.30]$ vs. $p[\$6.00] \oplus (1-p)[\$4.80]$ with $p \in \{0, 0.1, 0.2,..., 0.9\}$. Of this choice list, one lottery was randomly selected, the decision implemented and the corresponding lottery played out. Most lottery-choice experiments of this kind observe heterogeneous individual risk attitudes, with a majority of people being slightly to strongly risk averse.[20] Our experiment studies how these individual risk preferences are aggregated to a group risk attitude when the group has to make a decision that applies to all members.

In each session, we had n groups of five participants (with n varying between 3, 4, and 6, depending on session enrollment numbers), plus "individual decision-makers" (see below). Assignment to these roles was random by letting participants draw a numbered card.

For *group members*, each session consisted of three phases. In the first phase, the group made three decisions: two choices as second movers in the gift-exchange game (with two different first movers), and one choice in the lottery task. For each of the three choice situations, group members first made individual decisions (i.e. their choice if they can dictate the group decision), and submitted their decisions secretly to a research assistant. Then the research assistant left the room, and the group members freely discussed and made a group decision. After the three decisions, the initial random assignment to the n groups g was reshuffled by assigning each group member i to her new group $g + i(mod n)$. The newly formed groups made again two gift-exchange decisions and one lottery decision, following the same procedures as in Phase 1. After that, group participants were reshuffled again according to the same procedure, and made another two gift-exchange decisionsand one lottery decision in Phase 3.

The additional six *individual decision-makers* in each session were only used to generate first-mover choices for the six group gift-exchange decisions. Specifically, they made n first-mover decisions in a row at the beginning of the experiment, without any feedback, with n equal to the number of groups in the session.[21] Afterwards they had to stay in the lab until the end of the session.

The experimental sessions took place in 2008 and 2010 in the Computer Laboratory for Experimental Research at Harvard Business School. We conducted seven sessions with a total of 172 student subjects, each session comprising either 21, 26, or 36 participants, and lasting approximately 1 h and 30 min. The experiment was computerized using the z-Tree software (Fischbacher, 2007). After subjects arrived instructions were distributed.[22] An experimenter (the same in all sessions) led subjects through the instructions and answered open questions. Then, subjects were randomly assigned to be group or individual participants, and group participants were led to small group rooms to make their decisions, while individual participants stayed in the main lab. After all decisions in all three phases had been made, group members filled in a post-experimental questionnaire asking for demographic data and containing open questions for motivations of subjects' decisions.[23]

At the end of the experiment all participants were paid in cash. Tokens for the gift-exchange game were converted to real money at a fixed exchange rate, plus subjects received an additional fixed show-up fee of \$10.[24] Group members earned the income from each gift-exchange game and from one randomly chosen of the three lottery questionnaire choices they were involved in. Subjects were told that for each of those choices with 50% probability one of the individual decisions would become the effective group choice, with equal probability allocated to every member's decision, and that in this case it would not be revealed which individual's decision was implemented. With the remaining 50% probability, the group's joint decision became the effective group choice.

In the last session, in addition to individuals' choices on behalf of the group, we also elicited individuals' choices on behalf of themselves. That is, if an individual-for-individual decision was randomly selected at the end of the experiment, then the corresponding choice would only be

[18] Cason and Mui (1997) and Luhan et al. (2009) also solicit individual-for-individual decisions, besides group decisions, but they did not solicit individual-for-group decisions, which is the central component of our analysis. Furthermore, in these studies, individuals interact with individuals, and groups interact with groups, while the first-mover in our gift-exchange game is always an individual.

[19] The term gift-exchange game was introduced by Fehr et al. (1993). Gift-exchange games are similar in structure to trust games, and can be more generally classified as sequential social dilemma games.

[20] A typical result observed in Holt and Laury (2002) lottery tasks is that some participants make non-monotonic/inconsistent choices, i.e. make more than one switch between the safe and risky lotteries when going down the ordered list of choices. As Holt and Laury (2002), we consider the total number of safe choices per lottery task as our outcome measure. Overall, we observe relatively little inconsistency. While 3.1% of individual lottery tasks exhibit non-monotonic choices (5.4% of the very first tasks in a session), none of the observed group decisions was inconsistent.

[21] About half of first movers in our sessions did vary their offers, despite receiving no feedback between offers, while the other half didn't. Each decision of a single first mover was played against a different group.

[22] Instructions are included in the Supplementary Appendix.

[23] We did not find any systematic effects of specific demographic characteristics (gender, age, field of studies) moderating influence on group decisions. Also see below our results from regressions with individual-specific effects that may absorb any effects of demographics.

[24] The exchange rate for gift-exchange games of \$0.10 per token was verbally announced at the beginning of sessions.

implemented for that individual, while the other four group members received a fixed payment of 15 tokens/$5.50 for that choice. This allows us to test in a small sample whether individual-for-group choices differ substantially from individual-for-individual choices in our two settings. In this session, the 50% probability of individual choices to matter for payoff was split equally between individual-for-group and individual-for-individual choices.

Overall, we collected five individual choices and one group choice for each of the 156 gift-exchange games and 78 ten-decision lottery tasks. In the last session, we collected an additional 90 individual-for-individual decisions for the gift-exchange game and 45 individual-for-individual decisions for lottery tasks.

3. Theoretical background

The question of group members negotiating over outcomes in a non-transferrable utility setting (no side payments among group members) has been analyzed theoretically. One approach is provided in the spatial bargaining model of Banks and Duggan (2000), which is an adaptation of the multilateral bargaining model of Baron and Ferejohn (1989), popular in modeling legislative bargaining in political science, to settings with non-transferrable utilities.[25] An alternative approach is provided in Compte and Jehiel (2010), who also propose a generalization that encompasses both model frameworks. The common features of the two approaches are the following:

1. They assume members' utilities over outcomes in a policy space as a negative function of the distance between the group's choice and the exogenously specified ideal point of the member.
2. They both consider a multi-stage game in which at every stage a proposal is being considered and group members vote whether the proposal is accepted or not. The proposal passes if the proposal is supported by a winning coalition. In the most natural specifications of the models, when group members are ex ante symmetric and the voting rule satisfies a monotonicity property, this corresponds to at least k members supporting the proposal, where k can range between the smallest integer larger than half of the group members (simple majority rule) and the total number of group members (unanimity rule).
3. Group members discount payoffs from policy outcomes and prefer to get to an agreement earlier rather than later.

The difference between the two approaches is in the way proposals come about. Banks and Duggan (2000) follow a more conventional approach, going back to Rubinstein (1982), in that the proposal at any stage is chosen by a group member. In particular, the proposer is chosen randomly and independently across time, according to the same probability distribution, at every stage. In contrast, in the basic model of Compte and Jehiel (2010), proposals arise according to an exogeneous process. Below we summarize some predictions from both of these modeling approaches. There is also a possibility that besides pure bargaining, members try to persuade and change the ideal points of each other, during group discussion. This is discussed at the end of the section.

First we consider a 1-dimensional version of the spatial bargaining model in Banks and Duggan (2000). In the simplest version of the model, players' ideal points are commonly known, and acceptance of a proposal requires a simple majority. Later we show that the predictions generated from this model extend, in weaker forms, to versions of the model in which ideal points are private information, and when accepting a proposal requires unanimity.

Formally, consider a 5-player dynamic bargaining game, in which in each period ($t = 0, 1, 2, \ldots$) the proposer is chosen i.i.d., with each of the five players chosen with probability 1/5. A proposer must propose an action on [0, 1]. Player i's type, or ideal point, is $x_i \in 0, 1]$, and her payoff when an agreement y is reached at period t is $\delta^t(1-(x_i-y)^2)$, where $\delta \in (0, 1)$ is a common discount factor. Without loss of generality, assume $0 \le x_1 \le x_2 \le x_3 \le x_4 \le x_5 \le 1$. After an offer is made, players announce sequentially, in a random order, whether they accept or reject the offer. The proposal is accepted if and only if at least 3 group members accept.

The assumption of a 1-dimensional state space corresponds to the existence of a natural ordering of choices in both types of decision problems featured in our experiments (amount to be reciprocated in the gift-exchange game, and the switching point from safe to risky lotteries in lottery choices). We assume a continuum of choices for analytical convenience only: the qualitative predictions of the model remain the same if the set of possible action choices is a finite subset of the unit interval. We focus on stationary subgame-perfect equillibria (SSPE) of the game, that is, subgame-perfect equilibria (SPE) in which players' proposal strategies and acceptance rules do not change over time.[26]

In this game, for any specification of ideal points and the discount factor, SSPE is unique, and it is characterized by a compact set of acceptable proposals $A \subseteq 0, 1]$, of which x_3 is an interior point (relative to [0, 1]). When given the opportunity, each player proposes the point of A closest to her ideal point. This implies that the median player always proposes her ideal point. Other players either propose their ideal points (if included in A) or the closest boundary point of A. The size of A decreases with δ, shrinking to the median member's ideal point as members get very patient. The theorem below formalizes these statements. The proof is in Appendix A, and it explicitly solves for the unique SSPE in every region of parameter values.[27]

Theorem 1. *In a bargaining game with publicly known ideal points and plurality decision rule, SSPE is unique. For any set of ideal points, the set of acceptable proposals A monotonically decreases in δ, it converges to the whole choice set as $\delta \to 0$, and it converges to the median ideal point as $\delta \to 1$.*

The proof in Appendix A also reveals that for generic specifications of ideal points – when the median is not an extreme point of the choice set – for low enough δ in fact $A = [0, 1]$ and hence all players can propose their ideal points. Conversely, for high enough δ only players with an ideal point equal to the median can propose their ideal points. This means that if players are impatient enough, then the expected action chosen by the group is equal to the mean of the ideal points, while if players are very patient then the expected group choice is very close to the median ideal point. In general, the equilibrium prediction of the model for the expected group decision, for different levels of patience, spans the range between the mean and median ideal point. For intermediate levels of patience only players closer to the median can propose their ideal points, and player 2 (the less extreme member to the left of the median) is more likely to influence the expected group decision than player 1 (the left extreme member). Similarly, player 4 is more likely to influence the group decision than player 5. In short, the prediction is that if players are not very patient then non-median members' ideal points can also affect the expected group decision, but extreme members' ideal points are less likely to be influential than the ideal points of non-median members next to the median.[28]

[25] The origins of random proposal bargaining, which is at the heart of these models, go back to Binmore (1987).

[26] There are several recent papers that provide micro-foundations for selecting Markov perfect equilibria (equivalent to SSPE in our context) in sequential-move games. See Bhaskar and Vega-Redondo (2002) and Bhaskar et al. (2013) for general classes of asynchronous-move games, and Ambrus and Lu (2015) in the context of dynamic bargaining.

[27] The result on convergence to the median when players become patient also follows from Theorem 5 of Banks and Duggan (2000).

[28] When applying the model directly to the groups we observe in our experiment and predicting group choices from elicited individual preferences, and then regressing those predictions on the individual choices using an OLS model (as we do in our main empirical analysis reported below), we find exactly the discussed pyramid shape in coefficients on ordered individual opinions, the steepness of which is moderated by δ. In addition, all linear regressions have $R^2 > 0.99$, indicating that a linear approximation is able to capture the model's predictions almost perfectly (non-linear components of the model only play a minor role).

The above analysis assumes that the ideal points of group members are known at the beginning of the bargaining game. In the experiments, we kept the individual decisions preceding the group discussion private information. Therefore, taken literally, the above model implicitly assumes that members truthfully reveal their ideal points to each other before they start bargaining. This can be the case for example when there are costs associated with lying or misrepresenting one's true preferences. If, however, subjects are strategic, then the bargaining game is one of multi-sided asymmetric information. Such games are notoriously difficult to analyze, and they tend to have a severe multiplicity of equilibria.[29] Nevertheless, below we show that some of the qualitative features carry over to situations in which ideal points are privately known, and players are strategic in reporting their ideal points.

To formally analyze the situation with incomplete information, we modify the previous model in two ways (besides assuming that ideal points are private information). First, we augment the game with one round of cheap talk before bargaining, in which members simultaneously and publicly announce ideal points. Second, here we assume that the set of possible action choices, and hence the set of possible types of players, is a finite subset of (0, 1). Assume that the distribution of types is i.i.d. across members, with distribution function F allocating positive probability to each possible type. The next result shows that for both very impatient, and for very patient players, there exist sequential equilibria in which players truthfully announce their types at the beginning of the game, and then bargain the same way as in the game with publicly known ideal points. In particular, this results in all members proposing their ideal points when the discount factor is low enough (hence the expected group decision is the mean ideal point), and all members proposing the median ideal point when the discount factor is high enough. The proof of Theorem 2 can be found in Appendix A.

Theorem 2. In a bargaining game with private ideal points and simple majority rule, preceded by a round of cheap talk, for low enough discount factors there is a sequential equilibrium in which members announce their types truthfully, and then all members propose their ideal points in the bargaining phase. Moreover, all sequential equilibria are outcome-equivalent to the equilibrium above. Conversely, for high enough discount factors there is a sequential equilibrium in which members announce their types truthfully, and then all members propose the median announced ideal point in the bargaining phase.

In both versions of the model above, it is assumed that groups use simple majority rule to get to an agreement. However, similar results hold, with some caveats in the case of very patient players, when the group adopts some supermajority rule or unanimity as decision criterion. In a Supplementary Appendix we report a full analysis of the unanimity rule case. Similar to the case of simple majority, the acceptance set in an SSPE converges to the set of all possible choices as $\delta \to 0$ (and hence when members are very impatient, all of them can propose their ideal points), it monotonically decreases with δ, and converges to a single action choice as $\delta \to 1$. The difference is that in general the latter limit point is between the mean and the median ideal point under unanimity, although when the median is equidistant to the two extreme ideal points, then the limit point is exactly the median.[30]

Summarizing the predictions of various versions of the Banks and Duggan (2000) type spatial bargaining model:

i. the expected group decision is a convex combination of individual decisions;
ii. if players are very impatient then the expected group decision is close to the mean individual decision;
iii. if players are very patient then the expected group decision tend to be close to the median individual decision (requires extra assumptions if the group decision rule is not simple majority);
iv. for intermediate levels of patience, the further a member is from the median, the less effect she has in expectation on the group decision (same caveat as for the previous prediction).

The bargaining model analyzed in Compte and Jehiel (2010) has the same basic structure as the complete information version of the Banks and Duggan (2000) model described above. The difference is that the proposal to be voted on, x, at every period $t = 1,\ldots,$ is independently drawn from a set of possible proposals X, from the same distribution $F(x)$ with density $f(x) \in \Delta(X)$. The latter distribution is assumed to satisfy certain technical requirements. In our context, a sufficient condition for the latter is that the support is the whole interval [0, 1] and $f(x)$ is continuously differentiable and bounded away from 0.

In this model, just like in Banks and Duggan (2000), there is a unique SSPE, characterized by a set of acceptable proposals A. If $F(x)$ has full support then A is an interval contained in [0, 1]. A further similarity is that as $\delta \to 0$, A becomes X (any proposal is accepted). In this model this does not imply though that the expected group decision is the mean ideal point of group members. Instead, it is the expectation of $F(x)$, an exogenous parameter of the model (and hence unaffected by members' preferences). As δ increases, A shrinks, and in a 1-dimensional setting as in ours, it converges to a single point as $\delta \to 1$. Compte and Jehiel (2010) show that for general δ, set A is influenced by at most two group members in a 1-dimensional setting. In case of an odd number of members and simple majority rule, only the median member has an impact on A. For stricter majority rules, there are two key members, and their identities depend on the parameters of the model. Under unanimity rule the two key members are those with extreme ideal points.

To summarize, the model of Compte and Jehiel (2010) implies that members' preferences only influence the expected group decision if members are patient enough. If members are patient, the identity of the influential member(s) depends crucially on the voting rule: in case of simple majority only the median member has an influence, in case of unanimity, the two extreme members, while for in-between voting rules there are exactly two influential members, whose identities depend on various parameters of the model. Compte and Jehiel (2010) also show that these basic predictions extend to a generalization of the model in which members have some imperfect influence on the emerging proposals, subject to a technical condition.

Lastly, the models above can be extended to incorporate the possibility of persuasion, in particular others in the group persuading a member to change her ideal point (as opposed to just agreeing upon a compromise). Note that since the tasks in our experiments are not intellective, persuasion in our context does not involve transmission of hard information as in standard economic models of persuasion (for example Kamenica and Gentzkow, 2011), and hence has to include a psychological component, as in Cialdini and Goldstein (2004) and Cooper and Rege (2011). There are two straightforward ways to incorporate this into the theoretical framework. One possibility is that it is the group decision itself that persuades members to change their ideal points, presumably towards the group decision. Formally, the shifted ideal point is: $\hat{x}_i = g^i(x_i, y - x_i)$, where x_i is the member's ideal point before the group decision, and $y - x_i$ is the difference between the group decision and the original ideal point, with $g^i_1, g^i_2 \geq 0$. The function determining the shifted ideal point is indexed by the member's position in the group, allowing for relative position within the group influencing the magnitude of the shift. In this case persuasion does not directly affect the analysis of the current group decision process, instead the outcome of the decision process changes the ideal points thatsubjects start out within the next group decision problem. Another possibility is that persuasion happens during communication among group members,

[29] See for example Fudenberg and Tirole (1983) and Cramton (1984, 1992).

[30] Note that this condition does not restrict x_2 and x_3 to be the same distance from the median, and therefore for specifications satisfying this condition the median typically differs from the mean.

before the bargaining stage. This leaves the formal analysis unchanged, with the caveat that the relevant ideal points in the bargaining stage are not the originals, but the ideal points after persuasion. In this case, the effect of an individual's original ideal point on the eventual group decision measures both the individual's capacity to affect others' ideal points through persuasion as well as the individual's leverage to influence the expected outcome of the resulting bargaining game.

4. Competing hypotheses

Our stylized theoretical considerations motivate a number of competing hypotheses for the experiment. We investigate situations in which five individuals first make an individual choice, and then jointly decide on a group choice, for the same decision problem. Our main empirical model can be generally written as:

$$y_{gt} = f(x_{i_1gt}, x_{i_2gt}, x_{i_3gt}, x_{i_4gt}, x_{i_5gt}, X_{gt})$$

where t stands for a time period (at which every group is associated with one particular decision problem), y_{gt} is group g's observed decision in period t, x_{igt} is the observed decision of individual i in the same period, and X_{gt} is a vector of characteristics of the decision situation. We will use $x_{gt}^{(j)}$ to refer to the jth lowest decision among the individuals in group g (in particular $x_{gt}^{(1)}$ referring to the lowest, and $x_{gt}^{(5)}$ referring to the highest decision).

For gift-exchange games, variable x_{igt} corresponds to member i's individual reciprocation decision to the first mover, relative to the first mover's offer.[31] Thus, $x_{gt}^{(1)}$ represents the reciprocation of the most selfish member, minus the first mover's offer, while $x_{gt}^{(5)}$ represents the reciprocation of the most generous member, minus the first mover's offer. In the lottery tasks, choices might be inconsistent in that individuals might switch more than once from the safe to the risky lottery (about 3% of our subjects do so), such that their "cutoff" is not well-defined. We follow Holt and Laury (2002) and use the total number of safe choices per lottery task as a measure of risk aversion. (Our results presented below would not be qualitatively different when we used the position of the first switch to the risky lottery instead.) So for lottery choices, $x_{gt}^{(1)}$ equals the lowest number of safe choices in a group and corresponds to the most risk-loving group member, while $x_{gt}^{(5)}$ equals the number of safe choices of the most risk-averse group member.

We focus on models in which the group decision is a linear function of $(x_{igt})_{i=1,\dots,5}$:

$$y_{gt} = \alpha + \beta_1 x_{gt}^{(1)} + \beta_2 x_{gt}^{(2)} + \beta_3 x_{gt}^{(3)} + \beta_4 x_{gt}^{(4)} + \beta_5 x_{gt}^{(5)} + \epsilon_{gt}. \quad (1)$$

First we investigate whether there is a systematic shift between the group's decision and the individual decision. In particular, we test the hypothesis that $\alpha = 0$. Note that $\alpha \neq 0$ would imply that there is a difference between individual and group decisions that is independent of members' preferences.

Next, we investigate the hypothesis that the coefficients of individual decisions sum up to one:

$$\sum_{i=1}^{5} \beta_i = 1.$$

Note that $\alpha = 0$ and $\sum_{i=1}^{5} \beta_i = 1$ imply that the group decision is a convex combination of individual decisions; hence, the coefficients of the latter can be interpreted as the weights of different members in shaping the group decision.

[31] In the Supplementary Appendix we also present an alternative specification for gift-exchange games in which we do not normalize the reciprocation decisions by the first offer, and instead add the first offer as an additional control variable in the regressions. The results from this alternative specification are qualitatively the same.

Table 1
Average choices at individual and group level in gift-exchange games and lottery tasks.

	Gift-exchange decisions (Normalized amount returned)			Lottery decisions (Number of safe choices)		
	N	Avg	StdDev	N	Avg	StdDev
Individual	780	−3.86	3.70	390	6.31	1.30
Group median	156	−4.21	3.34	78	6.44	0.73
Group	156	−4.31	3.29	78	6.41	0.86
Only last session						
Individual for individual	90	−2.29	2.71	45	6.56	1.16
Individual for group	90	−2.44	2.18	45	6.29	1.24

Next we examine the *mean hypothesis*, which implies that the group's decision is simply a function of the mean individual decision:

$$y_{gt} = \alpha + \beta\left(\frac{1}{5}\sum_{i=1}^{5} x_{gt}^{(i)}\right).$$

That is, the mean is a sufficient statistic for the group's decision. If $\beta = 1$, then the mean exactly predicts the component of the group decision that varies with individual preferences. In our econometric tests, we test whether we can reject the hypothesis that $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5$. The version of the mean hypothesis which further requires the mean to exactly predict the group decision, what we call the *strong mean hypothesis*, involves tests of the hypothesis that $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \frac{1}{5}$.

A competing hypothesis, the *median hypothesis*, implies that the group's decision is a function of the median individual decision only, so that

$$y_{gt} = \alpha + \beta x_{gt}^{(3)}.$$

In our econometric tests, we estimate Eq. (1), and test whether we can reject the hypothesis that $\beta_1 = \beta_2 = \beta_4 = \beta_5 = 0$. The version of the hypothesis which further requires the median to exactly predict the group decision, what we call the *strong median hypothesis*, involves testing whether we can reject both $\beta_1 = \beta_2 = \beta_4 = \beta_5 = 0$ and $\beta_3 = 1$.

The last hypothesis, the *extreme-irrelevance hypothesis*, implies that extreme opinions are ignored in the group decision: $\beta_1 = \beta_5 = 0$.

5. Results

5.1. Summary statistics

The upper part of Table 1 lists averages of gift-exchange and lottery decisions at the individual and group level. We observe that on average, both individuals and groups reciprocate less as second movers than what they received from the first mover (the returned gifts, normalized by the received gifts, are negative), and are generally risk-averse. Compared to the mean individual decision, the median decision is less altruistic and more risk-averse. On average, group decisions are between the mean and median individual decisions, but closer to the median.

We use Datta and Satten's (2008) signed-rank test for clustered data to compare our matched data on group means/medians and group decisions, with clustering at the session level.[32] For gift-exchange decisions, we find that groups reciprocate less than the average member ($p = 0.035$), while there is no difference between group decisions and the

[32] Due to our group rematching procedure, our data is clustered at the session level. The tests we use are similar to Wilcoxon signed-rank tests or Sign tests, but allow the distribution of pair-wise differences to differ between clusters. Details are given in Datta and Satten (2008). We are grateful to Tom Wilkening for providing us with an Excel macro implementation of the test. We obtain the same qualitative results when using session averages and traditional Wilcoxon signed-ranks tests.

median members' decisions ($p = 0.277$). For lottery decisions, we observe a similar pattern: group decisions are significantly more risk averse than the average of members' decisions ($p = 0.037$), but not different to the median member's decision ($p = 0.449$).

Standard deviations of group and individual choices are reported in columns 3 and 6 of rows 1 and 3 of Table 1, respectively. We find that the variance of group choices is smaller than the variance of individual decisions, for both decision tasks. To provide statistical corroboration for this observation, which controls for the clustering in our data, we calculate the variance of group and individual decisions for each session separately, and use these session data points in non-parametric Wilcoxon signed-ranks tests. The tests yield $p = 0.018$ for gift-exchange decisions and $p = 0.018$ for lottery decisions, allowing us to reject the Null hypothesis of equal variances of group and individual choices.

The individual decisions considered above were taken on behalf of the group, such that the object and scope of the group and individual decisions were the same. In the last session, we elicited two types of individual decisions: individual-for-individual and individual-for-group. The lower part of Table 1 contains summary statistics for these two types of decisions in the last session. Individuals in this session give slightly less when deciding for the group rather than themselves (0.15 fewer tokens on average), but are slightly more risk-taking (0.27 fewer safe choices on average) in this case. Testing these differences statistically is difficult, since our independent unit of observation is the session.[33] In general, the magnitude of the difference between individual-for-individual and individual-for-group decisions is similar to the difference between individual-for-group and group decisions caused by preference aggregation.

5.2. Main results

Table 2 reports the results of our main econometric specifications, regressing the group decision on the ordered individual decisions in the group. Models (1) and (3) represent the basic linear specifications, while Models (2) and (4) verify robustness with respect to including session-phase fixed effects into the model.[34]

In all models, the coefficient on the median member's choice, β_3, is positive and significant. In gift-exchange decisions, also the second highest and the fourth highest individual choices contribute to explaining the group choice, while the coefficients on the extreme individual choices β_1 and β_5 are insignificant. In lottery choices, on the other hand, the member one position less risk-averse than the median, as well as the most risk-averse member have a significant effect, besides the median member. When controlling for session-phase effects, the coefficient on $x_{gt}^{(5)}$ stays significant, while the coefficient on $x_{gt}^{(2)}$ just misses the 10%-level with a p-value of $p = 0.115$.[35]

Our results do not support the hypothesis of a level shift — in both models (1) and (3) we do not observe a significant constant α. Table 3 reports the results from post-estimation hypothesis tests which directly test the competing hypotheses in Section 4 for the four OLS models reported in Table 2. We reject both the weak and strong versions of the mean and the median hypotheses for all models: neither are all five individual choices equally important, nor is only the median important

Table 2
OLS regressions of group choices on individual choices.

	Gift-exchange decisions (Normalized amount returned)		Lottery decisions (Number of safe choices)	
Model	(1)	(2)	(3)	(4)
Constant	−0.233	−0.164	−0.224	−0.930
	[0.244]	[0.229]	[0.671]	[1.090]
$x_{gt}^{(1)}$	0.031	0.021	0.062	0.034
	[0.069]	[0.070]	[0.044]	[0.117]
$x_{gt}^{(2)}$	0.347**	0.332**	0.348**	0.313
	[0.101]	[0.120]	[0.101]	[0.170]
$x_{gt}^{(3)}$	0.274*	0.343**	0.318*	0.525***
	[0.118]	[0.095]	[0.151]	[0.141]
$x_{gt}^{(4)}$	0.298**	0.342**	0.101	−0.03
	[0.099]	[0.110]	[0.123]	[0.125]
$x_{gt}^{(5)}$	0.062	0.014	0.206*	0.224**
	[0.042]	[0.093]	[0.089]	[0.067]
Session-phase FE	N	Y	N	Y
N	156	156	78	78
R^2	0.771	0.811	0.505	0.618

Notes: *, **, and *** denote significance at the 10%, 5%, and 1%-levels, respectively. Standard errors are clustered at the experiment session level, and shown in brackets.

and none of the other individual decisions. Consistent with our finding that the constant is not significant in models (1) and (3), we cannot reject the complementary hypothesis that the coefficients on individual choices add up to 1 in any of our models. We also cannot reject the hypothesis that extreme choices do not matter, except for model (4) on lottery decisions which includes session-phase fixed effects.

5.3. Order versus subject-specific effects

The results in the previous section establish that the median member, and certain other members have significant influence on the group decision. However, they do not rule out that there are unobserved individual characteristics, correlated with relative positions in groups, that determine how influential different members are with respect to the group decision. Since each individual participates in 6 gift-exchange and 3 lottery decisions, in three different groups, we can investigate this issue further, at the individual level.

Below we compare three empirical models, in all of which the dependent variable is the absolute difference between the group choice and an individual's choice (Δ_{gi}). In the first model, we regress this variable on a set of independents indicating the relative position of the individual in the given group. In the second model, we aim to explain the same dependent variable purely by subject-specific individual fixed effects. In the third model, we allow both the relative position within the group and individual effects to influence the variable of interest. Formally, the models we investigate are:

$$\Delta_{gi} = \alpha + \beta_1 p_{gi}^{(1)} + \beta_2 p_{gi}^{(2)} + \beta_4 p_{gi}^{(4)} + \beta_5 p_{gi}^{(5)} + \epsilon_{gi}.$$

$$\Delta_{gi} = \alpha + \gamma_i + \epsilon_{gi}.$$

$$\Delta_{gi} = \alpha + \beta_1 p_{gi}^{(1)} + \beta_2 p_{gi}^{(2)} + \beta_4 p_{gi}^{(4)} + \beta_5 p_{gi}^{(5)} + \gamma_i + \epsilon_{gi}.$$

In the above regression equations, γ_i with $i \in \{1,..., N\}$ represent subject fixed-effects, with the constraint $\sum_{i=1}^{N} \gamma_i = 0$. The variables $p_{gi}^{(k)}$ indicate the (tie-weighted) relative position of individual i in group g. If individual i's position in the group is the unique kth lowest within the group, $p_{gi}^{(k)}$ takes the value of 1, while all $p_{gi}^{(k')}$ for $k' \neq k$ take the value 0. If individuals $i_{l_1}, \ldots, i_{l_m}$ are tied at positions $k,..., k+m$ then all $p_{gi}^{(k)}$ for $k \in \{k,..., k+m\}$ take the value $1/|\{k,..., k+m\}|$ and all other $p_{gi}^{(k')}$

[33] Wilcoxon signed-ranks tests, applied to the respective 15 very first (and thereby statistically independent) individual decisions for self and group in this session yield p-values of $p = 0.629$ and $p = 0.103$ for gift-exchange and lottery choices, respectively.

[34] Our hypotheses developed in Section 4 assume a linear structural model of how individual decisions are aggregated to group decisions. Thus, reporting OLS estimates and running post-estimation hypotheses tests on their results (see Table 3) is appropriate here. To test for robustness of our results, we ran the same models as Tobit specifications, and with absolute gifts rather than normalized gifts, including the first mover's offer as a control. The results are reported in the Supplementary Appendix in Tables 7, 8, and 9. In general, all our results and conclusions are robust to the specification.

[35] In Tobit versions of these analyses reported in Table 7 in the Supplementary Appendix, the coefficient of $x_{gt}^{(2)}$ is significant in model (4).

Table 3
Results from post-estimation hypothesis tests, p-values.

Hypothesis	Model (1)	(2)	(3)	(4)
Weak mean	0.024**	0.051*	0.068*	0.061*
Strong mean	0.032**	<0.001***	0.075*	<0.001***
Weak median	0.004***	0.003***	0.016**	<0.001***
Strong median	0.005***	0.005***	<0.001***	<0.001***
Extreme-irrelevance	0.339	0.953	0.131	<0.001***
Convex combination	0.741	0.427	0.739	0.688

Note: *, **, and *** denote significance at the 10%, 5%, and 1%-levels, respectively.

take the value of 0. Thus, for example, if an individual i is tied with another group member at the lowest value in the group, we would have $p_{gi}^{(1)} = 0.5, p_{gi}^{(2)} = 0.5, p_{gi}^{(3)} = 0, p_{gi}^{(4)} = 0, p_{gi}^{(5)} = 0$.

In our estimations, the indicator variable for the 3rd position (median) is omitted to avoid perfect multicollinearity, in the presence of the constant term. Given this, the constant α represents the average difference of the group choice from the median member. Coefficients $\beta_1, \beta_2, \beta_4, \beta_5$ represent how much further away from the group choice the individual decisions of individuals at positions 1, 2, 4, and 5 are. To return to our example of individual i tied with another group member at positions 1 and 2, the post-estimation prediction of the difference between the group and the individual choice of this subject according to the third model above would be the sum of constant, individual effect, and weighted average of position-effects 1 and 2, $\Delta_{gi} = \alpha + 0.5\beta_1 + 0.5\beta_2 + \gamma_i$.

Table 4 summarizes the results of our investigation, which are remarkably similar across our two domains of decisions, gift-exchange and lottery choices. Models (1) and (4) show the estimates from the first equation, which only includes order indicators. While being at positions 1 and 5 implies that the group decision is significantly further away from the individual's decision than from the median's decision, for positions 2 and 4 the difference is not significant or depends on the specification. In models (2) and (4), using only subject fixed-effects as explanatory variables of the difference between an individual's choice and the choice of her group (the second equation above), we observe a drop in explanatory power. Adding these subject fixed-effects to the base model with order indicators (the third equation) increases the explanatory power only modestly, and has basically no effect on the size and significance of the order coefficients.

In sum, the relative position of an individual choice within a group seems to be the most important indicator of the individual's impact on the group's choice, and is particularly robust against controlling for possible (unobserved) individual characteristics which might additionally influence an individual's effect on the group choice.

Table 4
OLS regressions of difference between individual and group choice.

	Gift-exchange decisions			Lottery decisions		
Model	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.314	1.990***	0.391	0.124	0.923***	0.184
	[0.335]	[0.000]	[0.307]	[0.183]	[0.000]	[0.107]
$p_{gi}^{(1)}$	2.780***		2.615***	1.790***		1.630***
	[0.678]		[0.645]	[0.186]		[0.253]
$p_{gi}^{(2)}$	0.385		0.616	0.516*		0.298
	[0.349]		[0.374]	[0.236]		[0.249]
$p_{gi}^{(4)}$	0.911*		0.952*	0.225		0.266
	[0.426]		[0.431]	[0.245]		[0.247]
$p_{gi}^{(5)}$	4.305***		3.812***	1.465***		1.500***
	[0.477]		[0.464]	[0.137]		[0.172]
Subject FE	N	Y	Y	N	Y	Y
N	780	780	780	390	390	390
Adj. R^2	0.225	0.156	0.291	0.363	0.198	0.472

Notes: *, **, and *** denote significance at the 10%, 5%, and 1%-levels, respectively. Standard errors are clustered at the experiment session level, and shown in brackets. Subject fixed-effects are constrained in that their sum is equal to zero.

Table 5
OLS regressions of current individual choices on choices made before in a different group or the same group.

	Gift-exchange (Normalized amount returned)		Lottery (Nb safe choices)
Decision in phase	1st	2nd	
Constant	−1.295***	−1.717***	−0.130
	[0.329]	[0.395]	[0.157]
Own decision in $t-1$	0.734***	0.659***	1.022***
	[0.064]	[0.046]	[0.027]
Diff to group in $t-1$	0.061	0.176***	0.393***
	[0.042]	[0.045]	[0.048]
N	260	390	260
R^2	0.478	0.338	0.679

Notes: For the first gift-exchange decision in a phase, and for all lottery decisions, decisions in t−1 were made in a different group. For the second gift-exchange decision in a phase, decisions in t−1 were made in the same group. *, **, and *** denote significance at the 10%, 5%, and 1%-levels, respectively. Standard errors are clustered at the experiment session level, and shown in brackets.

5.4. Social influence

Our experimental design allows us to study the effect of group decisions on subsequent decisions of the involved individuals. Subjects made two gift-exchange decisions and one lottery decision within the same group, before being rematched to another group in the next of three phases. Thus, for gift-exchange decisions, when looking at the first decision in a phase, the previous decision (in $t-1$) was made in a different group. For the second decision in a phase, however, the previous decision (in $t-1$) was made in the same group. The first two models presented in Table 5 regress the first and second individual choice in a phase, respectively, on the subject's own decisions in $t-1$ as well as the difference between the group's choice and the own choice in $t-1$. We find that the first individual choice in a phase is correlated with the own previous decision, but not with the group choice in the previous, different group. The second choice in a phase, however, is correlated both with the own previous choice and with the difference between group decision and own decision in the same group in $t-1$ (such that when the group decision was more generous than the own decision in $t-1$, the own decision becomes more generous in t).

A similar analysis for lottery decisions is provided in the third model in Table 5. We find that the current decision of an individual, besides being highly correlated with the own previous decision, is positively correlated with the difference between the previous group choice and previous own individual choice.

These results indicate the existence of social influence. In gift-exchange games, the previous group choice only has a significant effect when the group members stay together (that is, for the second gift-exchange situation within a phase).[36] However, the second individual gift-exchange decisions within a phase already incorporate the influence of the first group choice in the same phase, so this just means that the same group members do not have a further influence on each others' individual choices the second time around. The influence of the first group choice in a phase is permanent, in the sense that we do not observe a correction in the opposite direction when individuals part from their previous group members and get into new groups.

To explore whether group members at different positions are differently affected by the group choice in their subsequent decision,

[36] Tobit versions of these analyses yield the same results, see Table 10 in the Supplementary Appendix.

we run extended versions of the models reported in Table 5, with results reported in Table 6. In particular, we interact the difference between an individuals' own choice in $t-1$ and the group choice in $t-1$ with a dummy indicating whether this person's individual choice was at position (1), (2), (3), (4), or (5) in the group in the previous decision situation. In other words, we estimate the effect of the group choice separately by position. For gift-exchange decisions, we find that for second choices within a phase, there is only a significant effect for the least altruistic member, with this member's individual choice becoming less egoistic. For lottery choices, we find relatively strong significant effects for both the most risk-averse and the most risk-loving person, both of whom adjust their subsequent lottery task decision towards the previous group choice.[37]

These results suggest that deliberation not only suppresses extreme opinions in the current decision, but that it also has a long-lasting effect by changing the opinions of extreme group members, bringing them closer to the median.

5.5. Relating the empirical results to predictions of spatial bargaining models

Our experiments were not particularly designed to test the predictions of spatial bargaining models, or to compare the validity of different types of spatial bargaining models. In particular, we do not control the level of patience of group members, or the decision rule the group adopts. Posing no restriction in these dimensions implies that both a Banks and Duggan (2000) and a Compte and Jehiel (2010) type model can explain a wide range of observed behavior, especially if groups are heterogeneous in the decision rule they adopt. Nevertheless, we can identify the range of parameter values for which the models' predictions are broadly consistent with the empirical findings. In the Banks and Duggan (2000) model framework, this simply requires an intermediate level of patience for the members. For too impatient members, the group decision would be the mean individual opinion and all members would be influential, while for too patient members the group decision would be the median individual opinion (subject to an additional condition if the decision rule is not simple majority). It is for intermediate levels of patience that extreme individual opinions are not likely to be included in the acceptance set, but all non-extreme individual opinions are likely enough to influence the expected group decision.

A Compte and Jehiel (2010) type model can also generate predictions that are consistent with the empirical findings. This first of all requires that the level of patience is not too low, because in that case no member would be influential. Second, it cannot be that either all groups adopt simple majority rule or that all groups adopt unanimity rule. In the first case only the median member would be influential, while in the second case only the extreme members. However, if it is either the case that groups tend to adopt a supermajority rule but not unanimity, or that some groups adopt a simple majority rule while other groups a supermajority rule then it can be that exactly the non-extreme members are influential in expectation.[38]

[37] For lottery decisions, within-group variance of individual decisions is statistically significantly lower in phase 2 than in phase 1, and in phase 3 than in phase 2. We do not observe such decrease of within-group variance over time for gift-exchange decisions. The time needed for group deliberation (the time from when the group discussion started until the group decision was entered into the experiment program) is statistically significantly lower for phase 2 than for phase 1 in both decision contexts, but does not decrease further in phase 3. In both gift-exchange and lottery decision environments, we do not observe a significant correlation between within-group variance and group deliberation time.

[38] For supermajority rules other than unanimity, Compte and Jehiel (2010) predicts that in each group exactly two members are influential, but the identity of these two members depend on the specific individual opinions in the group. So it can be the case that for some type of individual opinion compositions the median and another member are influential, while for some other type of compositions two non-median members are influential.

Table 6
OLS regressions of current individual choices on choices made before, conditional on previous position.

	Gift-exchange 2nd decision in phase (Normalized amount returned)	Lottery (Nb safe choices)
Constant	−1.939***	0.012
	[0.406]	[0.342]
Own decision in $t-1$	0.665***	1.003***
	[0.04]	[0.052]
Was (1)	0.296*	0.312**
× Diff to group in $t-1$	[0.146]	[0.121]
Was (2)	0.063	0.16
× Diff to group in $t-1$	[0.194]	[0.103]
Was (3)	0.189	0.07
× Diff to group in $t-1$	[0.115]	[0.102]
Was (4)	0.069	−0.305
× Diff to group in $t-1$	[0.056]	[0.185]
Was (5)	0.024	0.566***
× Diff to group in $t-1$	[0.102]	[0.111]
N	260	260
R^2	0.338	0.679

Notes: *, **, and *** denote significance at the 10%, 5%, and 1%-levels, respectively. Standard errors are clustered at the experiment session level, and shown in brackets.

6. Conclusion

This paper investigates how groups come to an agreement through deliberation, and which group members have a significant influence on the group choice. We find that the member with the median opinion has the greatest influence on the choice, but there are members in both directions from the median who can also influence the choice. We observe that extreme opinions tend to be suppressed in group decision making. We also find evidence that persuasion is part of deliberation, as individual opinions tend to move towards previous group decisions the individuals participated in. This particularly holds for individuals with extreme opinions, implying that deliberation has a permanent depolarizing effect in the population. This is consistent with the idea behind deliberative democracy, in that deliberation, not merely voting, provides real authenticity to social choices.

Appendix A

Theorem 1. *In a bargaining game with publicly known ideal points and plurality decision rule, SSPE is unique. For any set of ideal points, the set of acceptable proposals A monotonically decreases in* δ, *it converges to the whole choice set as* $\delta \to 0$, *and it converges to the median ideal point as* $\delta \to 1$.

Proof. Let y_i denote member i's proposal, and define the following:

$$\lambda_1 \equiv arg \min_{i \in \{1,2,4,5\}} |x_i - x_3|$$
$$\lambda_{k+1} \equiv arg \min_{i \in \{1,2,4,5\} \setminus \{\lambda_1,\ldots,\lambda_k\}} |x_i - x_3|$$

Thus, $|x_{\lambda_1} - x_3| \le |x_{\lambda_2} - x_3| \le |x_{\lambda_3} - x_3| \le |x_{\lambda_4} - x_3|$, so that player λ_k is the non-median player whose ideal point is the k-th closest (among non-median players) to that of the median player.

First, Theorem 1 of Banks and Duggan (2000) implies that there is no delay in SSPE (along the equilibrium path all proposals are accepted).

Next, note that in any SSPE, the acceptable proposals are exactly those that are acceptable to player 3. This follows because for all players, in any SSPE, the expected continuation value if there is no agreement in the current round depends only on the expectation and the variance of the proposal in the next round. Let the expectation of the proposal in the SSPE be $\overline{y}$. Since the variance affects all players' payoffs the same way, the difference in expected continuation payoffs across players only depends on the distance between ideal points and $\bar{y}$. This means that if a proposal y is preferred by player 3 to no agreement in the current round and $y \le \bar{y}$ then players 1 and 2 also prefer it. Conversely, if a proposal y is preferred by player 3 to no agreement in the current round and $y \le \bar{y}$ then players 4 and 5 also prefer it.

Because payoffs are quasi-concave, in any SSPE, the set of proposals acceptable to player 3 is a convex interval. Moreover, this interval contains x_3, the action that yields the highest possible payoff to player 3. Given this, we need to consider the following cases:

Case (i). The acceptance set contains x_3, but does not contain any other player's ideal point.

The incentive compatibility condition to lure player 3 is

$$v \ge \delta(4v+1)/5.$$

Letting v denote player 3's payoff and setting $y = x_{\lambda_1}$ gives the bound on δ:

$$\delta > \frac{5\left(1-\left(x_{\lambda_1}-x_3\right)^2\right)}{5-4\left(x_{\lambda_1}-x_3\right)^2}.$$

Then, given the above inequality holds,

$$y_1 = y_2 = x_3-\sqrt{\frac{5(1-\delta)}{5-4\delta}},\;\; y_3 = x_3, y_4 = y_5 = x_3+\sqrt{\frac{5(1-\delta)}{5-4\delta}}.$$

This candidate equilibrium cannot exist for $\delta \le \frac{5\left(1-\left(x_{\lambda_1}-x_3\right)^2\right)}{5-4\left(x_{\lambda_1}-x_3\right)^2}$. In this case, $|x_3-x_{\lambda_1}| \le \sqrt{\frac{5(1-\delta)}{5-4\delta}}$, so that x_{λ_1} is in the acceptance set given above (making it suboptimal for player λ_1 to propose anything other than x_{λ_1}).

Case (ii). The acceptance set contains two consecutive players' ideal points, including x_3.

The incentive compatibility condition to lure player 3 is

$$v \ge \delta\left(3v+2-\left(x_{\lambda_1}-x_3\right)^2\right)/5.$$

Setting the payoff of the median voter equal to v and setting $y = x_{\lambda_2}$ gives the next bound on δ:

$$\frac{5\left(1-\left(x_{\lambda_2}-x_3\right)^2\right)}{5-\left(x_{\lambda_1}-x_3\right)^2-3\left(x_{\lambda_2}-x_3\right)^2} < \delta \le \frac{5\left(1-\left(x_{\lambda_1}-x_3\right)^2\right)}{5-4\left(x_{\lambda_1}-x_3\right)^2}.$$

Then, given the above inequality holds,

$$y_3 = x_3, \quad y_{\lambda_k} = \begin{cases} x_{\lambda_k} & \text{if } k=1 \\ x_3-\sqrt{\dfrac{5(1-\delta)+\delta\left(x_{\lambda_1}-x_3\right)^2}{5-3\delta}} & \text{if } k>1,\; \lambda_k<3 \\ x_3+\sqrt{\dfrac{5(1-\delta)+\delta\left(x_{\lambda_1}-x_3\right)^2}{5-3\delta}} & \text{if } k>1,\; \lambda_k>3 \end{cases}$$

This candidate equilibrium cannot exist for $\delta < \frac{5\left(1-\left(x_{\lambda_2}-x_3\right)^2\right)}{5-\left(x_{\lambda_1}-x_3\right)^2-3\left(x_{\lambda_2}-x_3\right)^2}$. In this case, $|x_3-x_{\lambda_2}| < \sqrt{\frac{5(1-\delta)+\delta\left(x_{\lambda_1}-x_3\right)^2}{5-3\delta}}$, so that x_{λ_2} is strictly between x_3 and the prescribed proposal given above. Consequently, player λ_2's ideal point x_{λ_2} would be accepted if proposed, so it would then be suboptimal for player λ_2 to propose anything other than x_{λ_2}.

This candidate equilibrium cannot exist for $\delta > \frac{5\left(1-\left(x_{\lambda_1}-x_3\right)^2\right)}{5-4\left(x_{\lambda_1}-x_3\right)^2}$ either. In this case, $1-\left(x_3-x_{\lambda_1}\right)^2 < \delta\left[5-\left(x_3-x_{\lambda_1}\right)^2-\sum_{i\ne\lambda_1}\left(x_3-y_i\right)^2\right]/5$, where the y_i's are the proposals given above, so that $y_{\lambda_1} = x_{\lambda_1}$ is not incentive compatible with respect to player 3. Consequently, player 3 would reject x_{λ_1} so that it would be suboptimal for player λ_1 to propose x_{λ_1}.

Case (iii). The acceptance set contains three consecutive players' ideal points, including x_3.

The incentive compatibility condition to lure player 3 is

$$v \ge \delta\left(2v+3-\left(x_{\lambda_1}-x_3\right)^2-\left(x_{\lambda_2}-x_3\right)^2\right)/5.$$

Setting the payoff of the median voter equal to v and setting $y = x_{\lambda_3}$ gives the next bound on δ:

$$\frac{5\left(1-\left(x_{\lambda_3}-x_3\right)^2\right)}{5-\left(x_{\lambda_1}-x_3\right)^2-\left(x_{\lambda_2}-x_3\right)^2-2\left(x_{\lambda_3}-x_3\right)^2} < \delta \le \frac{5\left(1-\left(x_{\lambda_2}-x_3\right)^2\right)}{5-\left(x_{\lambda_1}-x_3\right)^2-3\left(x_{\lambda_2}-x_3\right)^2}.$$

Then, given the above inequality holds,

$$y_3 = x_3, \quad y_{\lambda_k} = \begin{cases} x_{\lambda_k} & \text{if } k\le 2 \\ x_3-\sqrt{\dfrac{5(1-\delta)+\delta\left[\left(x_{\lambda_1}-x_3\right)^2,+,\left(x_{\lambda_2}-x_3\right)^2\right]}{5-2\delta}} & \text{if } k>2,\; \lambda_k<3 \\ x_3+\sqrt{\dfrac{5(1-\delta)+\delta\left[\left(x_{\lambda_1}-x_3\right)^2,+,\left(x_{\lambda_2}-x_3\right)^2\right]}{5-2\delta}} & \text{if } k>2,\; \lambda_k>3. \end{cases}$$

This candidate equilibrium cannot exist for $\delta < \frac{5\left(1-\left(x_{\lambda_3}-x_3\right)^2\right)}{5-\left(x_{\lambda_1}-x_3\right)^2-\left(x_{\lambda_2}-x_3\right)^2-2\left(x_{\lambda_3}-x_3\right)^2}$. In this case, $|x_3-x_{\lambda_3}| < \sqrt{\frac{5(1-\delta)+\delta\left[\left(x_{\lambda_1}-x_3\right)^2,+,\left(x_{\lambda_2}-x_3\right)^2\right]}{5-2\delta}}$, so that x_{λ_3} is strictly between x_3 and the prescribed proposal given above. Consequently, player λ_3's ideal point x_{λ_3} would be accepted if proposed, so it would then be suboptimal for player λ_3 to propose anything other than x_{λ_3}.

This candidate equilibrium cannot exist for $\delta > \frac{5\left(1-\left(x_{\lambda_2}-x_3\right)^2\right)}{5-\left(x_{\lambda_1}-x_3\right)^2-3\left(x_{\lambda_2}-x_3\right)^2}$ either. In this case, $1-\left(x_3-x_{\lambda_2}\right)^2 < \delta\left[5-\left(x_3-x_{\lambda_2}\right)^2-\sum_{i\ne\lambda_2}\left(x_3-y_i\right)^2\right]/5$, where the y_i's are the proposals given above, so that $y_{\lambda_2} = x_{\lambda_2}$ is not incentive compatible with respect to player 3. Consequently, player 3 would reject x_{λ_2} so that it would be suboptimal for player λ_2 to propose x_{λ_2}.

Case (iv). The acceptance set contains four consecutive players' ideal points, including x_3.

The incentive compatibility condition to lure player 3 is

$$v \ge \delta\left(v+4-\left(x_{\lambda_1}-x_3\right)^2-\left(x_{\lambda_2}-x_3\right)^2-\left(x_{\lambda_3}-x_3\right)^2\right)/5.$$

Setting the payoff of player 3 equal to v and setting $y = x_{\lambda_4}$ gives the next bound on δ:

$$\frac{5\left(1-\left(x_{\lambda_4}-x_3\right)^2\right)}{5-\sum_{i=1}^{5}\left(x_i-x_3\right)^2} < \delta \le \frac{5\left(1-\left(x_{\lambda_3}-x_3\right)^2\right)}{5-\left(x_{\lambda_1}-x_3\right)^2-\left(x_{\lambda_2}-x_3\right)^2-2\left(x_{\lambda_3}-x_3\right)^2}.$$

Then, given the above inequality holds,

$$y_3 = x_3,\ \ y_{\lambda_k} = \begin{cases} x_{\lambda_k} & \text{if } k \le 3 \\ x_3 - \sqrt{\dfrac{5(1-\delta)+\delta\left[(x_{\lambda_1}-x_3)^2, +, (x_{\lambda_2}-x_3)^2, +, (x_{\lambda_3}-x_3)^2\right]}{5-\delta}} & \text{if } k=4,\ \lambda_4 < 3 \\ x_3 + \sqrt{\dfrac{5(1-\delta)+\delta\left[(x_{\lambda_1}-x_3)^2, +, (x_{\lambda_2}-x_3)^2, +, (x_{\lambda_3}-x_3)^2\right]}{5-\delta}} & \text{if } k=4,\ \lambda_4 > 3. \end{cases}$$

This candidate equilibrium cannot exist for $\delta < \frac{5\left(1-(x_{\lambda_4}-x_3)^2\right)}{5-\sum_{i=1}^{5}(x_i-x_3)^2}$. In this case, $|x_3 - x_{\lambda_4}| < \sqrt{\frac{5(1-\delta)+\delta\left[(x_{\lambda_1}-x_3)^2,+,(x_{\lambda_2}-x_3)^2,+,(x_{\lambda_3}-x_3)^2\right]}{5-\delta}}$, so that x_{λ_4} is strictly between x_3 and the prescribed proposal given above. Consequently, player λ_4's ideal point x_{λ_4} would be accepted if proposed, so it would then be suboptimal for player λ_4 to propose anything other than x_{λ_4}.

This candidate equilibrium cannot exist for $\delta > 5\frac{\left(1-(x_{\lambda_3}-x_3)^2\right)}{5-(x_{\lambda_1}-x_3)^2-(x_{\lambda_2}-x_3)^2-2(x_{\lambda_3}-x_3)^2}$ either. In this case, $1-(x_3-x_{\lambda_3})^2 < \delta\left[5-(x_3-x_{\lambda_3})^2-\sum_{i\neq\lambda_3}(x_3-y_i)^2\right]/5$, where the y_i's are the proposals given above, so that $y_{\lambda_3} = x_{\lambda_3}$ is not incentive compatible with respect to player 3. Consequently, player 3 would reject x_{λ_3} so that it would be suboptimal for player λ_3 to propose x_{λ_3}.

Case (v). All members propose their ideal points.

Given

$$\delta \le \frac{5\left(1-(x_{\lambda_4}-x_3)^2\right)}{5-\sum_{i=1}^{5}(x_i-x_3)^2},$$

we have

$$y_i = x_i \ \text{ forall } \ i.$$

This candidate equilibrium cannot exist for $\delta > \frac{5\left(1-(x_{\lambda_4}-x_3)^2\right)}{5-\sum_{i=1}^{5}(x_i-x_3)^2}$. In this case, $1-(x_3-x_{\lambda_4})^2 < \delta\sum\left(1-(x_3-y_i)^2\right)/5$, where the y_i's are the proposals given above, so that $y_4 = x_{\lambda_4}$ is no incentive compatible to player 3. Consequently, it would be suboptimal for player λ_4 to propose x_{λ_4}.

This concludes the analysis of possible cases. Since $\lambda_1,..., \lambda_4$ are not restricted to be strictly positive, the above analysis also covers parameter values when multiple players' ideal points are at the median position. Relatedly, for some ideal point configurations, only some of the above cases apply. For example, if there are two players with ideal points exactly at the median, then the set of δ for which exactly one player proposes her ideal point is empty: for any $\delta \in (0, 1)$ at least the two median players can propose their ideal points in SSPE.

Since the regions of discount factors for which different types of equilibria (Cases (i)–(v)) apply partition (0, 1) for any $(x_1,..., x_5)$, SSPE is always unique. The above characterization of the unique SSPE also implies all of the other claims in the statement of the theorem. ■

Theorem 2. In a bargaining game with private ideal points and simple majority rule, preceded by a round of cheap talk, for low enough discount factors there is a sequential equilibrium in which members announce their types truthfully, and then all members propose their ideal points in the bargaining phase. Moreover, all sequential equilibria are outcome-equivalent to the equilibrium above. Conversely, for high enough discount factors there is a sequential equilibrium in which members announce their types truthfully, and then all members propose the median announced ideal point in the bargaining phase.

Proof. Let $a_1 < \ldots < a_N$ be the set of possible actions and player types. Define $\Delta \equiv min_i(a_{i+1} - a_i)$.

Consider first the continuation game on the equilibrium path following the reporting stage. Note that player 3's continuation payoff at any stage, in case of no agreement, is δ, as according to the candidate equilibrium profile all players propose x_3 subsequently. Therefore $\delta \ge 1 - \Delta^2$ implies that player 3 rejects any proposal other than x_3. Moreover, any proposal $y < x_3$ gets rejected by players 4 and 5, and any proposal $y > x_3$ gets rejected by players 1 and 2. Conversely, a proposal $y = x_3$gets accepted by any player i, since it yields a payoff of $1 - (x_i - x_3)^2 > 0$, which is strictly larger than the continuation payoff $\delta(1 - (x_i - x_3)^2)$. This also implies that any player i strictly prefers proposing x_3 to any other proposal.

Now we consider the stage in which players report their valuations. Suppose a player with true ideal point x_i reports $x_i^{'} \neq x_i$, and consider each of the possible cases: (a) misreporting leads other players to believe the median ideal point is not x_i, when in fact it is x_i, (b) misreporting leads other players to believe the median ideal point is not x_j, when in fact it is $x_j \neq x_i$, (c) misreporting has no effect on other players' beliefs regarding the median ideal point.

For case (a), misreporting strictly decreases player i's payoff because other players will no longer propose player i's ideal point if recognized in the first period. Therefore no matter how the game proceeds thereafter, player i is strictly worse than if other players were to propose the true median of x_i in the first period.

For case (b), other players' beliefs regarding the median ideal point can only be affected by player i misreporting if x_j is strictly between x_i and $x_i^{'}$. Therefore, the perceived median, $\tilde{x}_j$ is further from player i's ideal point than the true median: $|\tilde{x}_j - x_i| > |x_j - x_i|$. Consequently, other players (if recognized) propose $\tilde{x}_j$, which gives player i a strictly worse payoff regardless of whether or not the proposal is accepted than the payoff if other players propose (and accept) x_j, as in the truthful equilibrium.

For case (c), player i's payoff is not affected by the deviation.

Therefore, player i cannot profit from misreporting in any of the three cases, and is strictly hurt by misreporting in cases (a) and (b). Moreover, for any x_i, player i's ex-ante expectation of x_i being the true median ideal point is nonzero for a finite type space with independently distributed types. The likelihood of strictly hurting oneself by misreporting is strictly positive for any x_i. Thus, truth-telling is the unique optimal action in the reporting stage. ■

Appendix B. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jpubeco.2015.05.008.

References

Ambrus, A., Lu, S.-E., 2015. A continuous-time model of multilateral bargaining. Am. Econ. J. Microecon. 7 (1), 208–249.

Ambrus, A., Pathak, P., 2012. Cooperation over finite horizons: a theory and experiments. J. Public Econ. 95, 500–512.

Arora, N., Allenby, G., 1999. Measuring the influence of individual preference structures in group decision making. J. Mark. Res. 36, 476–487.

Baker, R.J., Laury, S.K., Williams, A.W., 2008. Comparing small-group and individual behavior in lottery-choice experiments. South. Econ. J. 75, 367–382.

Banks, J.S., Duggan, J., 2000. A bargaining model of collective choice. Am. Polit. Sci. Rev. 94, 73–88.

Baron, D.P., Ferejohn, J.A., 1989. Bargaining in legislatures. Am. Polit. Sci. Rev. 83 (4), 1181–1206.

Besedes, T., Deck, C., Quintanar, S., Sarangi, S., Shor, M., 2014. Effort and performance: what distinguishes interacting and non-interacting groups from individuals? South. Econ. J. 81, 294–322.

Bhaskar, V., Vega-Redondo, F., 2002. Asynchronous choice and Markov equilibria. J. Econ. Theory 103 (2), 334–350.

Bhaskar, V., Mailath, G., Morris, S., 2013. A foundation for Markov equilibria in sequential games with finite social memory. Rev. Econ. Stud. 80, 925–948.

Binmore, K., 1987. Perfect equilibria in bargaining models. In: Binmore, K., Dasgupta, P. (Eds.), The Economics of Bargaining. Basil Blackwell, Oxford, pp. 77–105.

Blinder, A., Morgan, J., 2005. Are two heads better than one? An experimental analysis of group versus individual decision making. J. Money, Credit, Bank. 37, 789–811.
Bornstein, G., Yaniv, I., 1998. Individual and group behavior in the ultimatum game: are groups more "rational" players? Exp. Econ. 1, 101–108.
Bornstein, G., Kugler, T., Ziegelmeyer, A., 2004. Individual and group decisions in the centipede game: are groups more "rational" players? J. Exp. Soc. Psychol. 40, 599–605.
Bower, J., 1965. The role of conflict in economic decision-making groups: some empirical results. Q. J. Econ. 79, 263–277.
Brandts, J., Charness, G., 2004. Do labour market conditions affect gift exchange? Some experimental evidence. Econ. J. 114 (497), 684–708.
Brandts, J., Schram, A., 2001. Cooperation and noise in public goods experiments: applying the contribution function approach. J. Public Econ. 79, 399–427.
Brown, R., 1974. Further comment on the risky shift. Am. Psychol. 29, 468–470.
Brown, M., Falk, A., Fehr, E., 2004. Relational contracts and the nature of market interactions. Econometrica 72, 747–780.
Burnstein, E., 1982. Persuasion as argument processing. In: Brandstätter, H., Davis, J.H., Stocker-Kreichgauer, G. (Eds.), Group Decision Making. Academic Press.
Burnstein, E., Vinokur, A., Trope, Y., 1973. Interpersonal comparisons versus persuasive argument: a more direct test of alternative explanations for group-induced shifts in individual choices. J. Exp. Soc. Psychol. 9, 236–245.
Casari, M., Zhang, J., Jackson, C., 2012. When do groups perform better than individuals? A company takeover experiment. Technical report, Working Paper. University of Bologna.
Cason, T., Mui, V., 1997. A laboratory study of group polarisation in the team dictator game. Econ. J. 107, 1465–1483.
Charness, G., 2004. Attribution and reciprocity in an experimental labor market. J. Labor Econ. 22, 665–688.
Charness, G., Jackson, M., 2009. The role of responsibility in strategic risk-taking. J. Econ. Behav. Organ. 69, 241–247.
Charness, G., Sutter, M., 2012. Groups make better self-interested decisions. J. Econ. Perspect. 26 (3), 157–176.
Charness, G., Rigotti, L., Rustichini, A., 2007. Individual behavior and group membership. Am. Econ. Rev. 97, 1340–1352.
Charness, G., Cobo-Reyes, R., Jiménez, N., Lacomba, J., Lagos, F., 2012. The hidden advantage of delegation: Pareto improvements in a gift exchange game. Am. Econ. Rev. 102, 2358–2379.
Chen, Y., Li, S.X., 2009. Group identity and social preferences. Am. Econ. Rev. 99, 431–457.
Cialdini, R.B., Goldstein, N.J., 2004. Social influence: compliance and conformity. Annu. Rev. Psychol. 55, 591–621.
Compte, O., Jehiel, P., 2010. Bargaining and majority rules: a collective search perspective. J. Polit. Econ. 118, 189–221.
Cooper, D.J., Kagel, J., 2005. Are two heads better than one? Team versus individual play in signaling games. Am. Econ. Rev. 95, 477–509.
Cooper, D.J., Rege, M., 2011. Misery loves company: social regret and social interaction effects in choices under risk and uncertainty. Games Econ. Behav. 73 (1), 91–110.
Corfman, K.P., Harlam, B.A., 1998. Relative influence of parent and child in joint purchase decisions: the role of product characteristics. Working Paper, SternSchool of Business. New York University.
Cramton, P.C., 1984. Bargaining with incomplete information: an infinite-horizon model with two-sided uncertainty. Rev. Econ. Stud. 51 (4), 579–593.
Cramton, P.C., 1992. Strategic delay in bargaining with two-sided uncertainty. Rev. Econ. Stud. 59 (1), 205–225.
Datta, S., Satten, G., 2008. A signed-rank test for clustered data. Biometrics 64 (2), 501–507.
Davis, J., 1973. Group decision and social interaction: a theory of social decision schemes. Psychol. Rev. 80, 97–125.
Elbittar, A., Gomberg, A., 2012. "My friends: It would be an error to accept": communication in a bargaining setting. Working Paper. Centro de Investigacion Economica, ITAM.
Eliaz, K., Ray, D., Razin, R., 2005. Choice shifts in groups and the Allais paradox. Am. Econ. Rev. 96, 1321–1332.
Fehr, E., Gächter, S., 1998. How effective are trust and reciprocity-based incentives? In: Ben-Ner, A., Putterman, L. (Eds.), Economics, Values, and Organization. Cambridge University Press, Cambridge, pp. 337–363.
Fehr, E., Kirchsteiger, G., Riedl, A., 1993. Does fairness prevent market clearing? An experimental investigation. Q. J. Econ. 108 (2), 437–459.
Ferguson, D.A., Vidmar, N., 1971. Effects of group discussion on estimates of culturally appropriate risk levels. J. Pers. Soc. Psychol. 20, 436–445.
Fiorina, M., Plott, C., 1978. Committee decisions under majority rule: an experimental study. Am. Polit. Sci. Rev. 72, 575–598.
Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Exp. Econ. 10 (2), 171–178.
Fischbacher, U., Gächter, S., Fehr, E., 2001. Are people conditionally cooperative? Evidence from a public goods experiment. Econ. Lett. 71, 397–404.
Fudenberg, D., Tirole, J., 1983. Sequential bargaining with incomplete information. Rev. Econ. Stud. 50 (2), 221–247.
Glaeser, E.L., Sunstein, C.R., 2009. Extremism and social learning. J. Legal Anal. 1, 263–324.
Goeree, J., Yariv, L., 2011. An experimental study of collective deliberation. Econometrica 79, 893–921.
Holt, C.A., Laury, S.K., 2002. Risk aversion and incentive effects. Am. Econ. Rev. 92, 1644–1655.
Hong, L.K., 1978. Risky shift and cautious shift: some direct evidence on the culture-value theory. Soc. Psychol. 41, 342–346.
Insko, C.A., Schopler, J., Hoyle, R.H., Dardis, G.J., Graetz, K.A., 1990. Individual–group discontinuity as a function of fear and greed. J. Pers. Soc. Psychol. 58, 68–79.
Kamenica, E., Gentzkow, M., 2011. Bayesian persuasion. Am. Econ. Rev. 101, 2590–2615.
Kerr, N.L., MacCoun, R.J., Kramer, G.P., 1996. Bias in judgement: comparing individuals and groups. Psychol. Rev. 103, 687–719.
Kocher, M.G., Sutter, M., 2005. The decision-maker matters: individual versus group behaviour in experimental beauty-contest games. Econ. J. 115, 200–223.
Kocher, M.G., Sutter, M., 2007. Individual versus group behavior and the role of the decision making process in gift-exchange experiments. Empirica 34, 63–88.
Kramer, R.M., 1991. Intergroup relations and organizational dilemmas: the role of categorization processes. In: Cummings, L.L., Staw, B.M. (Eds.), Research in Organizational Behavior 13. JAI Press, Greenwich, CT.
Laughlin, P.R., 1980. Social combination process of cooperative problem-solving groups at verbal intellective tasks. In: Fishbein, M. (Ed.), Progress in Social Psychology. Erlbaum, Hillsdale, pp. 127–155.
Laughlin, P.R., Ellis, A.L., 1986. Demonstrability and social combination processes on mathematical intellective tasks. J. Exp. Soc. Psychol. 22, 170–189.
Ledyard, J., 1995. Public goods. In: Kagel, J., Roth, A.E. (Eds.), Handbook of Experimental Economic. Princeton University Press, Princeton.
Levinger, G., Schneider, D.J., 1969. Test of the "risk is value" hypothesis. J. Pers. Soc. Psychol. 11, 165–169.
Lindskold, S., McElwain, D.C., Wayner, M., 1977. Cooperation and the use of coercion by groups and by individuals. J. Confl. Resolut. 21, 531–550.
Luhan, W., Kocher, M., Sutter, M., 2009. Group polarization in the team dictator game reconsidered. Exp. Econ. 12, 26–41.
Manin, B., 2005. Deliberation: why we should focus on debate rather than discussion. Paper delivered at the Program in Ethics and Public Affairs Seminar. Princeton University (13 October 2005).
Masclet, D., Colombier, N., Denant-Boemont, L., Loheac, Y., 2009. Group and individual risk preferences: a lottery-choice experiment with self-employed and salaried workers. J. Econ. Behav. Organ. 70, 470–484.
Moulin, H., 1980. On strategy-proofness and single peakedness. Public Choice 35, 437–455.
Myers, D.G., Arenson, S.J., 1972. Enhancement of dominant risk tendencies in group discussion. Psychol. Rep. 30, 615–623.
Palfrey, T., Prisbey, J., 1997. Anomalous behavior in linear public goods experiments: how much and why? Am. Econ. Rev. 87, 829–846.
Pylyshyn, Z., Agnew, N., Illingworth, J., 1966. Comparison of individuals and pairs as participants in a mixed-motive game. J. Confl. Resolut. 10, 211–220.
Rabbie, J.M., 1982. The effects of intergroup competition and cooperation on intragroup and intergroup relations. In: Derlega, V.J., Grzelak, J. (Eds.), Cooperation and Helping Behavior: Theories and Research. Academic Press, New York, pp. 123–149.
Robert, C., Carnevale, P.J., 1997. Group choice in ultimatum bargaining. Organ. Behav. Hum. Decis. Process. 72, 256–279.
Rubinstein, A., 1982. Perfect equilibrium in a bargaining model. Econometrica 50 (1), 97–109.
Schopler, J., Insko, C.A., 1992. The discontinuity effect in personal and intergroup relations: generality and mediation. In: Stroebe, W., Hewstone, M. (Eds.), European review of social psychology Vol. 3. Wiley, Chichester, England, pp. 121–151.
Shupp, R.S., Williams, A., 2008. Risk preference differentials of small groups and individuals. Econ. J. 118, 258–283.
Stoner, J. A. F. 1961 , 'A comparison of individual and group decisions under risk', Unpublished thesis, MIT School of Management.
Sunstein, C.R., 2000. Deliberative trouble? Why groups go to extremes. Yale Law J. 110, 71–119.
Sunstein, C.R., 2002. The law of group polarization. J. Polit. Philos. 10, 175–195.
Sutter, M., 2009. Individual behavior and group membership: comment. Am. Econ. Rev. 99, 2247–2257.
Tajfel, H., Billig, M.G., Bundy, R.F., Flament, C., 1971. Social categorization and intergroup behavior. Eur. J. Soc. Psychol. 1, 149–177.
Teger, A.I., Pruitt, D.G., 1967. Components of group risk taking. J. Exp. Soc. Psychol. 3, 189–205.
Wallach, M.A., Kogan, N., Bem, D.J., 1962. Group influence on individual risk taking. J. Abnorm. Soc. Psychol. 65, 75–86.
Wallach, M.A., Kogan, N., Bem, D.J., 1964. Diffusion of responsibility and level of risk taking in groups. J. Abnorm. Soc. Psychol. 68, 263–274.
Wildschut, T., Pinter, B., Vevea, J.L., Insko, C.A., Schopler, J., 2003. Beyond the group mind: a quantitative review of the interindividual–intergroup discontinuity effect. Psychol. Bull. 129, 698–722.
Wolosin, R.J., Sherman, S.J., Maynatt, C.R., 1975. When self-interest and altruism conflict. J. Pers. Soc. Psychol. 32, 752–760.
Zhang, J., Casari, M., 2012. How groups reach agreements in risky choices: an experiment. Econ. Inq. 50 (2), 502–515.

J Risk Uncertain (2016) 52:99–116
DOI 10.1007/s11166-016-9237-8

Group decision rules and group rationality under risk

Aurélien Baillon[1] · Han Bleichrodt[1] · Ning Liu[1] ·
Peter P. Wakker[1]

Published online: 15 June 2016

Abstract This paper investigates the rationality of group decisions versus individual decisions under risk. We study two group decision rules, majority and unanimity, in stochastic dominance and Allais paradox tasks. We distinguish communication effects (the effects of group discussions and interactions) from aggregation effects (mere impact of the voting procedure), which makes it possible to better understand the complex dynamics of group decision making. In an experiment, both effects occurred for intellective tasks whereas there were only aggregation effects in judgmental tasks. Communication effects always led to more rational choices; aggregation effects did so sometimes but not always. Groups violated stochastic dominance less often than individuals did, which was due to both aggregation and communication effects. In the Allais paradox tasks, there were almost no communication effects, and aggregation effects made groups deviate more from expected utility than individuals.

Keywords Group decisions under risk · Unanimity rule · Majority rule · Allais paradox

JEL Classification D81 · C91

Many economic decisions—e.g. family financial planning, corporate strategies, national laws—are made by groups. The literature comparing individual and group decision making is rich (Kugler et al. 2012). Groups have been found to attenuate, amplify, or replicate the biases found for individual decisions (Kerr et al. 1996), and these diverse findings highlight the closing remark of a recent review paper: "Ultimately, the goal of

Electronic supplementary material The online version of this article (doi:10.1007/s11166-016-9237-8) contains supplementary material, which is available to authorized users.

✉ Peter P. Wakker
Wakker@ese.eur.nl

[1] Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, The Netherlands

comparing individual and group decision making is to identify the contexts and types of decisions where each is likely to work best" (Charness and Sutter 2012, p. 174).

This paper compares group and individual decision making for three tasks: a test of non-transparent stochastic dominance, the Allais common consequence paradox, and the Allais common ratio paradox. Stochastic dominance is an objective, generally accepted criterion of rationality but recognizing it may require intellectual effort. We will therefore refer to this type of task as *intellective*. Previous research has shown that groups violate stochastic dominance less frequently than individuals (Charness et al. 2007), suggesting that group decisions improve rationality. Expected utility (EU) violations are to a large extent due to motivational or subjective aspects of human behavior. We will, therefore, call the Allais tasks *judgmental*. In this, we follow the psychological literature on group decisions (e.g. Laughlin and Ellis 1986; Kerr and Tindale 2011), where problems with a demonstrably correct answer are called intellective and tasks in which one cannot objectively defend one's preferred alternative as correct (e.g. aesthetic judgments or matters of personal taste) are called judgmental. In Allais' common consequence and common ratio paradoxes (Allais 1953), previous research has found that groups violate EU as often as individuals do (Rockenbach et al. 2007; Bone et al. 1999; Bateman and Munro 2005).

We distinguish two components in group decision making: *aggregation* and *communication*. Aggregation refers to the direct effect of the procedure on the group decision, without involving any change of an individual or any communication. In single choice tasks, aggregation effects tend to amplify the patterns exhibited by the majority, as we will show; these effects are merely procedural and statistical. For tasks involving two choices (as in the Allais paradoxes), however, aggregation effects can reverse the majority pattern. It then is, for example, possible that, while a majority of the group members behave according to EU, the group decision violates EU. Aggregation effects are fleeting and do not influence individual attitudes or subsequent individual decisions.

Communication effects do influence individuals. Communication effects are lasting and capture the effects of group decision making beyond pure aggregation. These effects include the impact of learning what others prefer and why they prefer it. For instance, group decision making can foster discussion and this may change members' preferences. Communication effects will persist after the group process has ended (Maciejovsky et al. 2013).

In our experiment, we measured the strength of aggregation and communication effects for majority and unanimity group decisions. We elicited individual preferences before and after the group decision stage. Changes in individual preferences were interpreted as communication effects. We also used a second approach to measure communication effects. To control for the effect of aggregation, we aggregated the individual decisions before the group decision stage into simulated group decisions. Differences between these simulated decisions and the actual group decisions constitute communication effects. For instance, in the actual group decisions a minority sometimes convinced a majority to change their preferences, which was clearly a communication effect.

The effects of decision rules are central in political economics (Feddersen and Pesendorfer 1998; Messner and Polborn 2004), and have also been studied in psychology (Kerr and Tindale 2004). However, they have rarely been considered for decisions under risk. An exception is Brunette et al. (2015), who implemented majority and unanimity voting rules, but did not permit verbal communication between group members. Most existing studies considered groups of two individuals, for which majority and unanimity rules make the same predictions and therefore cannot be distinguished. We used groups of three individuals, for which these rules can be distinguished. In intellective tasks, the results indicated that both rules led to aggregation and communication effects, which improved both group rationality and individual rationality. In the judgmental tasks there were mainly aggregation effects, which led to more deviations from expected utility for groups than for individuals in one of the tasks (the common ratio task).

1 Literature

According to social-decision scheme theory (Davis 1973), group decisions are determined by the majority view unless the minority view is demonstrably correct, which is the case where "truth wins" (Laughlin and Ellis 1986; Davis 1992; Kerr and Tindale 2011). This suggests that communication effects will be stronger when the reasons for choosing one option are easy to defend. Because answers in intellective tasks are easier to defend than answers in judgmental tasks, we expect that communication effects will be stronger in the non-transparent stochastic dominance (NTSD) tasks than in the Allais tasks.

The literature on group decision processes shows that the unanimity rule typically leads to more communication than the majority rule. Moreover, group members identify themselves more with unanimous decisions than with majority decisions (Kameda 1991). The deliberations of unanimity groups show more conflicts and more changes of opinion (Nemeth 1977). Accordingly, we hypothesize that the unanimity rule leads to more discussion than the majority rule and, consequently, to stronger communication effects.

Several studies have investigated the effect of different aggregation rules on the outcome of the decision process, both theoretically and experimentally. Dasgupta and Maskin (2008) showed theoretically that the majority rule satisfies five appealing conditions (Pareto optimality, anonymity, neutrality, independence of irrelevant alternatives, and decisiveness) over a larger class of preferences than any other voting rule does and, in this sense, is the most robust rule.

For jury decision making, Feddersen and Pesendorfer (1998) showed theoretically that the unanimity rule is more likely to convict the innocent and to acquit the guilty if jury members vote strategically. However, Goeree and Yariv (2011) found in an experiment that such differences in jury decisions vanish when deliberation before voting is allowed. Miller (1985) and Ohtsubo et al. (2004) showed experimentally that majority group decisions tend to ignore the minority's preferences whereas the unanimity decisions incorporate the minority's preferences when information about each group member's preference is available. Finally, Stasson et al. (1991) showed that subjects from majority groups performed marginally better than subjects from unanimity groups in mathematical tasks. However, groups consisting only of members who had answered incorrectly in the individual tasks were more likely to find the correct solution under the unanimity rule.

2 Examples of aggregation effects

To clarify the role of aggregation and to show how it can distort group decisions, we give two simple examples, one involving a single choice (as in the NTSD task) and one involving two choices (as in the Allais paradoxes). A similar aggregation effect was discussed by Evans (1997) in a market context.

Example 1: *Aggregation effects in a single choice*

Consider a choice between two lotteries where one lottery stochastically dominates the other. Assume that a minority of the population violates stochastic dominance. If people in this population form groups and the group choice is made by the majority rule, then sufficiently large groups will almost always satisfy stochastic dominance. Aggregation effects amplify the majority choice, and provided the majority is rational, will amplify rationality. This results purely from the aggregation procedure, without involving any change of any person or any communication.

Example 2: *Aggregation effects in a pair of choices*

Aggregation effects are more complex when two choices have to be made as in the Allais paradoxes. Consider two choices, each between a risky lottery R and a safe lottery S. Then there are four possible choice patterns: SS, RR, SR, and RS. The SS and RR patterns are consistent with EU (EU-consistent), and the SR and RS patterns are EU-inconsistent.

Suppose that 30% of the sample chooses SS, 30% chooses RR, and 40% chooses SR. Then the majority of the population (60%) is EU-consistent. If we randomly draw groups of three persons from this population and let the group choice be determined by majority then SR will be chosen in 56.8% of the cases and the majority thus exhibits an EU-inconsistent pattern. The opposite case (a majority of EU-inconsistent people leading to a majority of EU-consistent groups) can also occur.[1] Again, this results purely from the aggregation procedure.

3 Method

3.1 Experimental design

$N = 156$ students of Erasmus University Rotterdam participated in the experiment (63% males). We organized 52 sessions of three subjects each.

The experiment consisted of three stages (within-subjects) and three treatments (between-subjects; see Table 1). A stage consisted of a set of instructions and decision tasks, which subjects received on paper (called answer sheets hereafter). All decision tasks were choices between two lotteries. The subjects in a session each had their own colored pen (blue, red, or black) to mark their preferred lottery on the answer sheet so that we could identify individual choices.

[1] If the individuals exhibit patterns SS, RR, SR, and RS with probability 40%, 0%, 30%, and 30% respectively, then the groups will exhibit the same patterns with probability 56.8%, 0%, 21.6%, and 21.6%.

Table 1 Structure of the experiment

	Stage 1	Stage 2	Stage 3
Control	Individual	Individual	Individual
M-treatment	Individual	***Majority rule***	Individual
U-treatment	Individual	***Unanimity rule***	Individual

In the first and third stages of the experiment, subjects made individual choices and were not allowed to talk with each other. Only the second stage of the experiment differed between treatments. In the *control treatment*, subjects made individual decisions (without communication), as in the other two stages. This permitted us to check for any learning effect. In the *M-treatment*, the three subjects in a session made group decisions using the majority rule and they could communicate face-to-face in English. All subjects, after (possible) discussion, marked their preferred lotteries. The option with at least two ticks was taken as the group choice, but minority subjects could still express disagreement. In the *U-treatment* subjects had to reach unanimity. All three subjects of a session had to tick the same lottery, otherwise the choice was invalid and would not be paid (see the section on incentives below). The groups of the U-treatment always reached unanimity. Subjects could freely and directly communicate during the group decisions (stage 2 of M- and U-treatment). In the control treatment, answering all questions took 21 minutes on average. In the other two treatments, subjects needed about 7 minutes more due to group deliberations.

Stage 1 gave us information on subjects' behavior before any treatment manipulation. We used these choices to simulate the group decisions that would have been made based on the majority rule and, thereby, we inferred the effect of pure aggregation without any communication. We could then compare the actual group decisions in stage 2 with these simulated decisions to measure the effects beyond pure aggregation. Finally, stage 3 gave information about whether any changes in preferences that we observed in stage 2 were temporary (and thus probably due to aggregation effects) or lasting (which is likely for communication effects).

Each session of three subjects was randomly assigned to one treatment. We ran 13 sessions for the control treatment, 21 sessions for the M-treatment, and 18 sessions for the U-treatment. There were fewer sessions of the control treatments to have more observations in M- and U-treatments, which concerned the main research questions.

3.2 Stimuli

Subjects faced three types of decision tasks in each stage of the experiment: NTSD tasks, common ratio (CR) tasks, and common consequence (CC) tasks. We present these tasks below. The order of the choices was randomized within each of the three stages (it was therefore possible that two questions for the same CR task were far apart in the experiment), but the order was the same for all subjects within a particular treatment. Huber et al. (2008) showed that choices are affected by the first and the last choice made. No distortions result when there are no more than three choices, as in our experiment.

Subjects had to choose between two lotteries, A and B. Table 2 gives an example of the way choices were displayed. Uncertainty was resolved using the roll of a 20-sided

Table 2 An NTSD task

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
Lottery A	€0												€35				€30	€5			
Lottery B	€5	€10		€35					€0												

die. The first row of Table 2 shows the possible outcomes of the roll of the die. Rows 2 and 3 show the payoffs of the two lotteries under each possible roll of the die. Subjects indicated their choice by ticking the cell of the last column of their preferred lottery.

3.2.1 NTSD tasks

Charness et al. (2007) found that groups are more likely to satisfy stochastic dominance than individuals, but did not specify the aggregation rule that groups should follow and did not check whether the group effects persisted in later individual choices. Maciejovsky et al. (2013) found that group effects persisted in intellective tasks other than NTSD, but they did not distinguish between the majority rule and the unanimity rule as their group size was two people, for which the two rules lead to the same results. We adapted the NTSD tasks first used by Tversky and Kahneman (1986).

Table 2 is an example of an NTSD task faced by the subjects. Lottery B results from lottery A by improving €30 into €35 and by improving two outcomes €5 into €10. Consequently, lottery B stochastically dominates lottery A. However, this dominance is not immediately obvious and a subject who neglects probability differences and only compares the outcomes of the lotteries may have the false impression that A is better than B because it yields money amounts €5, €30, and €35, whereas Lottery B yields money amounts €5, €10, and €35. Because the violation of stochastic dominance in Table 2 is non-transparent, we refer to these tasks as NTSD. Some theories allow for such violations of stochastic dominance (Kahneman and Tversky 1979; Viscusi 1989). Table 3 summarizes the tests of non-transparent stochastic dominance that we carried out. The tests are comparable in structure but differ across the three stages.

Table 3 NTSD questions

	Dominated lottery				Dominant lottery			
	20%	5%	15%	60%	25%	10%	5%	60%
Stage 1	35	30	5	0	5	10	35	0
	40	30	8	0	8	12	40	0
	40	35	5	0	5	10	40	0
	45	40	10	0	10	15	45	0
	35	28	5	0	5	10	35	0
Stage 2	35	32	8	0	8	12	35	0
	35	30	8	0	8	12	35	0
	45	35	5	0	5	10	45	0
	40	35	10	0	10	15	40	0
	45	28	8	0	8	12	45	0
Stage 3	40	28	5	0	5	12	40	0
	45	30	8	0	8	15	45	0
	45	25	10	0	10	15	45	0
	40	25	5	0	5	10	40	0
	35	30	8	0	8	12	35	0

The second row indicates the probabilities of the outcomes. The numbers below the probabilities indicate payoffs in €

Table 4 A CC task

a) First choice

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
Lottery *S*	€***s***																				
Lottery *R*	€0				€***s***												€***r***				

b) Second choice

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
Lottery *S*	€***s***								€0												
Lottery *R*	€0																€***r***				

Actual payoffs *s* and *r* are in Table 5. In the experiment, lotteries were presented as Lottery A and Lottery B

Table 5 Payoffs and common probability of the CC tasks

	Stage 1				Stage 2			Stage 3		
s	8	8	9	8	10	9	10	10	8	10
r	25	25	30	20	45	30	45	40	20	40
q	60%	55%	55%	55%	60%	60%	55%	60%	60%	55%

The first column means that in Table 4, we used $s = €8$, $r = €25$, and $q = 60\%$. The first choice was therefore between a lottery S yielding €8 for sure and a lottery R yielding €25 with probability 20%, €8 with probability 60% and €0 otherwise. The second choice involved lottery S, yielding €8 with probability 40% (=$100\% - q$) and nothing otherwise, and lottery R yielding €25 with probability 20% and nothing otherwise

3.2.2 CC tasks

The CC tasks were close to those of Huck and Müller (2012) with a small adjustment to fit into our 20-sided die format.

Table 4 shows the structure of the CC tasks. Table 5 presents the different values of *s*, *r*, and *q* that we used in the experiment.[2] Each CC task involved two choices between lottery S and lottery R with the second choice similar to the first except that the (common) *q* chance (60% in the above example) of €s was changed into a *q* chance of nothing.

According to the sure-thing principle of EU (Savage 1954), replacing a common payoff by another common payoff does not affect preference. Hence EU predicts the same preference in both choices, either SS or RR. However, empirically, many subjects violate EU and display the pattern SR, which can be due to the *certainty effect* (see Starmer 2000 for a survey). The opposite pattern, RS, occurs less frequently (Starmer 1992; Wu et al. 2005; Blavatskyy 2013).

3.2.3 CR tasks

The CR questions were adapted from Loomes (1988), using the design of van de Kuilen and Wakker (2006).

Table 6 shows the way the CR choices were presented to the subjects. The second choice follows from the first by dividing all probabilities at nonzero outcomes by 4. According to the independence condition of EU (Fishburn and Wakker 1995), this should not affect preferences. Hence EU predicts the same choice in both situations, either SS or RR. However, empirically, the prevailing pattern is SR, which violates EU. The opposite violation, RS, occurs rarely. Table 7 presents the values used.

[2] Stage 1 had four CC tasks and stages 2 and 3 had three CC tasks.

Table 6 A CR task

a) First choice

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
Lottery S	€*s*																				
Lottery R	€*r*																€0				

b) Second choice

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
Lottery S	€*s*					€0															
Lottery R	€0																€*r*				

Actual payoffs s and r are in Table 7. In the experiment, lotteries were presented as Lottery A and Lottery B

Table 7 Payoffs of the CR tasks

	Stage 1					Stage 2					Stage 3				
s	18	9	14	10.5	20	16.5	15	9.5	11.5	12.5	13	10	15.5	17.5	19
r	24.5	12	18	14.5	25.5	22.5	20.5	13	15	17.5	18	14	21.5	24	25

3.2.4 Incentives

Upon completion of the experiment, subjects received a €5 participation fee. In addition, they played out one of their choices for real. First, one of the three stages was randomly selected. All subjects in a session played out a choice that was made in this stage. If the selected stage was an individual stage (stages 1 and 3 for the M- and U-treatments, and all stages for the control treatment), each subject randomly drew one choice and played out the lottery he preferred in that choice.

If the second stage was selected in the M- and U-treatments, the lottery the group had chosen was played out and the three subjects received the same payoff. In the group decision making stage, subjects shared the consequences of their choice. We explained this group incentive procedure in the M- and U-treatments at the beginning of stage 2. The instructions were the same across treatments, except for stage 2.

3.3 Analysis

In the individual decisions (all three stages of the control and stage 1 and 3 of the M- and U-treatments), we obviously used subjects' reported choices. In the group decisions (stage 2 for the M- and U-treatments), we assigned the choices of the subject's group to each subject. For the M-treatment in stage 2, we also had the information about the choices reported by each subject, because they were allowed to express disagreement with the group choice. Thirty-one percent of the majority group decisions revealed disagreement. Our conclusions did not depend on whether we used the group choices or the reported choices in the second stage of the M-treatment (see Online Appendix A).

We used probit models to study the likelihood of violations of stochastic dominance and multinomial probit models to study the likelihood of the four CC and CR patterns. Logit models, reported in Online Appendix B, gave similar results. We used clustered standard errors in the probit regressions. For the NTSD task, the dependent (binary) variable indicated whether or not the choice satisfied stochastic dominance. For the CC and the CR tasks, the dependent (categorical) variable described the choice pattern (SS, SR, RS, or RR).

We used the dummy variables *majority*, *unanimity*, *stage2*, and *stage3* to code stages and treatments. The main effects *stage2* and *stage3* captured the effect of learning. Our main independent variables represented the interaction terms of the treatments with the stages. The terms *majority*stage2* and *unanimity*stage2* identified the full effect of group decision making (both aggregation and communication effects). The terms *majority*stage3* and *unanimity*stage3* captured the changes in preferences after the group stage and provided a first method to estimate communication effects.

As explained before, the second method to disentangle aggregation and communication effects used the individual choices (of stage 1 for the M- and U-treatments and stages 1 and 2 for the control treatment) to simulate aggregate choices. For the majority rule, as each session consisted of three subjects, we could in each case simulate the majority group choice. These simulated group choices captured the pure aggregation effects. The difference between the simulated and the actual group choices provided an estimate of communication effects. The majority preference is also the most plausible benchmark for the U-treatment. A deviation from the majority choice then can only occur if some group members were willing to change their minds, and there must then have been communication effects. The analysis of the simulated and the group choices was done the same way as the analysis of the individual and the group choices, which was described above, except that we excluded stage 3 (which would give no useful insights) and thus used fewer data points.

Finally, we also recorded and studied the group discussions. In nearly 40% of the groups, arguments based on stochastic dominance were used (without using this term though). The dummy variable *dominance_discussion* identified the members of such groups.

4 Results

4.1 NTSD tasks

Figure 1 displays the proportion of choices satisfying stochastic dominance, for the actual choices (panel A) and when individual choices are replaced by the simulated choices (panel B). Panel A shows that only a minority of subjects chose the dominant lottery in the first stage. In the control treatment, the proportion of choices satisfying stochastic dominance increased from around 40% in stage 1 to around 50% in stage 2 to slightly below 60% in stage 3, suggesting modest learning. In the M- and U-treatments, we see a faster increase in the second stage, due to aggregation and communication effects. The comparison of the group choices with the simulated choices (panel B) confirms that even though subjects of all treatments better identify dominant lotteries as the experiment progresses, this improvement is faster in the M- and U-treatments than in the control treatment.

Table 8 reports the results of the probit regressions. The first column, which reports the results for the actual choices, shows that the use of stochastic dominance arguments in the group discussion was crucial. Groups in which no stochastic dominance arguments were used did not differ from the control treatment. However, there were strong group effects in stages 2 and 3 for groups that did talk about stochastic dominance (captured by the three-way interaction terms between stages, treatments and *dominance_discussion*). Hence, a subject who used stochastic dominance arguments

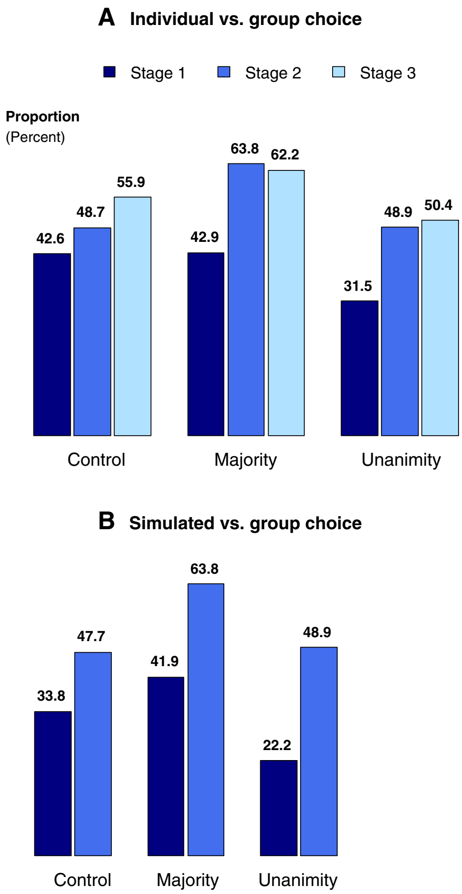


Fig. 1 Proportion of choices satisfying stochastic dominance across treatments and stages

strongly affected not only his group's choices in stage 2 but also the individual choices of his fellow group members in stage 3, indicating communication effects. In the M-treatment, the significant interaction term *dominance_discussion**majority* captures the difference in satisfaction of stochastic dominance in stage 1 between subjects who talked about stochastic dominance in stage 2 and those who did not.

The comparison of group choices with simulated choices confirmed that communication effects played a role in both the M-treatment and the U-treatment. Once aggregation effects were ruled out, groups that talked about stochastic dominance were still more likely to choose the dominant lotteries in all stages than groups that did not talk about stochastic dominance. The three-way interactions between *majority*, *dominance_discussion*, and *stage2* and between *unanimity*, *dominance_discussion*, and *stage2* were significant, showing that the choices of groups that talked about stochastic dominance were affected by communication effects beyond pure aggregation effects. Majority and unanimity seemed to have fostered discussions, which increased subjects' rationality. The importance of avoiding biases and increasing rationality has often been emphasized (Viscusi 1995, p. 108) and our analysis shows that group decision making can contribute here.

Table 8 Probit regressions for the NTSD tasks

	Group vs. individual choices		Group vs. simulated choices	
stage2	.06	(.04)	.14**	(.07)
stage3	.13**	(.06)		
majority	−.09	(.09)	−.05	(.17)
unanimity	−.17*	(.10)	−.17	(.17)
stage2*majority	.02	(.07)	.01	(.12)
stage2*unanimity	−.16	(.10)	−.14	(.14)
stage3*majority	−.02	(.08)		
stage3*unanimity	−.11	(.09)		
majority*dominance_discussion	.24**	(.11)	.34*	(.18)
unanimity*dominance_discussion	.12	(.12)	.08	(.17)
stage2*majority*dominance_discussion	.44***	(.14)	.50**	(.24)
stage2*unanimity*dominance_discussion	.84***	(.15)	.89***	(.21)
stage3*majority*dominance_discussion	.34***	(.13)		
stage3*unanimity*dominance_discussion	.47***	(.14)		
No. of observations	2340		520	
Wald chi2	149.69		74.69	
p-value	0.00		0.00	

Reported numbers are the marginal effects at the means of covariates, followed by significance and clustered standard errors between brackets. The standard errors in the left column are clustered at the individual level,[a] and those in the right column are clustered at the group level

*significant at 10% (two-sided test)

**significant at 5% (two-sided test)

***significant at 1% (two-sided test)

[a] Clustering at the group level did not change the results (see Online Appendix Table A1)

4.2 CC tasks

Neither the majority nor the unanimity group rule led to more EU-consistent choices in the common consequence Allais paradoxes (Fig. 2). In the first stage, half of our subjects' choices were consistent with EU (32% of the choice patterns were RR and 18% were SS). Surprisingly, the violations of EU were mainly of the RS type: 45% of the patterns were RS and only 5% of the empirically common SR. This finding does not confirm the certainty effect. The low occurrence of the SR pattern led to estimation problems and we excluded it from the analysis reported below. Keeping the SR pattern whenever possible gave similar results as those reported next (see Online Appendix C).

The proportion of EU violations slightly dropped in the second stage, but this held for all treatments and it disappeared in the third stage. The control and the M-treatment are very similar. In the U-treatment, there was a more pronounced increase in the proportion of RR choices, but, again, the effect did not last. Several studies have found that risk aversion is negatively related to cognitive ability (Frederick 2005; Dohmen et al. 2010, 2011). Hence the unanimity rule may also lead to more rationality in the CC task. This is

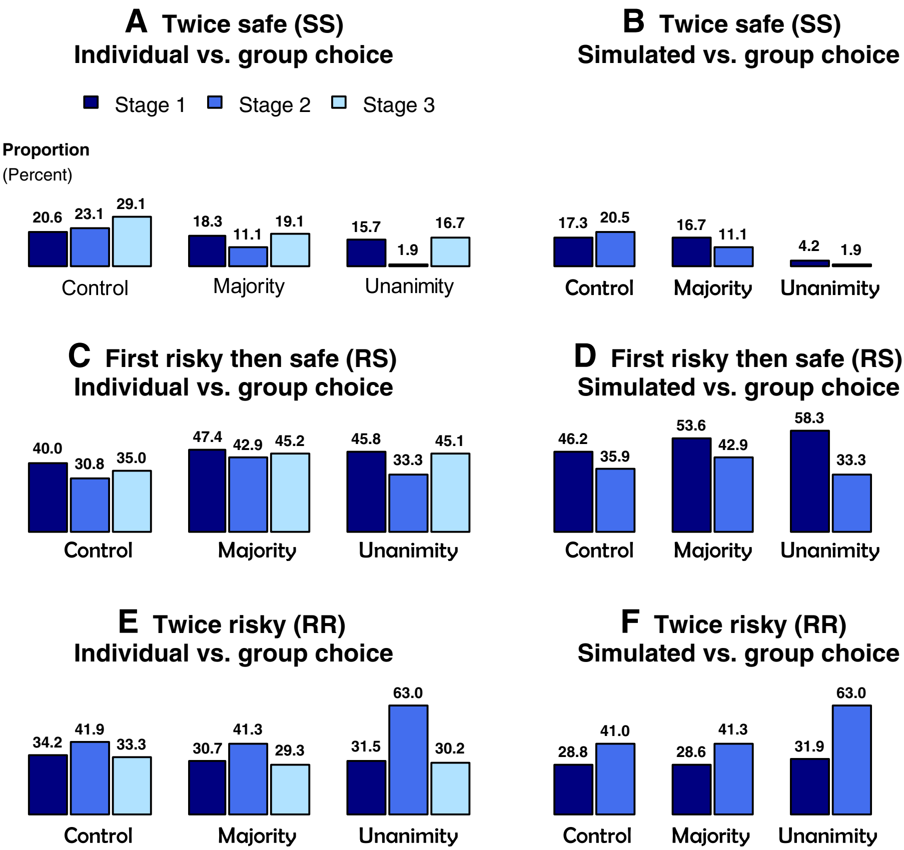


Fig. 2 Proportion of choices in the Allais common consequence tasks per treatment and stage

consistent with Keck et al.'s (2014) finding that groups are closer to EU (ambiguity neutrality) than individuals. However, the effect did not last.

The literature that compares risk attitudes between groups and individuals gives mixed results. Rockenbach et al. (2007) and Viscusi et al. (2011) found that groups were more risk seeking than individuals. On the other hand, Masclet et al. (2009) found that groups were more risk averse than individuals. Adams and Ferreira (2010) found that groups made decisions closer to risk neutrality than individuals do.

Table 9 shows the results of the multinomial probit regressions. Subjects became more risk averse during the experiment as indicated by the positive coefficient of stage 3 in the SS choices. Groups were more risk seeking than individuals: in stage 2, the prevalence of the SS pattern was less in the M- and U-treatments than in the control treatment. The RR pattern was also more common in the U-treatment than in the control group in stage 2. However, the two-way interactions in stage 3 (*stage3*majority* and *stage3*unanimity*) were not significant and there was no evidence of communication effects. When comparing group choices with simulated choices, the treatment effects on the

Table 9 Multinomial probit regressions for the CC tasks

	Group vs. individual choices			Group vs. simulated choices		
Choice pattern	SS	RR	RS	SS	RR	RS
stage2	.02	.08	−.10*	.02	.12	−.13
	(.02)	(.05)	(.05)	(.04)	(.10)	(.11)
stage3	.06**	−.01	−.06			
	(.03)	(.06)	(.06)			
majority	−.02	−.04	.07	−.01	−.03	.04
	(.06)	(.07)	(.07)	(.08)	(.12)	(.12)
unanimity	−.04	−.02	.07	−.15	.03	.12
	(.06)	(.07)	(.07)	(.10)	(.13)	(.12)
*stage2*majority*	−.10**	.03	.07	−.06	.02	.04
	(.05)	(.07)	(.08)	(.06)	(.13)	(.14)
*stage2*unanimity*	−.29***	.25***	.04	−.08	.18	−.10
	(.07)	(.07)	(.08)	(.05)	(.12)	(.13)
*stage3*majority*	−.05	.00	.05			
	(.04)	(.07)	(.07)			
*stage3*unanimity*	−.05	−.00	.05			
	(.06)	(.08)	(.08)			
No. of observations		1481			350	
Wald chi2		86.83			32.94	
p-value		0.00			0.00	

Reported numbers are the marginal effects at the means of covariates, followed by significance and clustered standard errors between brackets. The standard errors in the left three columns are clustered at the individual level, and those in the right three columns are clustered at the group level. Missing choices and SR patterns are excluded

* significant at 10% (two-sided test)

** significant at 5% (two-sided test)

*** significant at 1% (two-sided test)

SS and RR patterns at stage 2 disappeared, confirming the absence of communication effects. The tendency to make riskier group decisions was thus mainly due to aggregation effects.

4.3 CR tasks

In the CR tasks (Fig. 3), we again found that about half of the subjects satisfied EU in the first stage (29% SS pattern, 22% RR pattern). This time we found the usual SR violation of EU (41%). We do not report the results about the infrequent RS pattern, because it led to estimation problems in one of the regressions. Keeping the RS pattern did not affect the results of the other regressions as can be seen in Online Appendix C.

As in the CC tasks, group choices tended to be more risk seeking than individual choices. However, this time the shift was predominantly from SS to the EU-inconsistent SR pattern and less so to the EU-consistent RR pattern.

The multinomial probit regressions in Table 10 confirmed that the M- and U-treatments decreased the number of SS patterns and increased the number of SR patterns. These effects did not last in stage 3, suggesting that they were aggregation effects. This suggestion was confirmed by the comparison between the group choices and the simulated choices, in which these effects disappeared: the prevalence of the SS and SR patterns in the actual group choices (stage 2 of the M- and U-treatments) did not differ from the prevalence of these patterns in the simulated choices.

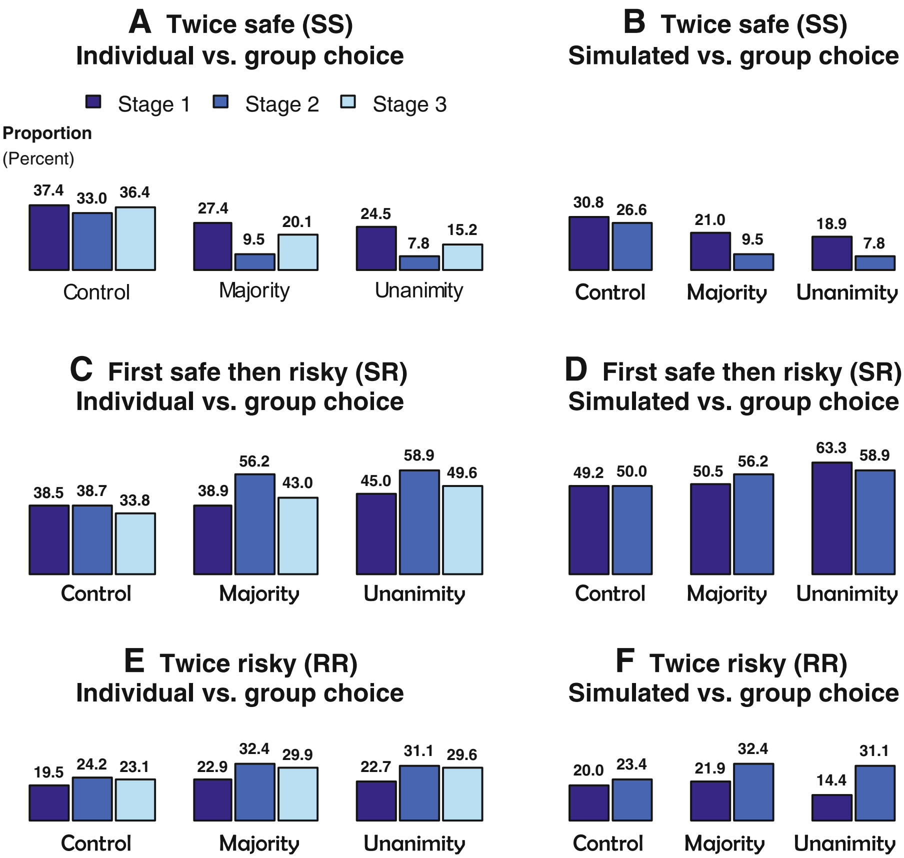


Fig. 3 Proportion of choices in the Allais common ratio tasks per treatment and stage

In the regression comparing groups with simulated choices, there were marginally more RR-choices in stage 2 than in stage 1 for treatment U, suggesting that communication increased risk seeking. This was not observed in the regression comparing groups with individual choices. This result indicates that aggregation effects and communication effects had opposite effects on the proportion of RR choices. Communication effects led to an increase in RR choices, whereas aggregation effects reduced them. These effects could not be observed in the actual choices because the aggregation effects and communication effects were close in magnitude and cancelled out. It shows the added value of analyzing the simulated choices. The finding that communication effects led to more risk seeking suggests that individual choices may be too risk averse and (some) more risk seeking may be rational. On the other hand, the effect was only modest and unlike the findings on NTSD, it did not persist in the third stage.

4.4 Additional analysis

We recorded the decision time at each stage. Table 11 reports the results of a linear regression of decision times on treatments and on stages. The M- and U-treatments took

Table 10 Multinomial probit regressions for the CR tasks

	Group vs. individual choices			Group vs. simulated choices		
Choice pattern	SS	RR	SR	SS	RR	SR
stage2	−.04	.05	−.01	−.03	.03	−.00
	(.03)	(.04)	(.06)	(.04)	(.06)	(.09)
stage3	−.00	.05	−.05			
	(.03)	(.04)	(.05)			
majority	−.07	.05	.02	−.06	.03	.04
	(.06)	(.07)	(.08)	(.09)	(.11)	(.13)
unanimity	−.10*	.03	.07	−.09	−.07	.16
	(.06)	(.08)	(.08)	(.09)	(.11)	(.13)
*stage2*majority*	−.19***	.03	.16**	−.10	.06	.05
	(.06)	(.06)	(.08)	(.08)	(.08)	(.12)
*stage2*unanimity*	−.19***	.04	.14*	−.11	.15*	−.04
	(.06)	(.07)	(.08)	(.09)	(.08)	(.12)
*stage3*majority*	−.08	.01	.07			
	(.06)	(.06)	(.08)			
*stage3*unanimity*	−.10*	.02	.09			
	(.06)	(.06)	(.07)			
No. of observations		2202			505	
Wald chi2		53.82			24.70	
p-value		0.00			0.00	

Reported numbers are the marginal effects at the means of covariates, followed by significance and clustered standard errors between brackets. The standard errors in the left three columns are clustered at the individual level, and those in the right three columns are clustered at the group level. Missing choices and RS patterns are excluded

* significant at 10% (two-sided test)

** significant at 5% (two-sided test)

*** significant at 1% (two-sided test)

more time than the control treatment in stage 2, but they did not differ from each other ($p = 0.20$). Response time decreased for the second and third stages of the control treatment and we also observed that the response time was less in the third stage than in the first stage in both the majority and the unanimity treatments.

Table 11 Linear regression on decision time

stage2	−1.77* (0.96)
stage3	−2.73*** (0.96)
majority	0.05 (0.86)
unanimity	0.50 (0.89)
stage2*majority	3.72*** (1.22)
stage2*unanimity	5.16*** (1.26)
stage3*majority	−0.27 (1.22)
stage3*unanimity	−0.57 (1.26)

Reported numbers are regression coefficients, followed by significance and standard errors. The unit of the dependent variable is minute

* significant at 10% (two-sided test)

** significant at 5% (two-sided test)

*** significant at 1% (two-sided test)

5 Conclusion

We disentangled communication effects and aggregation effects in group decisions. This allowed us to analyze the effects of unanimity and majority rules on intellective and judgmental tasks under risk. Our results show that:

a) Both aggregation effects and communication effects occurred, but these effects were stronger and occurred more often under the unanimity rule.
b) Aggregation effects were mixed and did not always lead to more rational choices. Aggregation effects reduced violations of stochastic dominance, but they also reduced EU-consistent patterns in Allais' common ratio tasks, and they changed the distribution of EU-consistent choice patterns in Allais' common consequence task without affecting the overall proportion of EU-consistent choice patterns.
c) Communication effects favored the justifiable (rational) choices and had more impact in intellective tasks than in judgmental tasks. When there was a clear argument for a particular choice, such as in the nontransparent stochastic dominance tasks, then communication effects could strongly increase the proportion of such choices in the group decisions. Communication effects led to more risk seeking in the common ratio tasks (with marginal significance), but not in the common consequence tasks. Communication effects increased rationality but they also increased decision times.

The separation of communication and aggregation effects introduced in this paper sheds new light on the pros and cons of group decision rules and on their differences with individual decisions. In particular, we can test whether these rules increase the rationality of group decisions, and have a lasting impact on individual decisions after the group process is over. Our results show that communication effects play an important role in group decisions.

Acknowledgments Gary Charness, Enrico Diecidue, Yoram Halevy, the editor W. Kip Viscusi, and an anonymous reviewer gave helpful suggestions. The research of Aurélien Baillon was made possible by a grant from the Netherlands Organization for Scientific Research (NWO).

References

Adams, R., & Ferreira, D. (2010). Moderation in groups: Evidence from betting on ice break-ups in Alaska. *The Review of Economic Studies, 77*(3), 882–913.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'ecole Americaine. *Econometrica, 21*(4), 503–546.

Bateman, I., & Munro, A. (2005). An experiment on risky choice amongst households. *The Economic Journal, 115*(502), C176–C189.

Blavatskyy, P. R. (2013). The reverse Allais paradox. *Economics Letters, 119*(1), 60–64.

Bone, J., Hey, J., & Suckling, J. (1999). Are groups more (or less) consistent than individuals? *Journal of Risk and Uncertainty, 18*(1), 63–81.

Brunette, M., Cabantous, L., & Couture, S. (2015). Are individuals more risk and ambiguity averse in a group environment or alone? Results from an experimental study. *Theory and Decision, 78*(3), 357–376.

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*(2), 129–148.

Charness, G., & Sutter, M. (2012). Groups make better self-interested decisions. *The Journal of Economic Perspectives, 26*(3), 157–176.

Dasgupta, P., & Maskin, E. (2008). On the robustness of majority rule. *Journal of the European Economic Association, 6*(5), 949–973.

Davis, J. H. (1973). Group decision and social interaction: A theory of social decision schemes. *Psychological Review, 80*(2), 97–125.

Davis, J. H. (1992). Some compelling intuitions about group consensus decisions, theoretical and empirical research, and interpersonal aggregation phenomena: Selected examples 1950–1990. *Organizational Behavior and Human Decision Processes, 52*(1), 3–38.

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review, 100*(3), 1238–1260.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association, 9*(3), 522–550.

Evans, D. A. (1997). The role of markets in reducing expected utility violations. *Journal of Political Economy, 105*, 622–636.

Feddersen, T., & Pesendorfer, W. (1998). Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting. *The American Political Science Review, 92*(1), 23–35.

Fishburn, P., & Wakker, P. P. (1995). The invention of the independence condition for preferences. *Management Science, 41*(7), 1130–1144.

Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives, 19*(4), 25–42.

Goeree, J. K., & Yariv, L. (2011). An experimental study of collective deliberation. *Econometrica, 79*(3), 893–921.

Huber, J., Viscusi, W. K., & Bell, J. (2008). Reference dependence in iterative choices. *Organizational Behavior and Human Decision Processes, 106*, 143–152.

Huck, S., & Müller, W. (2012). Allais for all: Revisiting the paradox in a large representative sample. *Journal of Risk and Uncertainty, 44*(3), 261–293.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–291.

Kameda, T. (1991). Procedural influence in small-group decision making: Deliberation style and assigned decision rule. *Journal of Personality and Social Psychology, 61*(2), 245–256.

Keck, S., Diecidue, E., & Budescu, D. V. (2014). Group decisions under ambiguity: Convergence to neutrality. *Journal of Economic Behavior & Organization, 103*, 60–71.

Kerr, N. L., MacCoun, R. J., & Kramer, G. P. (1996). Bias in judgment: Comparing individuals and groups. *Psychological Review, 103*(4), 687–719.

Kerr, N. L., & Tindale, R. S. (2004). Group performance and decision making. *Annual Review of Psychology, 55*(1), 623–655.

Kerr, N. L., & Tindale, R. S. (2011). Group-based forecasting? A social psychological analysis. *International Journal of Forecasting, 27*(1), 14–40.

Kugler, T., Kausel, E. E., & Kocher, M. G. (2012). Are groups more rational than individuals? A review of interactive decision making in groups. *Wiley Interdisciplinary Reviews: Cognitive Science, 3*(4), 471–482.

Laughlin, P. R., & Ellis, A. L. (1986). Demonstrability and social combination processes on mathematical intellective tasks. *Journal of Experimental Social Psychology, 22*(3), 177–189.

Loomes, G. (1988). Further evidence of the impact of regret and disappointment in choice under uncertainty. *Economica, 55*(217), 47–62.

Maciejovsky, B., Sutter, M., Budescu, D. V., & Bernau, P. (2013). Teams make you smarter: How exposure to teams improves individual decisions in probability and reasoning tasks. *Management Science, 59*(6), 1255–1270.

Masclet, D., Colombier, N., Denant-Boemont, L., & Lohéac, Y. (2009). Group and individual risk preferences: A lottery-choice experiment with self-employed and salaried workers. *Journal of Economic Behavior & Organization, 70*(3), 470–484.

Messner, M., & Polborn, M. K. (2004). Voting on majority rules. *The Review of Economic Studies, 71*(1), 115–132.

Miller, C. E. (1985). Group decision making under majority and unanimity decision rules. *Social Psychology Quarterly, 48*(1), 51–61.

Nemeth, C. (1977). Interactions between jurors as a function of majority vs. unanimity decision rules. *Journal of Applied Social Psychology, 7*, 38–56.

Ohtsubo, Y., Miller, C. E., Hayashi, N., & Masuchi, A. (2004). Effects of group decision rules on decisions involving continuous alternatives: The unanimity rule and extreme decisions in mock civil juries. *Journal of Experimental Social Psychology, 40*(3), 320–331.

Rockenbach, B., Sadrieh, A., & Mathauschek, B. (2007). Teams take the better risks. *Journal of Economic Behavior & Organization, 63*(3), 412–422.

Savage, L. J. (1954). *The Foundations of Statistics*. New York: Wiley.

Starmer, C. (1992). Testing new theories of choice under uncertainty using the common consequence effect. *The Review of Economic Studies, 59*(4), 813–830.

Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature, 38*(2), 332–382.

Stasson, M. F., Kameda, T., Parks, C. D., Zimmerman, S. K., & Davis, J. H. (1991). Effects of assigned group consensus requirement on group problem solving and group members' learning. *Social Psychology Quarterly, 54*(1), 25–35.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *The Journal of Business, 59*(4), S251–S278.

Van de Kuilen, G., & Wakker, P. P. (2006). Learning in the Allais paradox. *Journal of Risk and Uncertainty, 33*(3), 155–164.

Viscusi, W. K. (1989). Prospective reference theory: Toward an explanantion of the paradoxes. *Journal of Risk and Uncertainty, 2*, 235–264.

Viscusi, W. K. (1995). Government action, biases in risk perception, and insurance decisions. *Geneva Papers in Risk and Insurance Theory, 20*, 93–110.

Viscusi, W. K., Phillips, O. R., & Kroll, S. (2011). Risky investment decisions: How are individuals influenced by their groups? *Journal of Risk and Uncertainty, 43*(2), 81–106.

Wu, G., Zhang, J., & Abdellaoui, M. (2005). Testing prospect theories using probability tradeoff consistency. *Journal of Risk and Uncertainty, 30*(2), 107–131.

The Economic Journal, **115** (*March*), C176–C189. © Royal Economic Society 2005. Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

AN EXPERIMENT ON RISKY CHOICE AMONGST HOUSEHOLDS*

Ian Bateman and Alistair Munro

A host of experiments have examined theories of risky choice using *individuals*. However, many important economic decisions are taken within multi-adult *households*. This paper reports on the first economic experiment designed to test theories of household choice. We use established couples and face them individually and jointly with decisions involving monetary payoffs. We find that joint choices typically are more risk averse than those made by individuals. Meanwhile, choices made by couples exhibit the same kinds of departures from expected utility theory (e.g. the common ratio and common consequence effects) as are regularly recorded with individuals.

The expected utility-maximising household is one of the most common models employed to understand economic behaviour. This *standard model*, which is used to investigate saving, insurance decisions, labour supply etc., involves two important assumptions. First, that the household acts as if it has a single set of preferences and, second, that these preferences conform to the axioms of expected utility theory (EUT). The first of these assumptions has received scrutiny (Browning and Chiappori, 1998) but very little attention has been paid to the second assumption for households as opposed to individuals. In fact though there is copious experimental evidence on how individuals choose, to date there has been very little experimental investigation into how multi-adult households or couples make their decisions. For instance, in Starmer's (2000) extended survey of the field of risky choice, there is no discussion of evidence on household as opposed to individual behaviour and though there is an interesting body of work by psychologists on this issue (see Corfman and Lehmann (1987) for example), the questions asked provide little insight into the applicability of economists' models of choice.

This paper therefore presents results of an experiment designed to investigate the following issue: to what extent do the decisions made by couples and the decisions made separately by individuals who are part of a couple conform to the standard model? In outline the experiment is as follows: we use a sample of established couples[1] and present them with tasks of the kind depicted in Figure 1, all of which involve binary choices between lotteries. In Section 1 of the experiment the subjects are separated and face choices separately; in Section 2 they remain apart and must predict their partner's answers from Section 1; in Section 3 they rejoin their partner and make choices as a couple. There is some overlap in the tasks faced in each section. Each lottery has possible monetary payoffs for each

* The authors wish to thank the editor and two anonymous referees for their comments. The work reported here was financed by ESRC grant no. m545285002 and written up while Alistair Munro was visiting the University of Queensland, for whose hospitality he is grateful.

[1] Meaning that the couple are in a relationship of at least one year's standing and live together.

Question 2

Option A		
For numbers:	1–50	51–100
You receive	£20	£0
Your partner receives	£0	£20

Option B		
For numbers:	1–50	51–100
You receive	£20	£40
Your partner receives	£0	£20

I choose(*tick one*): Option A □ Option B □

Fig. 1. *A Typical Question from Section 1 of the Experiment*

individual within the couple[2] and these payoffs may be different. A random lottery device is used to provide incentives. We use the results of the choices to compare whether couples have similar revealed preferences to individuals. In particular we test whether couples' joint choices show the same a pattern of anomalies that has been regularly reported in individual choices.

Before presenting the experiment in detail it is worth making two points. First, it could reasonably be supposed that the results of individual choice experiments should carry over into household decision making. However, leaving aside the issue of differences in subject pool, the decision-making environment of the household might eliminate some anomalies commonly observed in individual choice. For instance, with two people scrutinising options rather than one, the kind of editing and framing effects which underlie Kahneman and Tversky's (1979) explanation of the Allais paradox might not apply. Conversely, even if individuals separately have preferences which satisfy EUT, the rule used to aggregate preferences within the household might produce choices for the household at variance with EUT. So, the existing theory and experimental evidence on individual choice does not therefore imply much about how households make choices in risky situations and in particular, whether households conform to the standard model.

Secondly, in the standard model of household choice referred to above, the household is assumed to be unitary – that is, the household is modelled as a single agent with a single set of preferences – either because there are no public goods local to the household and all members share the same preferences or because the structure of incentives within the household align individual preferences with those of the decision-maker, as in, for instance, the 'rotten kid theorem', Becker (1974). Empirical testing (Lundberg *et al.*, 1997; Alderman *et al.*, 1995) gives results largely hostile to the unitary model, particularly its prediction of income pooling (IP) which is the property that household behaviour may respond to changes in aggregate household income but not to who in the household earns that income. This has prompted a large number of alternative household models but nevertheless IP remains a convenient assumption in many contexts. For the experiment we design a mix of tests, some of which are conditional on households

[2] Bone *et al.* (1999, 2000) examines decisions made by pairs but using students paired at random, whereas we are interested in the behaviour of pre-existing decision-making units – i.e. established couples. Also, in their design, the pairs of students are given a collective payment and must decide how to divide it. Our lotteries assign payments to individuals (and our payment procedures reflect this), though this is not to deny the possibility of bargains being made or anticipated.

satisfying IP and some of which are not. One reason for having the conditional tests is that in many empirical situations it may not be possible to observe the sources of income in a household. We wish to see whether any departures from EUT are robust in the sense that they are still observable in the face of variation in the identity of the income recipient.

1. Theory

For simplicity we consider a two-person household. Let agent $i = 1,2$ receive payment m_{is} in state of the world $s = 1,\ldots,S$. A typical lottery $\boldsymbol{p}$ (or $\boldsymbol{q}$, $\boldsymbol{r}$ or $\boldsymbol{s}$) is then a vector $(p_1,\ldots,p_S)$. The standard sign, $\succeq$ denotes the weak preference relationship for the household, with strict preference denoted $\succ$ and indifference $\sim$, constructed in the usual manner.

A household obeys expected utility theory in its joint choices if there exists a strictly increasing function $w(m_{1s},m_{2s})$ such that the household ranks lotteries according to, $W(\boldsymbol{p}) \equiv \sum_{s=1}^{s=S} p_s w(m_{1s}, m_{2s})$. In other words $W(\boldsymbol{p}) \geq W(\boldsymbol{q}) \leftrightarrow \boldsymbol{p} \succeq \boldsymbol{q}$. In a similar manner it is also possible to define utility functions, W^i $(\boldsymbol{p})$ $i = 1,2$, for the two individuals. Note that the relationship between the $W^i s$ and W depends on the household aggregation rule. So the fact that the household choices conform to EUT does not imply that the $W^i s$ satisfy the axioms of EUT – or vice versa.

We shall say that the household *income pools* (IP) or that it is an *income pooler* if $w(m_{1s},m_{2s}) = w(m'_{1s},m'_{2s})$ whenever $m_{1s} + m_{2s} = m'_{1s} + m'_{2s}$ for all s.

Although w has two arguments rather than the one that is typical of individual choice, nevertheless for the household or individual which maximises $W(\cdot)$, preferences between lotteries should have the familiar properties of EUT. Figure 2 shows a standard unit probability triangle representing lotteries involving three possible values of w: w_1, w_2 and w_3, with $w_3 > w_2 > w_1$. In the Figure, the solid line connecting a and e is parallel to that between c and d.

In the unit probability triangle, EUT predicts that indifference curves are straight, parallel lines. However, individuals frequently fail to conform to the predictions of EUT in a number of ways. Possibly, three of the most robust anomalies (Starmer, 2000) are the *common ratio* effect, the *common consequence* effect

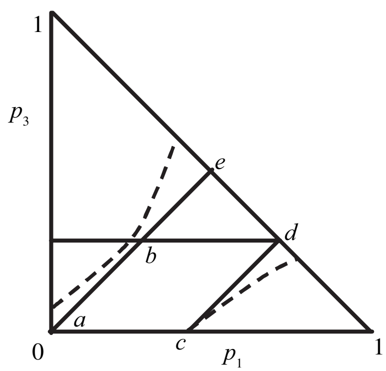


Fig. 2. *Tests of EUT in the Unit Probability Triangle*

and failure of the *betweenness* property. The common ratio effect occurs when individuals choose a out of the pair {*a*,*e*} and then *d* out of the pair {*c*,*d*}, when the lotteries, *a*, *c*, *d* and *e* are such that, as in Figure 2,

(*i*) the line between *a* and *e* is parallel to that between *c* and *d* and
(*ii*) in *e* and *d*, there is a zero probability of the intermediate outcome, w_2.

Note that in this anomaly, the safer option is chosen out of {*a*,*e*}, but then the riskier option is chosen out of {*c*,*d*}. The common consequence effect occurs when *a* is chosen out of a pair {*a*,*b*} and *d* is picked from {*c*,*d*}, where *a*, *b*, *c* and *d* are such that:

(*i*) the line between *a* and *b* is parallel to that between *c* and *d*;
(*ii*) as in Figure 2, lotteries *b* and *d* share a common consequence in the sense that they have the same probability of w_3 and the same is true for *a* and *c*.

Betweenness is the property that if, as in Figure 2, lottery *b* is a probability mixture of lotteries *a* and *e*, then the utility of *b* should between that of *a* and *e*. For example, EUT implies that individuals who choose *b* out of {*a*,*b*} in Figure 2 should choose *e* out of {*a*,*e*}, whereas often individuals choose *b* and then *a*. Indifference curves in the triangle therefore seem to be more like the broken lines depicted in Figure 2, than the straight and parallel lines implied by EUT. Now all of these anomalies are well documented in individuals. We use them to interpret the experimental test reported below where we see if couples exhibit similar patterns of revealed preferences.

2. Experimental Design

Figure 3 summarises the design. Upon entering the venue, one member of the couple was randomly allocated either a 'wave' or 'triangle' card, their partner receiving the other card. This allocation of cards was then used to separate the pair. The first two sections of the experiment were conducted with the separated partners in different rooms; pairs then rejoined each other for Section 3. Throughout the experiment the investigators used a script (available from the authors) and subjects received instructions (including a written summary) one section at a time.

In the experiment a *task* is a pair of lotteries, such as the one depicted in Figure 1, which could either be presented to an individual as a choice question (Section 1) or to an individual as a prediction question (Section 2) or to a couple as a joint choice question (Section 3). The description of each component lottery consisted of three elements: ranges of numbers were shown along the top, underneath which were shown corresponding payoffs for the subject,[3] below which were the corresponding payoffs for their partner. The numbers along the top corresponded to numbered discs in a bag of one hundred discs shown to the subjects by the experimenters.

[3] For the joint choice tasks, the triangle partner's payoffs were always shown first. We found no evidence that this order gave triangle partners more or less influence in the joint decisions.

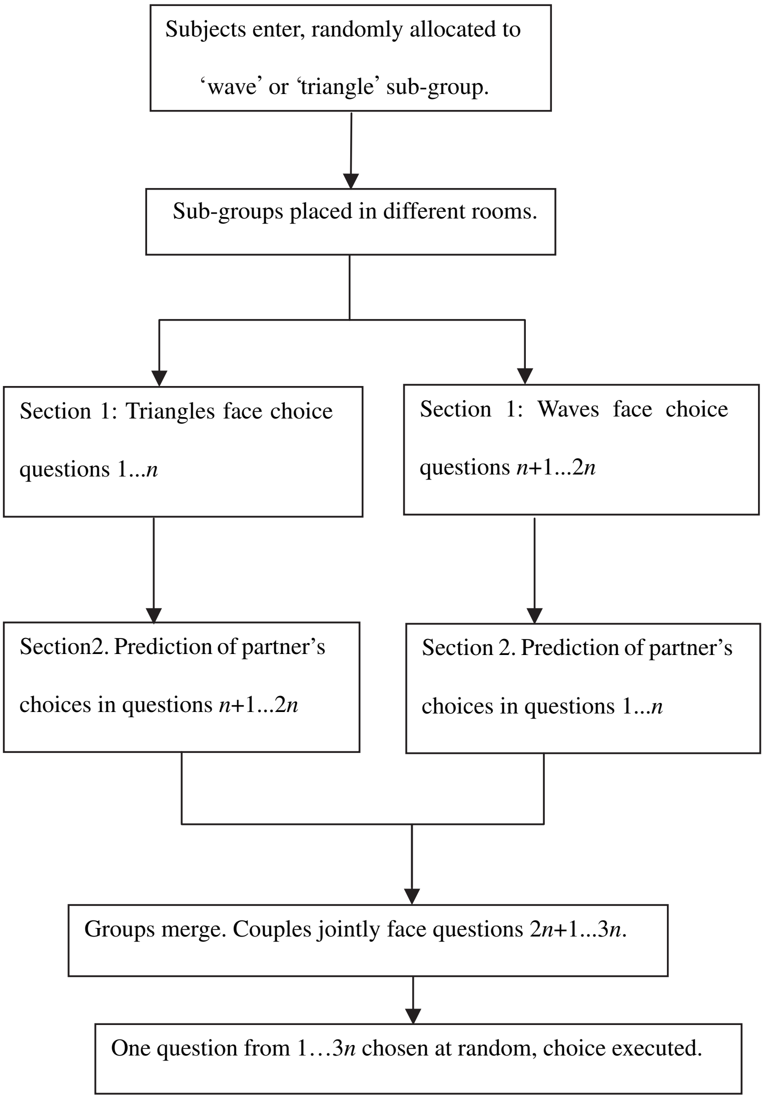


Fig. 3. *Experimental Procedure*

In Section 1 of the experiment the separated subjects had to choose one of the lotteries in each task. They were told that, at the end of the experiment, one task would be chosen at random for each couple and played out for real. If this was one of the tasks from Section 1, the subject would play out the lottery he or she chose in that task by taking a numbered disc from the bag, that number determining the resulting payoffs for them and their partner.

In Section 2, subjects (who remained separated from their partners) were asked to predict their partners' answers from Section 1. Before doing so, they were led through the relevant instructions, including those concerning incentives (see below). After making their predictions, subjects completed a short questionnaire which collected demographic details. Once this was filled in answer books were collected and subjects rejoined their partners for the final section.

In the final section, couples made choices jointly. At the start of this Section each couple selected a small envelope from a shuffled pile of envelopes placed in front of them but were told not to open it until instructed. No prompts were given as to which partner should make the selection. The pile of envelopes contained lottery ticket numbers for all questions from all sections, with one number inside each envelope. The randomly drawn number inside the envelope determined which question the couple would play out 'for real' at the end of the experiment. The subjects were given details of how the payout procedures would operate at the end of the experiment and led through the instructions for Section 3. Once all subjects had completed their tasks, we began opening the small envelopes, executing lotteries and making payoffs.

The incentive system was as follows: for lottery ticket numbers from 1 to n, the triangle partner played his or her choice for that Section 1 question and the wave partner received £0.50 for each correct prediction (in Section 2) of their partner's Section 1 answers. For numbers between $n + 1$ and $2n$, the wave partner played his or her choice from Section 1 while the triangle partner was paid £0.50 for each correct prediction (in Section 2) of their partner's Section 1 answers. For numbers from $2n + 1$ to $3n$, the couple played their joint choice for that question from Section 3 and no money was paid for predictions. This random lottery system[4] is incentive compatible if individuals are selfish (and make no binding agreements on *ex-post* trade), but it would be usual to suppose some degree of other-regarding preferences within couples. As a result, it is conceivable that an altruist might view the first two sections of the experiment as an exercise in co-ordination and possibly choose so as to maximise the predictive success of his or her partner. We aimed to guard against this possibility in three ways. First, prior to the experiment we told participants only that the experiments were aiming 'to help us understand how couples make decisions'. As a second measure the prediction questions (in Section 2) always came after the separate choice questions (in Section 1). We also saved the briefing for Section 2 until all subjects in a session had completed Section 1. So, subjects therefore had no reason to anticipate that they should answer in Section 1 so as to raise the possible payoffs of their partner. As a final measure we kept the payments for prediction to a relatively small fraction of the payments associated with the choice sections. The difference in payments meant that (for example) a risk neutral income pooler would not make an expected gain from switching choices in order to improve the predictive success of their partner.

We went to some trouble to preserve the confidentiality of the answers from sections 1 and 2. Partners were paid sequentially and separately with payments placed in envelopes. The payment process occurred in another room or in a position which masked the payments made. Subjects were not informed of their partner's answers in Section 1 of the experiment and they were not given information about the accuracy of their partner's predictions. Our main reason for

[4] Cubitt *et al.* (1998) provide evidence that random lottery schemes are a reliable means of eliciting preferences even when subjects are not EUT maximisers.

confidentiality is as follows: Many economic theories of the household relate collective choice to individual preferences over goods. To test such theories we normally require data about individual preferences over commodities. Revealing choices to partners might instead produce information on preferences over actions. For instance, it might create incentives for individuals to choose so as to garner approval from their partners. Such motives may actually be an important source of household behaviour but they are not typically the objects of preference in economic theories of the household. So, we opted for confidentiality in our design.[5]

3. Results

Following a successful pilot session, the experiments were carried out from December 2002 to March 2003.[6] Subjects were recruited from the city of Norwich and rural Norfolk via email, through community groups and using posters. Session sizes varied from two to ten couples and were held at a variety of venues, including a village hall and the experimental economics laboratory at the University of East Anglia. In recruiting we required all individuals to be over 21, to be living with their partners and to have been together as a couple for at least one year. We asked subjects to bring evidence of their relationship and made random checks.[7] We recruited 76 couples for our experiment. Average payoffs were just under £17 per individual – more than twice the median hourly post-tax wage for a UK adult in 2003. Ages ranged from 22 to 70, with a mean of 37.3. On average couples had been together for 11 years, with a maximum of 46 and a minimum of 1. Seventy-three per cent of individuals stated that they were married to their current partner and all the couples in our sample were heterosexual. The distribution of children per couple was bimodal with peaks at zero and two and a mean of 1.1. So, without being representative of the UK adult population, the subjects were generally older and more diverse than the typical sample of university students used in choice experiments.

In what follows, the tasks are labelled. Their details can be found in the Appendix. The number identifies the task, while the letter indicates the identity of the decision maker: 'T' means a task was faced by triangle subjects in their Section 1, 'W' stands for tasks faced by wave subjects in their Section 1 and 'J' indicates tasks faced jointly by couples in Section 3.

For both joint and separate choice we included four tasks (T13, J13, W13 and W14) such as the one shown in Figure 1 where one option first-order stochastically dominates the other. For these questions the dominated option was chosen in just under 6% of observations.

[5] If one partner predicted perfectly (or scored zero), then provided she or he had perfect recall, that subject could deduce a partner's choices. No participant raised this possibility with us during the conduct of the experiment or achieved perfection in their predictions (or scored zero).

[6] The initial design had 10 questions in each section, subsequently expanded to 12. Standard statistical tests indicated that there was no significant differences between the data for the questions common to both variants and so these responses were pooled within our analysis.

[7] Evidence included passports, photos, bills to the same address and, in three cases, children.

Recall that IP (income pooling) is a feature of what we termed the standard model. We had seven tasks where one of the options dominates the other, for subjects whose choices satisfy IP (but not necessarily otherwise). In 90% of cases the choice is in conformity with IP and this accordance is stronger for the choices made jointly than for those made when the individuals are separated. Suppose we hold the null hypothesis that in all cases subjects mean to choose the IP dominating option, but make a mistake in 6% of cases (i.e. the rate of 'error' in the choices with one dominating option discussed above). With the exception of one task (T11), the pattern of choices is consistent with this null hypothesis, suggesting that IP is a reasonable assumption in the context of this experiment. We also had a number of tests of IP based on pairs of tasks which are equivalent when faced by a chooser who satisfies the IP property. That data is more mixed in the conclusions it produces; see Bateman and Munro (2003), which focuses on the IP issue. So, the evidence for IP in our data is not overwhelming and we therefore conduct tests of EUT both with and without its presence as an auxiliary assumption.

Tables 1–3 summarise our tests of EUT: figures in Table 1 are based upon the individual choice data, Table 2 draws upon the prediction tasks and Table 3 is based upon joint choices. All the comparisons shown are within subject or within couple. In the first column of these three Tables, entries labelled CR represent common ratio tests, those marked CC represent common consequence tests, while BB indicates tests of the betweenness property of EUT. The next column states whether the comparison is conditional on the assumption of the IP property. If it is, then this means that pairs of tasks can only be plotted in the same unit probability triangle if IP holds. In the two 'proportions' columns, the numbers

Table 1

Tests of EUT Using Separate Choice Data

					Proportion choosing safer option,		
N	Type of comparison	IP assumed?	Task 1	Task 2	Task 1	Task 2	Probability
76	CR	No	T6	T8	0.66	0.43	0.000***
76	CR	Yes	T6	T7	0.66	0.54	0.073*
76	CR	Yes	T4	T8	0.78	0.43	0.000***
76	CR	Yes	T4	T7	0.78	0.54	0.001***
34	CR	Yes	W4	W7	0.77	0.53	0.001***
34	CR	Yes	W5	W7	0.71	0.53	0.054*
76	CC	No	T1	T8	0.50	0.43	0.190
34	CC	No	W3	W7	0.53	0.53	0.500
34	CC	Yes	T2	T8	0.53	0.32	0.055*
34	CC	Yes	W2	W7	0.71	0.47	0.028**
76	BB	No	T6	T1	0.66	0.50	0.048**
34	BB	Yes	W5	W3	0.71	0.56	0.042**
76	BB	Yes	T4	T1	0.78	0.50	0.000***

***indicates difference significant at 1% level, 1 tailed test;
**indicates significant at 5% level;
*indicates significant at 10% level.

Table 2

EUT and Prediction Data

			Comparison involves		Proportion predicting safer option		
N	Type of comparison	IP assumed?	Task 1	Task 2	Task 1	Task 2	Probability
76	CR	No	T6	T8	0.70	0.43	0.000***
76	CR	Yes	T6	T7	0.70	0.59	0.051*
76	CR	Yes	T4	T8	0.88	0.43	0.000***
76	CC	No	T1	T8	0.62	0.43	0.011**
34	CC	No	W3	W7	0.47	0.32	0.078*
34	CC	Yes	T2	T8	0.68	0.41	0.024**
34	CC	Yes	W2	W7	0.68	0.32	0.002***
76	BB	No	T6	T1	0.70	0.62	0.068*
34	BB	Yes	W5	W3	0.71	0.47	0.009***
76	BB	Yes	T4	T1	0.88	0.62	0.000***

***indicates difference significant at 1% level, 1 tailed test;
**indicates significant at 5% level;
*indicates significant at 10% level.

represent the fraction of the sample choosing the safer option.[8] According to EUT, the fraction should be the same across the relevant tasks. This is always the null hypothesis. According to the typical results of individual choice experiments the proportion in the task 2 column should be lower. This is always the alternative hypothesis. Taking all of the tables together we see only one instance of equality (the EUT prediction) compared to the remaining 32 cases all of which are in the direction of the alternative hypothesis. In the final column we report probability values for the null hypothesis that the sample proportions are equal, using a paired, one-sided z-test. A large number of these comparisons are statistically significant; in many cases at levels of significance well below 0.1%.

Table 1 summarises results for the separate choice. All the CR and BB comparisons are statistically significant at the 10% level or lower. For two CC cases where the test is not conditional on IP, the difference in responses to the two tasks is not statistically significant.[9] So, broadly speaking the evidence for a common ratio effect and for failure of the betweenness property is stronger than that for the common consequence effect. (Chew and Waller (1986) find a similar pattern in their experiment on individual choice. Later experiments have suggested that the relative strengths of the common ratio and common consequence effect are sensitive to the parameters of the experiment.) In terms of the stylised indifference curves in Figure 2, the pattern of our results suggests that the section between a and b is roughly parallel to the curve between c and d but that the indifference

[8] Rabin (2000), argues that choosing the safe option in choices of this kind is prima facie evidence against the EUT model, because it is incompatible with attitudes to risk displayed in other settings. Here, the same argument does not apply because the choice between options also typically reflects intrahousehold inequality aversion.

[9] In a between-subject CC comparison, a proportion 0.53 choose the safer option out of $\{a,b\}$ while 0.43 pick the safer option out of $\{c,d\}$. This is significant at the 10% level ($p = 0.095$).

Table 3
Tests of EUT Using Joint Choice Data

N	Type of comparison	IP assumed?	Comparison involves		Proportion choosing safer option,		Probability
			Task 1	Task 2	Task 1	Task 2	
34	CR	Yes	J4	J7	0.94	0.59	0.000***
34	CR	Yes	J5	J7	0.88	0.59	0.000***
34	CR	No	J10	J7	0.73	0.59	0.013**
34	CC	No	J1	J7	0.71	0.59	0.110
34	CC	Yes	J2	J7	0.71	0.59	0.130
34	CC	Yes	J9	J11	0.44	0.06	0.000***
76	BB	No	J4	J1	0.94	0.64	0.000***
34	BB	Yes	J5	J1	0.88	0.71	0.006***
34	BB	Yes	J4	J2	0.94	0.71	0.002***
34	BB	Yes	J5	J2	0.88	0.71	0.016**

***indicates difference significant at 1% level, 1 tailed test;
**indicates significant at 5% level;
*indicates significant at 10% level.

curves show increasing risk aversion (i.e. become steeper) between *b* and *e*. It is worth noting that this pattern persists, even when the identity of the recipient of payoffs changes – for some of the tasks the payoffs are to the choosing agent, but for many tasks both partners might possibly receive payment and in several cases it is only the partner that might receive payments. Nevertheless the pattern of choices is consistent.

Table 2 presents the data from the prediction section of the data. These tasks are the same as those in Table 1, but it is the other partner who is doing the predicting. All the comparisons are statistically significant, even with the CC examples. So, the results suggest that prediction deviates significantly from EUT. When we look at the prediction data in detail we find that partners predict correctly in 65% of cases. This is significantly better than 50–50; it is also better than the success rate if they supposed (as a benchmark example) that their partner was a risk neutral income pooler. However, if individuals predict according to how they themselves choose and preferences are not correlated within couples then the predicted success rate is 64.7% – which is not statistically significantly different from the actual value.

The fact that, when separated, individual partners depart from the standard model in their choices does not mean that those individuals have non-EUT preferences. They may be altruists who believe that their partners have non-EUT preferences. Similarly, in the absence of common knowledge, it cannot be deduced that individuals who predict anomalous behaviour in their partners actually believe that their partners have non-EUT preferences. If, though, this was the case, then we would expect the possibilities for communication afforded by the joint decision-making responsibility of Section 3 to iron out any misunderstandings. In fact, as Table 3 shows, with and without the auxiliary assumption of IP the joint choice data exhibit the same patterns as the prediction and separate choice data. It suggests that the departure from EUT reported in Tables 1 and 2 is not due to misconceptions about the preferences of partners. Rather, it seems to be a persistent feature of choice in the context of multi-person households.

Table 4 summarises comparisons of choices made jointly (in Section 3) and when separated (in Section 1) for the tasks where one option is safer than the other and where all the trade-offs are in one partner's payoffs. Note that, when viewed by Wave and Triangle subjects, these tasks appear *reflected* in the sense that, for any given task, payoffs which belong to the self when Wave chooses belong to the partner when Triangle chooses and vice versa. Three things are particularly notable in Table 4. First, the proportions for Wave and Triangle subjects are very close – in other words subjects appear to place equal weight on their partner's payoffs as on their own and are not more or less risk averse when it is their partner who faces the risk rather than themselves. Second, perhaps surprisingly, choices made jointly are consistently more risk averse than those made separately, to the extent that, as the penultimate column shows, in three of the four cases the difference is statistically significant whichever partner is taken as the benchmark. It is not clear to us why this result occurred. For example, standard risk sharing arguments would predict joint choice would exhibit less risk aversion, since it gives opportunities for agreeing to ex-post risk sharing transfers. We can also rule out misperception of the partner's degree of risk aversion as the explanation, since there is no evidence for such a bias in the prediction data. Possibly, the result is due to the psychology of group choice, one robust feature of which (Kerr *et al.*, 1996) is that collective decisions are typically more extreme than their individual counterparts. Yet, in our case it is not clear why the safe option should be viewed as more extreme. A final possibility is suggested by anecdotal evidence from our participants, some of whom suggested a 'fear of recrimination' as a significant factor influencing joint choices. This could make some participants reluctant to be seen to be pressing for the risky option.

Table 4

Comparison Between Choices Made Jointly and Separately

	Proportion choosing safer option			Tests	
	Separate choice		Joint Choice		
Task	Triangle	Wave	Joint	Test of Equality	Correlation (p-value)
1	0.5	0.53	0.64	0.030**	0.053 (0.649)
4	0.78	0.77	0.94	0.000***	0.175 (0.322)
7	0.54	0.53	0.64	0.059*	0.016 (0.889)
9	0.35	0.39	0.44	0.364	−0.092 (0.611)

'Test of Equality' is the p-value associated with the test of equality between the joint choice value and the closest value from separate choice. 'Correlation' is Kendall's Tau B correlation between the separate choices; the associated p-value is for the null hypothesis that the correlation coefficient is zero.
***indicates difference significant at 1% level, 1 tailed test;
**indicates significant at 5% level;
*indicates significant at 10% level.

The final column of the Table also gives pause for thought as it shows the lack of correlation between the separate choices of the partners. These correlations are not just low, they are never even remotely significantly different from zero. We do find high levels of correlation within the choices made separately by individuals, so the data in the final column is not simply evidence of randomness. Rather it suggests that there is little correlation in risk attitudes towards monetary lotteries within couples.

4. Discussion

The fact that the preferences of two individuals separately conform to the assumptions of EUT does not imply that their collective decisions will always obey the same axioms. Conversely, depending on the household decision process, it is possible that two individuals with non-EUT preferences can produce collective choices that do satisfy the predictions of EUT. It follows that tests of whether the decisions of established couples conform to EUT are logically separate from the issue of whether individual decisions satisfy the theory. Nevertheless, in this experiment a clear result is that couples show the same anomalous patterns in their risky choices as have been frequently observed in individual choice experiments. When separated from their partners, individuals who are part of a couple also show the same patterns and predict the same patterns in their partner's choices. The results of the experiment also suggest that the results are robust in the face of changes in the identity of who in the household receives the payoffs.

In the face of individual choice anomalies many alternatives to EUT have been proposed (e.g. regret theory, prospect theory, etc.). To a significant degree, these theories have been motivated by ideas drawn from the psychology of the individual. The same ideas may not automatically apply for the household, where decisions are typically made interactively; other forces may be at work in the results found here. For instance, we found an unexpectedly high incidence of examples where joint choice was more risk averse than choices made separately. If such patterns are a feature of many households it would suggest that behavioural models of collective decision making may be quite different to their individual counterparts.

University of East Anglia and University of Western Australia
University of East Anglia

Appendix

The Tasks

		Lottery 1				Lottery 2			
		Triangle		Wave		Triangle		Wave	
Task number	Subjects	£20	£40	£20	£40	£20	£40	£20	£40
1	T, W, J	1–100	–	–	–	21–70	71–100	–	–
2	T, W, J	21–100	–	1–20	–	21–100	–	71–100	–
3	W	–	–	1–100	–	–	–	21–70	71–100

Table (*Continued*)

		Lottery 1				Lottery 2			
		Triangle		Wave		Triangle		Wave	
Task number	Subjects	£20	£40	£20	£40	£20	£40	£20	£40
4	T, W, J	1–100	–	–	–	–	–	–	41–100
5	W, J	1–50	–	51–100	–	41–100	–	41–100	–
6	T	1–100	–	–	–	–	41–100	–	–
7	T, W, J	–	–	51–100	–	–	–	–	71–100
8	T	51–100	–	–	–	–	71–100	–	–
9	T, W, J	1–100	–	1–100	–	1–100	–	21–70	71–100
10	T, W, J	–	–	1–100	–	–	1–70	–	–
11	T, J	51–100	–	1–50		1–100	–	–	71–100
12	W	31–100	–	1–30	–	–	71–100	1–100	–
13	T, W, J	1–50	–	51–100	–	1–50	51–100	51–100	–
14	W	1–70	–	21–70	71–100	71–100	–	1–70	–
15	J	–	71–100	1–70	–	1–60	61–100	–	–
16	J	–	–	21–70	71–100	1–40	–	–	71–100

Note: Task numbers do not match the order of questions, but are for reference purposes. The letters under 'Subjects' identify the groups that faced the task as a choice question. Numbers show the ranges of disc values for which the corresponding payoffs were awarded, with the numbers omitted for disc values where the payoff was zero. We also omit tasks which are not relevant for this paper.

References

Alderman, Harold, Chiappori, Pierre-Andre, Haddad, Lawrence, Hoddinott, John and Kanbur, Ravi (1995). 'Unitary versus collective models of the household - is it time to shift the burden of proof', *World Bank Research Observer*, vol. 10(1), (February), pp. 1–19.

Bateman, Ian and Munro, Alistair (2003), 'Testing economic models of the household: an experiment', CSERGE Working Paper, University of East Anglia.

Becker, Gary S. (1974), 'A theory of social interactions', *Journal of Political Economy*, vol. 82 (6), pp. 1063–93.

Bone, John, Hey, John and Suckling, John (1999), 'Are groups more or less consistent than individuals?', *Journal of Risk and Uncertainty*, vol. 8, pp. 63–81.

Bone, John, Hey, John and Suckling, John (2000). 'A simple risk-sharing experiment', University of York Discussion papers in Economics, No. 2000/36.

Browning, Martin, and Chiappori, Pierre-Andre (1998). 'Efficient intra-household allocations: a general characterisation and empirical tests', *Econometrica*, vol. 66(6), pp. 1241–78.

Chew, Soo Hong and Waller, William S. (1986). 'Empirical tests of weighted utility theory', *Journal of Mathematical Psychology*, vol. 30, pp. 55–72.

Corfman, Kim. P. and Lehmann, Donald. R. (1987). 'Models of cooperative group decision-making and relative influence: an experimental investigation of family purchase decisions', *Journal of Consumer Research*, vol. 14, pp. 1–13.

Cubitt, Robin P., Starmer, Chris. V. and Sugden, Robert (1998). 'On the validity of the random lottery incentive system', *Experimental Economics*, vol. 1, pp. 115–31.

Kahneman, Daniel, and Tversky, Amos (1979). 'Prospect theory: an analysis of decision under risk', *Econometrica*, vol. 47, pp. 263–91.

Kerr, Norbert L., MacCoun, Robert J. and Kramer, Geoffrey P. (1996). 'Bias in judgment: comparing individuals and groups', *Psychological Review*, vol. 103(4), (October), pp. 687–719.

Lundberg, Shelly J., Pollak, Robert A. and Wales, Terence J. (1997). 'Do husbands and wives pool their resources? Evidence from the United Kingdom child benefit', *Journal of Human Resources*, vol. 32(3), (Summer), pp. 463–80.

Rabin, Matthew (2000). 'Risk aversion and expected-utility theory: a calibration theorem', *Econometrica*, vol. 68, pp. 1281–92.

Starmer, Chris, V. (2000). 'Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk', *Journal of Economic Literature*, vol. 38, pp. 332–82.

Theory Dec. (2011) 70:45–64
DOI 10.1007/s11238-009-9179-6

Individual and couple decision behavior under risk: evidence on the dynamics of power balance

André de Palma · Nathalie Picard · Anthony Ziegelmeyer

Published online: 24 October 2009

Abstract This article reports results of an experiment designed to analyze the link between risky decisions made by couples and risky decisions made separately by each spouse. We estimate both the spouses and the couples' degrees of risk aversion, we assess how the risk preferences of the two spouses aggregate when they make risky decisions, and we shed light on the dynamics of the decision process that takes place

Electronic supplementary material The online version of this article (doi:10.1007/s11238-009-9179-6) contains supplementary material, which is available to authorized users.

A. de Palma
Ecole Normale Supérieure de Cachan, 61 Avenue du Président Wilson, 94230 Cachan, France
e-mail: andre.depalma@ens-cachan.fr

A. de Palma
Institut Universitaire de France, Paris, France

A. de Palma · N. Picard
Ecole Polytechnique, Palaiseau, France

N. Picard
THEMA, University of Cergy-Pontoise, 33, Bd du Port, 95011 Cergy-Pontoise Cedex, France
e-mail: nathalie.picard@u-cergy.fr

N. Picard
INED, 133 Boulevard Davout, Paris, France

A. Ziegelmeyer (✉)
Strategic Interaction Group, Max Planck Institute of Economics, Kahlaische Str. 10, 07745 Jena, Germany
e-mail: ziegelmeyer@econ.mpg.de

A. Ziegelmeyer
Faculty of Economics and Management, Technical University of Berlin, Straße des 17 Juni 135, 10623 Berlin, Germany

when couples make risky decisions. We find that, far from being fixed, the balance of power within the household is malleable. In most couples, men have, initially, more decision-making power than women but women who ultimately implement the joint decisions gain more and more power over the course of decision making.

Keywords Balance of power · Experiments · Household decision-making · Risk

1 Introduction

Almost every important economic decision involves risk, and a substantial body of research investigates how individuals incorporate risk into their decision process. In this body of literature, only a tiny portion is devoted to the study of household decision-making under risk. However, in many day-to-day life contexts such as financial investments, insurance, retirement plans, or residential location, the decisions have consequences at the household level rather than at the individual level. These decisions are (or should be) made jointly. Even when these decisions are formally made by only one member of the household, they may modify (and/or can be modified by) other decisions in the household.

A growing literature in economics shows that household savings and financial investments are significantly affected by how decision-making power is allocated between men and women.[1] This empirical work observes household outcomes and changes in members' incomes to draw conclusions about underlying gender preferences. As argued below, we strongly believe that this evidence should be interpreted with caution and that such empirical results are not necessarily reflective of intrinsic or immutable preference differences between women and men. In order to identify the link between risky decisions made by couples and risky decisions made separately by each spouse, we use an experimental approach. We observe intra-household financial decisions in an artifactual field experiment.[2] Our experiment was explicitly designed to investigate the decision process that takes place when couples make risky decisions.[3]

Until recently, household decisions were treated in the standard neoclassical framework of economic theory. This approach corresponds to the *unitary model*, which

[1] For example, income given to women is more likely to be used for investments in education and housing than income given to men (Duflo 2003).

[2] An artifactual field experiment is identical to a typical laboratory experiment but one which makes use of a non-standard subject pool (see the terminology of Harrison and List 2004).

[3] Though there exists copious experimental evidence on how individuals choose, there has been very little experimental investigation into how households or couples make their decisions. Bateman and Munro (2005) present results of an experiment designed to investigate the extent to which decisions made by couples and decisions made separately by spouses are consistent with the axioms of Expected Utility Theory (EUT). They find that choices made by couples exhibit the same kinds of patterns (e.g., the common ratio and common consequence effects) as are regularly recorded with individuals. Bateman and Munro (2009) report on a choice experiment using reductions in dietary health risks as the vehicle. In one treatment, a random individual is chosen from the couple and takes part in a face-to-face interview; in the other treatment, both partners are asked questions jointly, again in a face-to-face interview. They find significant differences in the values elicited in the two treatments, and the values elicited from couples are not a simple average of those elicited from men and women.

involves a unique decision-maker representing the household. From a methodological point of view, unitary models are open to criticism, since they hide the divergences of interest that may arise among household members. Unitary models implicitly assume that the household's members pursue consensual objectives. However, individual preferences cannot be easily aggregated. As pointed out by Chiappori (1988), joint decision making has a different meaning within a couple than in other contexts such as professional interactions. A poor understanding of decision mechanisms (and therefore, of resources allocation within the household) may introduce biases at the descriptive level (interpretation of empirical results) and at the normative level (optimal taxation of households). Further emphasis on the bargaining process in which men and women interact can shed greater light on how individual incomes turn into household outcomes.

The main differences between couples (or more generally families) and other groups are that (i) a large degree of altruism usually takes place within the couple, and (ii) spouses usually have more occasions and willingness to share information. In riskless situations, Chiappori (1988) assumes that the utility of a family is a weighted average of the utilities of its members; the (endogenous) weights depend on all individual characteristics and reflect the respective bargaining powers of each member in the household. Chiappori's approach amounts to assume that the negotiation leads to a Pareto-optimal solution, which is consistent with any efficient negotiation process. The weights of each spouse's utility are then called Pareto weights. If the Pareto weights are constant (i.e., do not depend on any individual or family characteristics such as wages or individual wealths), then the family can be represented by a single standard utility function. This corresponds to the (above mentioned) unitary approach, which ignores the various decision-making processes and transactions occurring among the household members. Unitary models imply the *income pooling condition*: decisions made by the family should not be affected by the source of income or wealth.

On the contrary, if the Pareto weights are not constant, then bargaining powers change with individual wages or wealths. In this case, there is no simple and intuitive relation between the spouses' and the couple risk aversions. Income pooling has been repeatedly rejected empirically in different cultural contexts (Vermeulen 2002). Therefore, more and more studies (both theoretical and empirical) concerned with couple decisions in a deterministic environment are now written within the collective framework (à la Chiappori).

When risk dimensions are involved in the decision process, most of the literature still relies on the unitary approach. Among the very few exceptions is Mazzocco (2004) who shows that, in a collective model, an increase in the degree of risk aversion of one household member may induce the household to take more risk (see also Donni 2003). This counter-intuitive phenomenon results from the opposing impacts of individual degrees of risk aversion on individual decisions and Pareto weights. Indeed, the Households Retirement Survey data show that the risk aversion of couples in which the woman's risk aversion is very high, and is a U—shaped function of the man's risk aversion.

This article reports on an experimental test of couple decision-making under risk. Couples are presented with tasks involving binary choices between a lottery and a sure payoff. In the first part of the experiment, spouses are separated and choose independently. In the second part of the experiment, male spouses rejoin their partner and they make joint decisions. Couples are video-recorded while interacting and discussing to make joint decisions. We estimate both the spouses and the couples' degrees of risk aversion, we assess how the risk preferences of the two spouses aggregate when they make risky decisions, and we shed light on the dynamics of the decision process that takes place when couples make risky decisions. We find that, far from being fixed, the balance of power within the household is malleable. In most couples, men have, initially, more decision-making power than women but women who ultimately implement the joint decisions gain more and more power over the course of decision making.

The road map of the article is as follows. Section 2 describes our experimental design. In Sect. 3, we first provide ordinal measures of couples and spouses risk aversion, and then we investigate the dynamics of power balance using discrete choice models techniques. A quantitative analysis of the couples discussions is provided in Sect. 4. Finally, Sect. 5 concludes.

2 Experimental design

We elicited measures of risk aversion by means of *choice bracketing procedures*, also referred to as investment series. In each step of the bracketing procedure, the decision maker (either an individual or a couple) had to choose between a safe and a risky alternative. Risky alternatives were simple monetary lotteries, modeling the toss of a fair coin, i.e., yielding a low (respectively high) payoff with probability 1/2. Potential payoffs and probabilities were always known to the decision makers and, in a given bracketing procedure, the safe alternative was a sure amount ranging from the low outcome of the lottery to the high outcome of the lottery. At the end of the experimental session, one of the steps was randomly selected for payoff, and the decision maker's chosen option was then played out as the reward.[4] All details concerning the bracketing procedures and the lotteries are to be found in the Supplementary Material.[5]

Our elicitation method has two main advantages: First, we expect it to provide reliable estimates of risk aversion due to the simplicity of the task and the transparency of the incentives to respond truthfully; Second, it enables us to directly infer a risk attitude from the pattern of the decision maker's responses in a given investment series (see Sect. 3.1). The main disadvantage of our elicitation instrument is that it cannot be used to make inferences about non-EUT models of choice behavior. As we restrict probabilities to 1/2, we cannot use the decision maker's responses to make inferences

[4] The random lottery incentive system avoids income and house money effects, and it has become the almost exclusively used incentive system in individual decision-making experimental studies today. Holt (1986) argued that if subjects do not separate each task under this incentive system, then it would lead them to a different behavior from that in a single-shot task. However, several experimental studies demonstrate that the random lottery incentive system induces an almost identical behavior to that in a single-shot task (Starmer and Sugden 1991; Cubitt et al. 1998; Hey and Lee 2005).

[5] Our elicitation method is remotely related to the *Random Lottery Pairs* design which has been generally used to test the predictions of EUT (see, among others, Hey and Orme 1994). The main differences are that probabilities were always equal to 1/2 in our design, and one of the two alternatives was a safe option.

about probability weighting, which plays a major role in rank-dependent alternatives to EUT. Consequently, we default to thinking of risk attitudes as synonymous with the properties of the utility function, consistent with traditional EUT representations.

2.1 Experimental sessions and participants

Seven experimental sessions were carried out from January 2005 to February 2005. Subjects were recruited from the city of Jena (Germany) via local newspaper advertisements, through community groups, and using posters in the city center. Session sizes varied from 2 to 4 couples, and were held at the experimental economics video laboratory of the Max Planck Institute of Economics in Jena. In recruiting, we required all individuals to be over 30, to be living with their partners and to have been together as a couple for at least one year. We recruited 22 couples for our experiment. They answered to a total of 3,828 lotteries (either individually or with their spouse). At the beginning of the experiment, we asked a few warm-up questions to the spouses separately about themselves and about the couple (see step 1 of Sect. 1 in the experimental procedures sum-up below). The main characteristics are briefly summarized now.

Average payoffs were just above 50 € per individual—more than five times the median hourly post-tax wage for an adult working in the former East Germany in 2005. Ages ranged from 21 to 64, with a mean of 43.[6] Approximately 73% of individuals stated that they were married to their current partner, and all the couples in our sample were heterosexual. On average, couples had been together for 15 years (median of 17), with a maximum of 42 and a minimum of less than 1.[7] Interestingly, the union duration stated by women is on average 1 year more than the duration stated by men, with a maximum difference of 12 years. This difference may be explained by the fact that the man only considered marriage duration, whereas the woman considered the total duration, including the period they were living together before they got married. On average, couples had 1.3 children together. In addition, the women (men) had on average 0.3 (respectively 0.5) children from previous union(s). These figures are quite representative of the German population (see Lechner 2001).

2.2 Progress of an experimental session

Before entering the video laboratory, couples were reminded that decisions would be implemented on computers (this information had already been provided in the invitation mail), and they were told that they could ask for help at any point in time during the experimental session. Couples were also informed that the session would take place in a video laboratory and that part of the session would be video-recorded.[8]

[6] One couple was below the required age of 30 years. Both were students aged 21–22.

[7] Only the couple of students had a union duration of less than one year.

[8] Couples were also told that if they did not feel like being recorded, then they could leave immediately and that they would get a compensation of 20 € per person. All couples decided to stay and take part in the experiment.

Finally, it was mentioned to the couples that the session would consist of several parts (no details concerning the different parts were provided at that point of time) and that instructions for each part would be delivered in due time.

On entering the video laboratory, couples were separated: each male entered one of the odd numbered cabins and each female entered one of the even numbered cabins.[9] The experiment involved two sections. The first section was conducted with the two spouses located in different cabins; pairs then rejoined each other for the second section.

The first section of the experiment started with the elicitation of the participants' socioeconomic characteristics (level of education, post-tax monthly salary, etc.). Next, the separated subjects had to estimate their influence on the couple decision in every day life situations. After answering this questionnaire, each subject was endowed with 40 €. Finally, the separated subjects went through six investment series: in the first three series, separated subjects had to invest part or all of their own endowment into risky options, whereas during the last three series each subject had to invest part or all of the couple endowment into risky options. Before going through the six series of risky investments, subjects were told that they would have to go through 12 investment series and that each of their answers could possibly determine their payoff.[10] The subjects were given details of how the payout procedures would operate only at the end of the experiment.

In the second section of the experiment, couples made choices jointly, and this section has been video-recorded. Male spouses were asked to join their female partners in their cabin, and choices were made on the computer previously used by the female spouse. Couples went through six investment series. They had the possibility to discuss but no specific instructions as to how the couple decisions should be made were provided (and no explicit time limit was given). Most couples went through the six series of risky investments in less than 15 min, which indicates that agreements were quite easily reached. Except for five couples, the female spouse always physically entered the couple decisions into the computer. It is rather unsurprising that in most cases women implemented the couple decisions since the second section of the experiment took place in the women cabins and couple decisions were made on the same computers women used to make their individual decisions.[11] Although the

[9] The experimental economics video laboratory of the Max Planck Institute of Economics in Jena comprises eight soundproof cabins. Each cabin provides in- and out-put for video and audio signals. In addition, each cabin is equipped with a personal computer.

[10] Payoff-relevant investments were preceded by a training series of 10 investments.

[11] The likelihood that the man implements the couple decisions increases with the income difference between the two spouses. In the five couples in which the man holds the mouse, the average income difference is 3.8 categories. The average income difference is only 1.23 categories in the 17 couples in which the woman holds the mouse. This difference is statistically significant at the 10% level. In addition, men who hold the mouse are on average 0.4 levels more educated than their wives, and women who hold the mouse are on average 0.47 levels more educated than their husbands. This difference is not statistically significant (p-value > 0.1).

assignment of the mouse is not induced experimentally,[12] our experimental design clearly favors women over men for the control of the mouse.

The incentive system was as follows. In the beginning, one of the two spouses had to randomly draw a card from a pile of five cards, one card being numbered one, two cards being numbered two, and two cards being numbered three. If the card numbered one was randomly drawn then the payoff-relevant decision was determined separately for each spouse. Indeed, the payoff-relevant decision for each spouse was randomly selected among the decisions in which the spouse invested part or all of his/her endowment. The male spouse went back to his cabin, and each spouse's paid decision was determined according to two random draws, one random draw to determine the series (series 1, 2, or 3) and the other random draw to determine which decision in the series. If a card numbered two was randomly drawn, then the payoff-relevant decision for the couple was determined. First, a random draw decided whether one of the female or one of the male decisions to invest the couple endowment would be paid, and second, two additional random draws were made to select the series (series 4, 5, or 6) and the decision in the randomly selected series. If a card numbered three was randomly drawn, then the payoff-relevant decision for the couple was determined. Two random draws were made to select the series (series 7, 8, 9, 10, 11, or 12) and the decision in the randomly selected series.

Our payoff scheme incentivizes subjects to truly reveal their individual preferences in the first part of the experiment unless spouses expect their choices to be undone later at home and/or spouses try to efficiently coordinate their choices with each other. The first possibility is a natural consequence of the fact that our subjects have an ongoing relationship with each other which implies that choices made during the experiment can be undone when spouses go home. Any experimental study on intra-household decision making suffers from this weakness, and we do not know of a satisfactory procedure to handle it. Still, we observed that in all cases where spouses collected money from having invested their personal endowment, earnings were placed in individual wallets immediately after payment (in those cases, spouses were paid in private and not informed about each other's earnings).[13] Although this evidence is not fully convincing, it suggests the absence of binding agreements on individual earnings. The second possibility relates to the fact that individual incentives in the first three investment series differ, at least theoretically, from those in the next three investment series. Spouses might choose more risky options in investment series 1, 2, and 3 than in investment series 4, 5, and 6 as a result of coordination attempts with their partner since only personal endowments are invested. It seems rather unlikely that such coordination attempts took place. Indeed, spouses were in separate soundproof cabins in the first section of the experiment not being aware of the decision tasks faced by their partner. More importantly, details of how the payout procedures operate were

[12] Given the nature of our subject pool, we feared that the assignment of the mouse to a specific spouse would be a rather intrusive procedure. In a follow-up study, we investigate an alternative procedure in which both spouses enter the couple decisions into the computer.

[13] On the contrary, when spouses earnings were derived from investments of the couple endowment, one spouse always collected the joint earnings as a result of mutual consent even though these earnings had been divided by the experimenters into two equal shares.

provided only at the end of the experiment. Before the start of the experiment, subjects were only told that each of their choices could possibly determine their payoff.

The computer screens that subjects saw while going through the two sections of the experiment have been translated (see the Supplementary Material). Additional material of the experimental sessions, such as the written instructions and the payment procedures, is available on request from the authors. Below, we summarize our experimental procedures.

Experimental procedures

Section 1 of the experiment: Spouses are separated

In step 1, each spouse is asked to answer questions concerning his/her personal characteristics as well as concerning the couple characteristics. In the last three steps, each spouse goes through several investment series. In each series, the spouse has to invest a certain amount of money either in a lottery, modeling the toss of a fair coin, or in a sure payoff. Sure payoffs range from the low outcome of the lottery to the high outcome of the lottery.

- **Step 1. Characteristics of the individual/couple**: First, each spouse is asked to answer questions concerning his/her personal characteristics (age, job status etc.). Second, each spouse is asked to answer questions concerning his/her financial status (income, real estate etc.). Finally, the decision-making power of each spouse in some of the couple decisions is elicited.
  After answering all the questions, each spouse collects 40 € as a reward.
- **Step 2. Training investment series**: Each spouse goes through an investment series which is not payoff-relevant. Each investment decision consists in investing 50 €.
- **Step 3. Investment series 1, 2, and 3**: Each spouse goes through three payoff-relevant investment series. In the first series, each spouse invests 20 out of the 40 € he/she collected. In the second and third series, each spouse invests the entire 40 €.
- **Step 4. Investment series 4, 5, and 6**: Each spouse goes through three payoff-relevant investment series. In the first series, each spouse invests 40 out of the 80 € the couple collected. In the second and third series, each spouse invests the entire 80 €.

Section 2 of the experiment: Spouses are together

In step 5, the couple goes through three investment series. In each series, the couple has to invest a certain amount of money either in a lottery, modeling the toss of a fair coin, or in a sure payoff. Sure payoffs range from the low outcome of the lottery to the high outcome of the lottery. In step 6, the couple goes through three investment series, including 3 questions each. In each series, the couple has to invest a certain amount of money either in a lottery (specific to each question), modeling the toss of a fair coin, or in a sure payoff (which does not vary within a series). In each series, the lottery proposed in the second question depends on the answer to the first question, and the lottery proposed in the third question depends on the answer to the first and second questions.

- **Step 5. Investment series 7, 8, and 9**: The couple goes through three payoff-relevant investment series. In the first series, the couple invests 40 out of the 80 € the couple collected. In the second and third series, the couple invests the entire 80 €. The figures in series 7 (respectively 8, 9) are exactly the same as the ones in series 4 (respectively 5, 6).
- **Step 6. Investment series 10, 11, and 12**: Both the amount invested and the sure payoff are 80 €. In the first series, the couple may loose half of the 80 € in the worst case and increase their payoff up to 140 € in the best case. The expected payoff of all lotteries is 90 €, and the variability of the payoff is increased if the couple previously selected the lottery, decreased if they previously selected the sure payoff. The second series is similar, except that the safe payoff is 90 € (all amounts in the first question are increased by 10 €). In the third series, there is no risk of any loss (the payoff in the worst case is 80 €), and instead of increasing/decreasing the variance, only one outcome is increased/decreased depending on the answer to the previous question.

3 Results

In this section, we first assess the decision makers' degrees of risk aversion by relying on an ordinal approach. We restrict ourselves to the choices made by the spouses separately in investment series 4–6 (Step 4), and to the choices made by the couples in investment series 7–9 (Step 5). Indeed, in investment series 4 (respectively 5 and 6) each spouse is assigned to the same lottery, and this lottery is also the one used in investment series 7 (respectively 8 and 9) when both spouses decide jointly. Therefore, the individual and couple answers can be compared directly. Second, we rely on a cardinal approach, and we assume that spouses are expected utility maximizers with a constant absolute risk aversion utility function. We use individual choices in investment series 1–6 and joint choices in investment series 7–12 to study the evolution of the balance of decision-making power within the household. Both in the cardinal and in the ordinal approach, we allow the choices to violate the assumption that preferences are monotonic with respect to money.

3.1 Man, woman, and couple risk attitudes

In each investment series j, the decision maker faces 11 choices ($i = 1, \ldots, 11$) between a lottery L_j and a sure payoff $S_j(i)$. The lottery yields the low payoff $S_j(11)$ and the high payoff $S_j(1)$ with equal probabilities. The sequence of sure payoffs is given by $S_j(i) = S_j(11) + \left(\frac{11-i}{10}\right)\left(S_j(1) - S_j(11)\right), i = 1, \ldots, 11$. As the expected value of the lottery equals $S_j(6)$, a risk-neutral decision maker is indifferent between the lottery and $S_j(6)$.

The set of choices made by a decision maker facing investment series j is inconsistent if monotonic and transitive preferences cannot rationalize those choices. Table 1 shows, for each investment series, the relative frequency of inconsistent sets of choices for women, men, and couples. Altogether, there were 23% (respectively 13 and 9%) of inconsistent sets of choices for women (respectively for men and for couples). Most of the women inconsistent sets of choices were made in the early investment series, which suggests that women need more than one training investment series to get acquainted with the task.[14]

A consistent set of choices is characterized by a unique switching point, $i \in \{0, \ldots, 11\}$: for a given investment series j, decision maker k in class i prefers lottery L_j to all deterministic amounts lower than or equal to $S_j(i+1)$ and prefers all amounts larger than or equal to $S_j(i)$ to lottery L_j. In this case, we denote by $\succ_k$ the risk preference relation of decision maker k, uniquely defined on the set

[14] Given the nature of our risk elicitation mechanism and subject pool, the amount of inconsistencies observed is rather low. Experimental studies which have employed a multiple price list (MPL) risk elicitation mechanism report a proportion of inconsistent preferences usually larger than 10%. Among others, both Eckel and Wilson (2004) and Holt and Laury (2002) report 13% of inconsistent preferences (this proportion drops to 5.5% when payoffs are scaled by a factor of 50 or 90 in the latter study), and Bruner et al. (2008) report 20% of inconsistent preferences. Note that these studies used college students as experimental subjects (who are generally quick on understanding their task in the experiment) and that the MPL mechanism allows subjects to see all choices in one frame which might lead them to make more consistent choices than they would otherwise (but seems to induce a framing effect).

Table 1 Relative frequencies of inconsistent series of choice

Investment series	Woman	Man	Couple
1 (Woman and Man)	9/22	2/22	
2 (Woman and Man)	7/22	3/22	
3 (Woman and Man)	6/22	3/22	
4 (Woman and Man)/7 (Couple)	2/22	3/22	2/22
5 (Woman and Man)/8 (Couple)	2/22	2/22	1/22
6 (Woman and Man)/9 (Couple)	4/22	4/22	3/22

Table 2 The 12 sets of consistent choices

Switching point	Set of consistent choices	Investment series: Woman, Man; Couple					
		1	2	3	4; 7	5; 8	6; 9
0	$L_j \succ_k S_j(1)$	1, 0					
1	$S_j(1) \succ_k L_j \succ_k S_j(2)$						
2	$S_j(2) \succ_k L_j \succ_k S_j(3)$						
3	$S_j(3) \succ_k L_j \succ_k S_j(4)$	1, 1					
4	$S_j(4) \succ_k L_j \succ_k S_j(5)$	0, 1	0, 1	0, 1	1, 0; 0	1, 0; 0	
5	$S_j(5) \succ_k L_j \succ_k S_j(6)$	0, 3	1, 2	5, 1	1, 1; 2	1, 1; 2	0, 0; 1
6	$S_j(6) \succ_k L_j \succ_k S_j(7)$	5, 7	1, 3	6, 8	4, 7; 6	4, 6; 2	6, 7; 5
7	$S_j(7) \succ_k L_j \succ_k S_j(8)$	1, 0	2, 6	1, 3	4, 5; 4	5, 3; 0	4, 3; 4
8	$S_j(8) \succ_k L_j \succ_k S_j(9)$	3, 5	5, 3	1, 4	5, 2; 7	2, 6; 13	3, 4; 5
9	$S_j(9) \succ_k L_j \succ_k S_j(10)$		2, 3	0, 1	2, 2; 1	2, 2; 4	0, 1; 4
10	$S_j(10) \succ_k L_j \succ_k S_j(11)$	1, 0	2, 0	1, 1	0, 1; 0	2, 1; 0	1, 1; 0
11	$S_j(11) \succ_k L_j \left(L_j - \text{OR}\right)$	1, 3	2, 1	2, 0	3, 1; 0	3, 1; 0	4, 2; 0

$\{L_j, S_j(i), i = 1, \ldots, 11\}$ by his/her set of replies to series j. More specifically, $S_j(i) \succ_k L_j$ means that decision maker k prefers the sure payoff $S_j(i)$ to the lottery L_j. Given the construction of the series, the classes are ranked by increasing risk aversion, which defines an ordinal measure of risk aversion.

Out of the 2^{11} potential sets of choices in a given investment series, only 12 are consistent, which defines 12 ordered classes of risk aversion. They are represented in Table 2, together with the frequencies of observed answers in each series, for women, men, and couples.

We observe that a significant proportion of individuals (especially women) are willing to receive *always* less money just for the benefit of avoiding any risk (15 out of $22 * 6$ women-series and 8 out of 22×6 men-series). We denote by Locally Opposed to Risk for lottery L_j (L_j-OR), those decision makers who consistently prefer any sure payoff $S_j(i), i = 1, \ldots, 11$, to lottery L_j in investment series $j. L_j-$OR preferences are never shared by both spouses in a couple nor by the two spouses together, i.e., no L_j-OR individual was able to convince his/her spouse. We observe that only one respondent (a female respondent) is L_j-OR for the six series L_j, $j = 1, \ldots, 6$ which implies an infinite level of risk aversion.

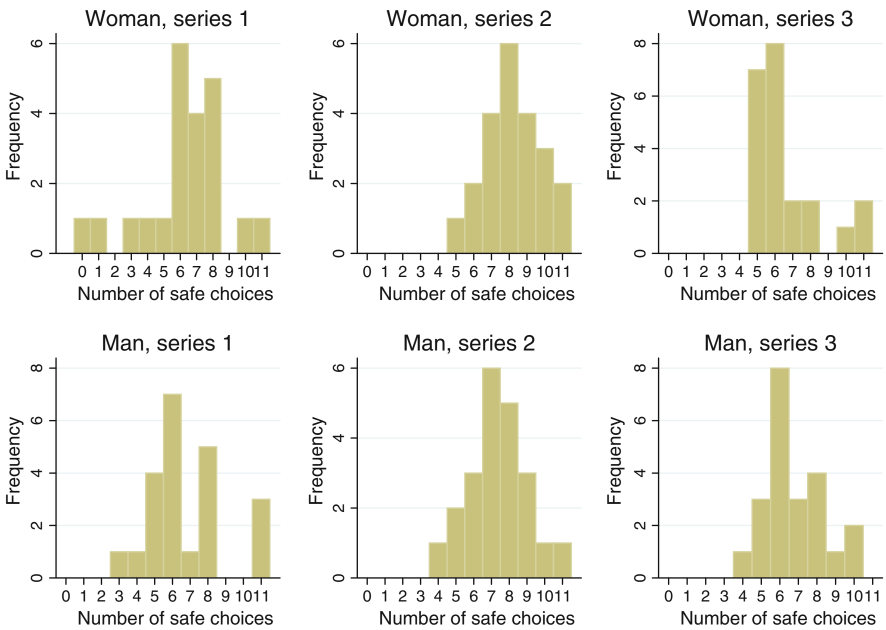


Fig. 1 Empirical distributions of safe choices, individual money

In order to take into account inconsistent sets of choices, the total number of "safe" choices is used as an indicator of risk aversion: for a given investment series, the measure of risk aversion is given by the frequency of choices where the decision maker picks the sure payoff instead of the lottery (in case of consistent series, the same measure is obtained with the switching point). Figure 1 shows the empirical distributions of safe choices in the three investment series concerned with individual money, separately for women and men. Both for women and men, the distribution is more spread for the first series, and some respondents appear extremely risk lovers. This may reflect the fact that one training series was not enough and that some respondents answered randomly in the first series because they were not acquainted with the task.

Figure 2 shows the empirical distributions of safe choices in the three investment series concerned with couple money, separately for women, men, and couples (spouses together). In all three investment series, the distribution of couple choices is more concentrated than the distribution of spouse choices.

Both figures suggest that women are slightly more risk averse than men, and that men and women answers are more heterogeneous than couples answers. This is confirmed in Table 3, which shows the average frequencies of safe choices for the woman, the man and for the couple in the different investment series, as well as their differences.

Concerning investment series 1–3 (individual money), women and men answers cannot be directly compared since the amounts involved were generated randomly, independently for the woman and for the man. Table 3 supports the idea that individuals (especially women) answered more randomly in series 1, since the average

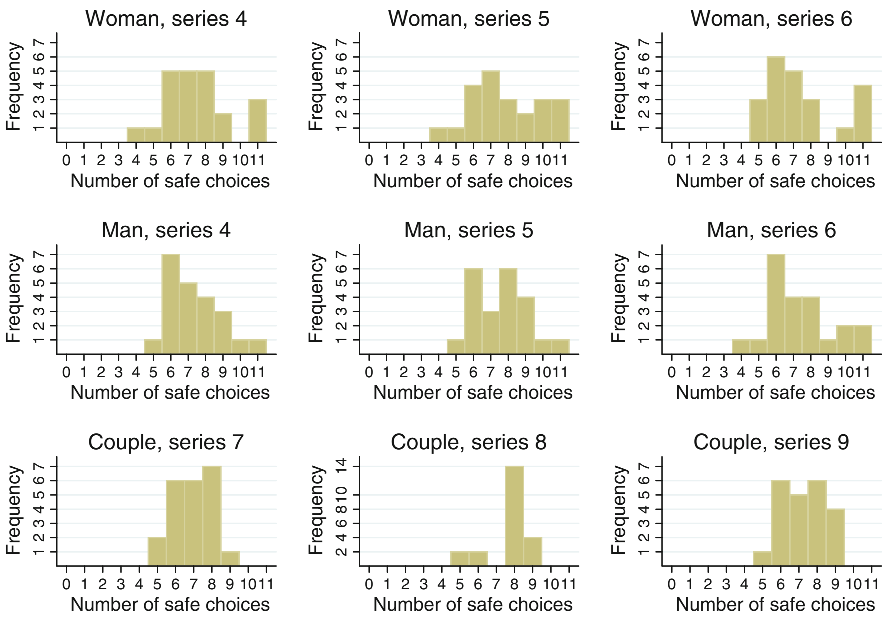


Fig. 2 Empirical distributions of safe choices, couple money

Table 3 Average frequencies of safe choices (standard deviations in parentheses)

Investment series	Woman	Man	Couple	Differences		
				Woman–Couple	Man–Couple	Woman–Man
1	6.27 (2.57)	6.73 (2.18)				−0.50 (3.69)
2	8.23 (1.60)	7.36 (1.68)				0.86 (2.57)
3	6.59 (1.89)	6.77 (1.60)				−0.18 (2.34)
4/7	7.50 (1.87)	7.36 (1.53)	6.95 (1.09)	0.55 (1.84)	0.41 (1.37)	0.14 (2.28)
5/8	7.86 (2.03)	7.59 (1.53)	7.73 (1.16)	0.14 (2.21)	−0.14 (1.49)	0.27 (2.86)
6/9	7.45 (2.06)	7.36 (1.89)	7.27 (1.20)	0.18 (1.92)	0.09 (2.14)	0.09 (3.00)

frequency of safe choices is lower[15] and the standard deviations (of individual answers and of their differences) are larger for series 1 than for series 2 and 3.

Concerning investment series 4–6 (couple money), women are (slightly) more risk averse than men.[16] Moreover, the average couple tends to be *less risk averse* than its

[15] The average would be 5.5 for pure random choices, which is lower than the observed average of 7–8 for the other series.

[16] Individual answers in investment series 4, 5, and 6 could reflect more the preferences of the couple than individual answers in investment series 1, 2, and 3 since individuals invest the couple's money. Our experimental data do not support this argument. Indeed, the correlation factor between the average degree of risk aversion in Step 3 and the average degree of risk aversion in Step 4 is highly positive for both spouses: It equals 0.682 for the woman and 0.574 for the man. These correlation factors are at least as positive as those obtained when comparing the average degrees of risk aversion of different investment series in the same step for a given spouse: In Step 3, the average correlation factor equals 0.419 for the woman and

average members (Munro et al. 2008 report a similar finding). Indeed, the average measure of risk aversion for couples is systematically lower than the average measure of risk aversion for women, and it is lower than the average measure of risk aversion for men in two out of the three investment series. The variance of the difference between men and women measures of risk aversion increases over time. The variance of the difference between the couples and the men measures of risk aversion also increases over time contrary to the variance of the difference between the couples and the women measures of risk aversion which exhibits no monotonic pattern. In conclusion, after controlling for the average difference between women, men, and couples measures of risk aversion, the distance between couples and men measures of risk aversion increases, whereas the distance between couples and women measures of risk aversion remains constant. This suggests that the relative decision-making power of the woman when the couple is facing a unique decision increases over time.

In the next section, we study the evolution of the balance of power within the household, and we relate our findings to the empirical/theoretical literature on intra-household decision-making.

3.2 The balance of power within the household

We now assume that the preferences of any decision maker in our sample can be represented by a utility function with constant absolute risk aversion (CARA) for money $x > 0$ (this choice is based on initial tests for various standard utility functions). Decision maker k's utility function is, therefore, given by $V_k(x) = V(x; \theta_k) = (1 - \exp(-\theta_k x))/\theta_k$ where θ_k is the level of absolute risk aversion. In investment series j, the utility of the safe alternative $S_j(i)$ is $V_k(S_j(i))$, while the expected utility of the lottery is given by $\mathbb{E}\left[V_k\left(L_j\right)\right] = \left(V_k(S_j(1)) + V_k(S_j(11))\right)/2$.

Experimental evidence suggests that stochastic variation is an essential feature of decision-making behavior.[17] Therefore, we assume that choices made by decision maker k in series j are consistent with an absolute risk aversion level $\tilde{\theta}_{jk}$, a random variable modeled as $\tilde{\theta}_{jk} = a + \mu_k + \delta_j + \varepsilon_{jk}$ where a is a constant term equal across series and decision makers, μ_k is an decision-maker-specific term with zero mean, δ_j measures potential systematic deviations specific to series j and ε_{jk} is a normal error term with zero mean and variance specific to series j. The variance of $\mu_k + \varepsilon_{jk}$ is denoted by σ_j^2. We employ an interval regression method to estimate the absolute risk aversion

Footnote 16 continued
0.288 for the man; In Step 4, the average correlation factor equals 0.718 for the woman and 0.467 for the man. On the contrary, correlation factors between the average degree of risk aversion of the woman and the average degree of risk aversion of the man are either weakly positive or negative depending on the considered step(s): In Step 3 (respectively Step 4), the correlation factor between the average degrees of risk aversion of the two spouses equals −0.125 (respectively −0.148); The correlation factor between the average degree of risk aversion of the woman in Step 3 and the average degree of risk aversion of the man in Step 4 equals –0.046; The correlation factor between the average degree of risk aversion of the woman in Step 4 and the average degree of risk aversion of the man in Step 3 equals 0.118.

[17] In controlled experiments in which subjects have confronted exactly the same choice problem on two occasions, the proportion of choice reversals is between 10 and 30%. See Table 1 in Stott (2006, p. 105).

Table 4 Regression of the couple CARA level on the individual weighted average CARA levels

Series	Case 1 (22 couples)			Case 2 (17 couples)		
	Woman	Man	Joint	Woman = UC	Man = No UC	Joint
7	0.13 (0.14)	0.20 (0.12)	3.77 (0.15)	0.22 (0.11)	0.19 (0.16)	4.32 (0.11)
8	0.03 (0.55)	$0.25^{▲}$ (0.00)	$8.41^{\triangle}$ (0.02)	0.07 (0.34)	$0.24^{▲}$ (0.01)	$8.52^{\triangle}$ (0.02)
9	$0.24^{\triangle}$ (0.02)	$0.40^{\triangle}$ (0.02)	$8.94^{\triangle}$ (0.02)	0.34* (0.06)	0.35* (0.06)	$6.69^{\triangle}$ (0.04)
10	$0.34^{\triangle}$ (0.02)	$0.56^{\triangle}$ (0.02)	$8.10^{\triangle}$ (0.02)	$0.93^{▲}$ (0.00)	0.32* (0.08)	$18.06^{▲}$ (0.00)
11	$0.27^{\triangle}$ (0.05)	0.21 (0.31)	4.16 (0.13)	$0.66^{▲}$ (0.00)	0.01 (0.97)	$12.11^{▲}$ (0.00)
12	0.11 (0.16)	0.14 (0.26)	2.72 (0.26)	$0.24^{\triangle}$ (0.05)	0.07 (0.63)	3.86 (0.15)

Notes: ▲, $\triangle$, $*$ indicate significance at 1, 5, and 10% level, respectively
Two-sided p-values in parentheses

levels. With a normal distribution, we obtain an unbiased estimate $\hat{\theta}_{jk}$ of $\tilde{\theta}_{jk}$ which, in turn, enables us to compute an unbiased estimate of θ_k. Average individual risk aversion levels are based on the weighted mean of the absolute risk aversions estimated for each of the first six investment series: $\hat{\theta}_k = \sum_{j=1}^{6} \left(\left(\hat{\theta}_{jk} - \hat{\delta}_j \right) / \hat{\sigma}_j \right) / \sum_{j=1}^{6} \left(1/\hat{\sigma}_j \right)$. The weights, inversely proportional to the estimated standard deviation of the residuals, improve the efficiency of the estimate when the variance of the residuals varies across series.

In order to study the evolution of the balance of power within the household, we use the individual risk aversions estimated from choices made in the first section of the experiment to explain the couple risk aversions estimated from choices made in the second section of the experiment. Concretely, we regress the couple CARA level on the respective spouses weighted average CARA levels for the different investment series of the second section of the experiment. Regressions are conducted on the entire sample (Case 1) and on the restricted sample composed of couples in which women ultimately implemented the joint decisions (Case 2). Our regression results, for each investment series of the second section of the experiment, are displayed in Table 4 where, in columns "Joint," we report the χ^2-statistic of the likelihood test for the null hypothesis that both coefficients are zero.[18]

According to the regression results for Case 1, neither the man nor the woman weighted average CARA level significantly influences the couple CARA level in the seventh investment series. This observation suggests that spouses needed to acquire some experience in making investment choices jointly. The influence of the man risk aversion on the couple risk aversion is highly significant in investment series 8, it becomes significant in investment series 9 and 10, and it is non-significant in the last two investment series. On the contrary, the influence of the woman risk aversion on the couple risk aversion is not significant in investment series 8, but it is significant in investment series 9, 10, and 11. These observations indicate that women gain more and more power over the course of decision making to the detriment of men.

[18] In the Table 4, we report separate regressions for the different investment series. Running a single regression on the pooled data with interaction terms between investment series and individual measures of risk aversion produces qualitatively equivalent results.

The balance of power within the household evolves even more clearly in couples where women ultimately implement the joint decisions. According to the regression results for Case 3, the influence of the man risk aversion on the couple risk aversion is highly significant in investment series 8, it becomes weakly significant in investment series 9 and 10, and it is non-significant in the last two investment series. On the contrary, the influence of the woman risk aversion on the couple risk aversion is not significant in investment series 8, it is weakly significant in investment series 9, it becomes highly significant in investment series 10 and 11, and it is still significant in the last investment series. In addition, the marginal impact of the woman risk aversion is identical to the marginal impact of the man risk aversion in investment series 9 (where both are weakly significant), but it is three times higher in investment series 10 (where the man risk aversion is weakly significant, and the woman risk aversion is highly significant). Finally, the marginal impact of the woman risk aversion decreases in the last two investment series (where the man risk aversion is not significant).[19]

Our estimation results suggest that in most couples men have, initially, a stronger decision-making power than women. This observation is in line with earlier studies which show that men tend to have more say in economic decision making and readier access to financial resources than their wives (e.g., Pahl 1995; Kirchler et al. 2001). More surprisingly, our estimation results also suggest that, far from being fixed, decision-making powers are malleable. Women who ultimately implement the joint decisions gain more and more decision-making power over the course of an experimental session.[20] Our evidence on the malleability of the balance of power within the household is clearly at odds with the collective approach based on static models of intra-household resource allocation that obeys a Pareto-efficient sharing rule. However, recent extensions of the collective approach to intertemporal settings where the assumption of intra-household commitment is relaxed allow for temporal variations in relative decision-making power (see, e.g., Mazzocco 2007). In a more drastic departure from the early collective approach, Basu (2006) discusses a model of household behavior under no-commitment where Pareto weights depend on choices variables. As a consequence, the Pareto weight assigned to each spouse is endogenous to the household decision-making process, i.e., the household balance of power is endogenously determined. The assumption of an endogenous intra-household balance of power has received empirical support by Lancaster et al. (2006) whose findings using household-level unit record data sets from India indicate that decision-making powers are determined jointly with the expenditure outcomes.

Although these recent extensions of the collective approach assume that intra-household allocations influence the balance of power within the household, they are not fully consistent with our experimental results since they predict that house-

[19] Regression of the couple safe choices on the individual safe choices leads to similar results for the subset of investment series where the couple and individual choices can be directly compared. See Table 6 in the Appendix.

[20] Note that we do not argue that women gain more decision-making power only because they ultimately implement the joint decisions, as it could well be that women ultimately implement the joint decisions because they initially have substantial decision-making power. We do, however, believe that there is a relative gain in decision-making power obtained by the spouse who ultimately implements the couple decisions.

hold decisions shape relative decision-making power gradually. On the contrary, we observe that power patterns within the household vary quickly, which suggests that the household decision process heavily depends on the context. This observation might cause discomfort to economists, but it is fully in line with the social-exchange formulation of family power dynamics in social psychology and recent evidence from the marketing literature. According to the social-exchange perspective, decision-making power resides in the characteristics of relationships and not in personal traits which implies that power patterns within families will vary from time to time and with task characteristics (Beckman-Brindley and Tavormina 1978). Su et al. (2003) examine the dynamics of spousal behavioral interactions in a questionnaire study based on a sequence of family purchase decisions. Their results suggest that spousal purchase–decision processes are adaptive. In early purchase episodes, few spouses are willing to yield when faced with coercion while, in later purchase episodes, spouses get their way by strong means of influence. The authors speculate that there is a learning curve underlying the spousal decision–behavior dynamics.

Before concluding, we provide in the next section a quantitative analysis of the couples discussions which corroborates our previous analyzes of the choice data.[21]

4 Quantitative analysis of the discussions within the couple

In this section, we present a basic quantitative analysis of the discussions that couples had while answering investment series 7–12 (a content analysis is beyond the scope of the present study). Two undergraduate native raters independently watched the videos of 17 couples several times and evaluated the talk duration of each spouse, i.e., the amount of time spent by each spouse talking to the other spouse about which joint decision to implement.[22] Both raters were instructed to exclude from talk duration the amount of time spent by each spouse discussing topics not closely related to the experiment. Table 5 shows the individual talk durations per investment series as well as the ratio between the woman talk duration (WTD) and the couple talk duration (CTD) for each of the 17 couples.

In all these couples except two, the man was always arguing more about which joint decision to implement than the woman. Unsurprisingly, both spouses talk on average more in the seventh investment series than in the latter investment series. Although there is no clear time trend in WTDs, men argue on average more in the first three investment series than in the last three investments series. It seems natural to relate the talk duration of an individual with his/her decision-making power: the more an individual is arguing the more he/she is trying to influence the joint decision (and, in most cases, he/she will probably be successful). In this respect, our quantitative analysis of the couples discussions corroborates our statistical analyzes of the choice data: the man leads the joint decision, at least initially. Our previous analyzes also suggested that the woman who ultimately implements the joint decisions gains power over the

[21] Verbal frequency measures constitute the main quantitative process measures to study family interactions in psychology. See, e.g., Jacob (1975) and the references therein.

[22] Unfortunately, five out of the 22 videos had to be discarded due to the low sound quality.

Table 5 Individual talk durations in seconds

Session	Cabin	Spouse	Investment series						Total	WTD/CTD
			7	8	9	10	11	12		
January 24, 2005	1	Man	45	40	28	35	43	25	216	0.388
7 pm		Woman	35	30	20	10	23	19	137	
January 24, 2005	2	Man	55	60	60	34	38	23	270	0.338
7 pm		Woman	40	25	18	15	29	11	138	
January 25, 2005	1	Man	29	19	11	10	24	25	118	0.433
7 pm		Woman	20	13	12	6	24	15	90	
January 25, 2005	2	Man	44	33	50	27	37	12	203	0.450
7 pm		Woman	23	18	49	17	44	15	166	
January 25, 2005	3	Man	44	38	36	14	20	37	189	0.357
7 pm		Woman	20	14	11	6	20	34	105	
January 25, 2005	4	Man	8	18	26	17	28	8	105	0.521
7 pm		Woman	16	14	17	23	28	16	114	
January 26, 2005	1	Man	25	20	22	30	31	36	164	0.416
7 pm		Woman	20	9	21	20	21	26	117	
January 26, 2005	2	Man	52	13	29	30	20	8	152	0.290
7 pm		Woman	26	8	5	6	11	6	62	
January 26, 2005	3	Man	4	3	7	3	31	17	65	0.356
7 pm		Woman	3	8	4	4	6	11	36	
January 27, 2005	2	Man	13	9	11	14	21	10	78	0.447
7 pm		Woman	13	7	7	9	14	13	63	
January 27, 2005	3	Man	70	51	42	43	34	28	268	0.396
7 pm		Woman	46	25	14	32	42	17	176	
January 28, 2005	1	Man	51	44	21	19	19	31	185	0.387
7 pm		Woman	26	19	24	15	18	15	117	
January 28, 2005	2	Man	24	34	25	24	34	25	166	0.362
7 pm		Woman	24	6	17	23	11	13	94	
February 19, 2005	1	Man	38	11	13	16	20	29	127	0.392
3 pm		Woman	24	8	10	12	9	19	82	
February 19, 2005	2	Man	48	30	22	30	30	35	195	0.449
3 pm		Woman	36	34	18	27	19	25	159	
February 19, 2005	3	Man	42	20	8	3	6	10	89	0.429
3 pm		Woman	20	7	11	8	5	16	67	
February 19, 2005	2	Man	20	18	10	6	28	20	102	0.512
5 pm		Woman	23	16	12	7	34	15	107	

Note: WTD/CTD denotes the ratio between the woman talk duration and the couple talk duration

course of decision making to the detriment of the man. We offer now a final evaluation of the impact of ultimate control on the evolution of power balance by comparing the woman relative talk duration when she has ultimate control to her relative talk duration when the man has ultimate control. Figure 3 shows the woman relative talk duration in each investment series averaged, on the one hand, over the 13 couples where the

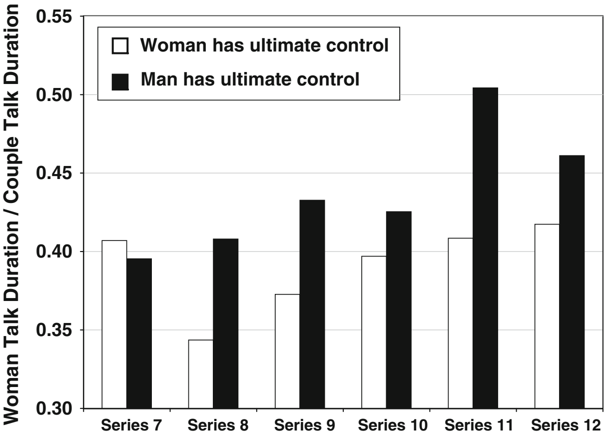


Fig. 3 Woman relative talk duration

woman had ultimate control and averaged, on the other hand, over the 4 couples where the man had ultimate control.

In the seventh investment series, whether the woman has ultimate control or not does not influence her talk duration. However, in investment series 8 and 9, a woman without ultimate control argues, in relative terms, much more than a woman who has ultimate control. A similar tendency is observed in the last part of the experimental session, i.e., in investment series 10–12. Under the natural assumption that talk duration is related to decision-making power, we again conclude that the spouse who has ultimate control gains additional influence on the decision of the couple.

5 Concluding comments

This article reports results on individual and couple choices in an experiment involving risk. Individuals and couples make binary choices between a lottery and a sure payoff. In the first part of the experiment, spouses are separated and choose independently. Individual choices express individual risk preferences. In the second part of the experiment, male spouses rejoin their partner, and they make joint choices. In most cases, the woman implemented the couple choices which express collective risk preferences. We investigate the evolution of the balance of power within the household by the individual risk aversions estimated from choices made in the first section of the experiment to explain the couple risk aversions estimated from choices made in the second section of the experiment. We find that the man is initially more successful than the woman in influencing couple choices but that the woman progressively gains power over the course of decision making, and we speculate that part of this power increase relates to the ultimate control over the implementation of joint choices.

Our evidence on the dynamics of power balance suggests that actual decision processes within the household are adaptive and depend on the context. This finding contradicts the early collective approach based on static models of intra-household resource allocation according to which the decision-making powers are regarded as exogenous to the household decision-making process. It is, however, compatible with recent extensions of the collective approach, which assume that intra-household allocations influence the balance of power within the household. More research on larger samples would be necessary to validate this preliminary finding. Similarly, larger samples would be necessary to link stated and revealed decision-making power to distribution factors (such as difference between spouses, educational levels, ages, or assets).

Acknowledgments Many thanks are due to two constructive and helpful referees whose comments improved this article considerably. We are especially grateful to one of the referees who encouraged us to develop our ideas more fully. Research assistance has been provided by Sascha Bähr, Andreas Dittrich, Hendrik Ehrhardt, Andreas Lehmann, Anika Littmann, Dirk Pogode, Christoph Röber, and Torsten Wollina. We acknowledge the financial and technical supports from the Institut National d'Etudes Démographiques, and the French National Research Agency (research program "*RiskAttitude*", ref. ANR-05-BLAN-0345-01). Valuable suggestions were provided by seminar participants at the University of Virginia, Cergy-Pontoise (THEMA), Paris I, Paris II (ERMES), and Montpellier, and at INED, as well as by the members of the *RiskAttitude* consortium. We especially thank Olivier Donni for very useful discussions about the literature on collective models.

Appendix: Dynamics of power balance in the ordinal approach

Contrary to the regression results reported in Table 4, the regression results reported in the table below are based on a subset of the observed choices, namely the individual choices made in investment series 4–6 (Step 4) and the couple choices made in investment series 7–9 (Step 5).

Table 6 Regression of couple safe choices on individual safe choices

Series	Case 1 (22 couples)			Case 2 (17 couples)		
	Woman	Man	Adjusted R^2	Woman = UC	Man = No UC	Adjusted R^2
7	0.15 (0.19)	$0.33^{\triangle}$ (0.03)	0.24	0.25 (0.12)	$0.34^{\triangle}$ (0.03)	0.36
8	0.15 (0.23)	$0.37^{\triangle}$ (0.03)	0.15	$0.38^{\blacktriangle}$ (0.01)	$0.41^{\blacktriangle}$ (0.01)	0.50
9	$0.25^{\triangle}$ (0.05)	0.11 (0.43)	0.11	0.29^{*} (0.10)	0.07 (0.70)	0.07

Notes: ▲, △, ∗ indicate significance at 1, 5, and 10% level, respectively
Two-sided p-values in parentheses

References

Basu, K. (2006). Gender and say: A model of household behavior with endogenously determined balance of power. *Economic Journal, 116*, 558–580.

Bateman, I., & Munro, A. (2005). An experiment on risky choice amongst households. *Economic Journal, 115*, C176–C189.

Bateman, I., & Munro, A. (2009). Household versus individual valuation: What's the difference? *Environmental and Resource Economics, 43*, 119–135.

Beckman-Brindley, S., & Tavormina, J. (1978). Power relationships in families: A social-exchange perspective. *Family Process, 17*, 423–436.

Bruner, D., McKee, M., & Santore, R. (2008). Hand in the cookie jar: An experimental investigation of equity-based compensation and managerial fraud. *Southern Economic Journal, 75*, 261–278.

Chiappori, P. A. (1988). Rational household labor supply. *Econometrica, 56*, 63–90.

Cubitt, R., Starmer, C., & Sugden, R. (1998). On the validity of the random lottery incentive system. *Experimental Economics, 1*, 115–131.

Donni, O. (2003). Collective household labor supply: Nonparticipation and income taxation. *Journal of Public Economics, 87*, 1179–1198.

Duflo, E. (2003). Grandmothers and granddaughters: Old age pension and intra-household allocation in South Africa. *World Bank Economic Review, 17*, 1–25.

Eckel, C., & Wilson, R. (2004). Is trust a risky decision? *Journal of Economic Behavior & Organization, 55*, 447–465.

Harrison, G., & List, J. (2004). Field experiments. *Journal of Economic Literature, 42*, 1013–1059.

Hey, J., & Lee, J. (2005). Do subjects separate (or are they sophisticated)? *Experimental Economics, 8*, 233–265.

Hey, J., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica, 62*, 1291–1326.

Holt, C. (1986). Preference reversals and the independence axiom. *American Economic Review, 76*, 508–513.

Holt, C., & Laury, S. (2002). Risk aversion and incentive effects. *American Economic Review, 92*, 1644–1655.

Jacob, T. (1975). Family interaction in disturbed and normal families: A methodological and substantive review. *Psychological Bulletin, 82*, 33–65.

Kirchler, E., Rodler, C., Holzl, E., & Meier, K. (2001). *Conflict and decision making in close relationships*. Hove: The Psychology Press.

Lancaster, G., Maitra, P., & Ray, R. (2006). Endogenous intra-household balance of power and its impact on expenditure patterns: Evidence from india. *Economica, 73*, 435–460.

Lechner, M. (2001). The empirical analysis of East German fertility after unification: An update. *European Journal of Population, 17*, 61–74.

Mazzocco, M. (2004). Savings, risk sharing and preferences for risk. *American Economic Review, 94*, 1169–1182.

Mazzocco, M. (2007). Household intertemporal behaviour: A collective characterization and a test of commitment. *Review of Economic Studies, 74*, 857–895.

Munro, A., Bateman, I., & McNally, T. (2008). The family under the microscope: An experiment testing economic models of household choice. MPRA Paper No. 8974.

Pahl, J. (1995). His money, her money: Recent research on financial organisation in marriage. *Journal of Economic Psychology, 16*, 361–376.

Starmer, C., & Sugden, R. (1991). Does the random-lottery incentive system elicit true preferences? An experimental investigation. *American Economic Review, 81*, 971–978.

Stott, H. (2006). Cumulative prospect theory's functional menagerie. *Journal of Risk and Uncertainty, 32*, 101–130.

Su, C., Fern, E., & Ye, K. (2003). A temporal dynamic model of spousal family purchase-decision behavior. *Journal of Marketing Research, 40*, 268–281.

Vermeulen, F. (2002). Collective household models: Principles and main results. *Journal of Economic Surveys, 16*, 533–564.

Southern Economic Journal 2012, 79(2), 277–299
DOI: 10.4284/0038-4038-2010.026

Risk-Taking Behavior: An Experimental Analysis of Individuals and Dyads

Cary Deck,* Jungmin Lee,† Javier Reyes,‡ and Chris Rosen§

The decision to undertake risk is often made by pairs (dyads), while much of the economics literature on risk taking focuses on the individual. We report the results of controlled laboratory experiments that compare behavior between individuals and pairs. Using the Holt and Laury (2002) procedure and a within-subjects design, we find that pair choices are largely consistent with subjects bargaining over the outcome rather than the pairs taking a more extreme stance than the individual members. Further, gender and age but not personality seem to influence relative bargaining weight. We also find that individuals are more willing to take risks after making decisions as part of a pair than beforehand. Both the personality of one's partner and nontask social interaction influence subsequent individual risk-taking behavior.

JEL Classification: C7, C9, D7, D8

1. Introduction

People routinely make decisions under uncertainty due to incomplete information. The perceived degree of uncertainty affects decisions regarding consumption, saving and investing, and the selection of warranties and insurance policies. It also impacts the decision to enter a given profession or engage in certain activities, such as crimes and extreme sports. Thus, it is no surprise that considerable attention has been given to measuring risk attitudes. To this end, researchers have used data from sources ranging from hypothetical responses in large-scale surveys (e.g., Barsky et al. 1997; Dohmen et al. 2005) to behavior in high-stakes game shows (e.g., Baltussen et al. 2008; Deck, Lee, and Reyes 2008). In the laboratory, researchers can manipulate the decision problem while obtaining complete information about the relevant payoffs and probabilities associated with a given choice. The generally observed pattern is that people tend to exhibit mild risk aversion in the laboratory.

* Department of Economics, University of Arkansas, WCOB 402, Fayetteville AR 72701, USA; E-mail cdeck@walton.uark.edu.

† School of Economics, Sogang University, 1 Sinsu-dong, Mapo-gu, Seoul 121-742, Korea; E-mail junglee@sogang.ac.kr; corresponding author.

‡ Department of Economics, University of Arkansas, WCOB 402, Fayetteville AR 72701, USA; E-mail cdeck@walton.uark.edu; E-mail jreyes@walton.uark.edu.

§ Department of Management, University of Arkansas, WCOB 407, Fayetteville AR 72701, USA; E-mail crosen@walton.uark.edu.

The authors wish to thank two anonymous referees, Philip Grossman, Charlie Holt, Robert Shupp, and participants at the Southern Economic Association Conference for helpful comments and suggestions. Deck, Reyes, and Rosen are grateful for Sutton Research Funding from the Walton College of Business Sutton Research. Lee is grateful for the financial support from the Sogang University Research Grant of 2010. This research was supported in part by NIH grant R21AG030184.

Received January 2010; accepted January 2012.

Risk-taking research has focused predominantly on choices by individuals, while many naturally occurring decisions are made in groups of various sizes. In some situations, decisions are made by large groups, such as boards or juries. In other situations, decisions are made by pairs or dyads, such as spouses or business partners. For dyads, the standard economic approach assumes that each person has a preferred outcome and that the two bargain over the single choice. For example, spouses may have different individual risk attitudes but have to make many joint decisions over a range of everyday economic activities (Mazzocco 2004).

The goal of this article is to explore differences in choices between individuals and pairs and determine what impacts relative bargaining power within a pair when making a joint decision. For this reason, we focus on *ad hoc* dyads so as to avoid repeated play concerns and unobservable (to the researcher) threat points that are invariably present with naturally occurring pairs. For example, a husband may concede on one decision so that he can get his way next time.[1]

This article contributes to the literature on risk taking by *ad hoc* dyads by considering the effect of demographic, psychological, and behavioral characteristics on the difference between individual and pair choices. The main conclusions from the experimental results are that we observe considerable evidence that people bargain over the action that the pair should take given their own personal preferences. We find that age and pair composition influence bargaining strength but that personality traits and observable behavior during (audio and video) recorded face-to-face pair discussions do not. We do, however, find that making a pair decision first increases individual risk taking. We also find that both a partner's personality traits and engagement in nontask socializing have an influence on subsequent individual choices.

2. Background Literature

Kocher, Strauss, and Sutter (2006) show that people have a preference for making decisions in groups, and there is a small but growing literature in experimental economics that studies group decision making. The results are not uniform. Cason and Mui (1997) explore group polarization in dictator game experiments and find that group behavior tends to resemble the behavior of the more generous members of the group, but Luhan, Kocher, and Sutter (2009) do not. Cox (2002) finds some evidence that groups are less trustworthy.[2] Some studies suggest that groups make *better* choices. For example, Cooper and Kagel (2005) find that pairs of subjects are dramatically better at a strategic signaling game than individuals. Kocher and Sutter (2005) show that groups are better at deducing optimal strategies in a beauty contest game. Borstein, Kugler, and Ziegelmeyer (2004) find that groups are more strategic in centipede games. However, other studies find that groups make poorer decisions. Cox and Hayne (2006) compare groups and individuals in common value auctions and report that groups are more susceptible to the winner's curse (see also Sutter, Kocher, and Strauss 2009).

[1] For a study that looks at the dynamics of bargaining power over time in couples, see de Palma, Picard, and Ziegelmeyer (2011).

[2] Group polarization refers to groups acting in a more extreme fashion than its members would individually. For background, see Stoner (1968) and Isenberg (1986).

Table 1. Decision Problem

	Option A	Option B
(Row) choice 1	$10.00 [–]	$19.25 [–]
	$8.00 [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]	$0.50 [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
(Row) choice 2	$10.00 [1]	$19.25 [1]
	$8.00 [2, 3, 4, 5, 6, 7, 8, 9, 10]	$0.50 [2, 3, 4, 5, 6, 7, 8, 9, 10]
(Row) choice 3	$10.00 [1, 2]	$19.25 [1, 2]
	$8.00 [3, 4, 5, 6, 7, 8, 9, 10]	$0.50 [3, 4, 5, 6, 7, 8, 9, 10]
(Row) choice 4	$10.00 [1, 2, 3]	$19.25 [1, 2, 3]
	$8.00 [4, 5, 6, 7, 8, 9, 10]	$0.50 [4, 5, 6, 7, 8, 9, 10]
(Row) choice 5	$10.00 [1, 2, 3, 4]	$19.25 [1, 2, 3, 4]
	$8.00 [5, 6, 7, 8, 9, 10]	$0.50 [5, 6, 7, 8, 9, 10]
(Row) choice 6	$10.00 [1, 2, 3, 4, 5]	$19.25 [1, 2, 3, 4, 5]
	$8.00 [6, 7, 8, 9, 10]	$0.50 [6, 7, 8, 9, 10]
(Row) choice 7	$10.00 [1, 2, 3, 4, 5, 6]	$19.25 [1, 2, 3, 4, 5, 6]
	$8.00 [7, 8, 9, 10]	$0.50 [7, 8, 9, 10]
(Row) choice 8	$10.00 [1, 2, 3, 4, 5, 6, 7]	$10.25 [1, 2, 3, 4, 5, 6, 7]
	$8.00 [8, 9, 10]	$0.50 [8, 9, 10]
(Row) choice 9	$10.00 [1, 2, 3, 4, 5, 6, 7, 8]	$19.25 [1, 2, 3, 4, 5, 6, 7, 8]
	$8.00 [9, 10]	$0.50 [9, 10]
(Row) choice 10	$10.00 [1, 2, 3, 4, 5, 6, 7, 8, 9]	$19.25 [1, 2, 3, 4, 5, 6, 7, 8, 9]
	$8.00 [10]	$0.50 [10]
(Row) choice 11	$10.00 [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]	$19.25 [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
	$8.00 [–]	$0.50 [–]

"(Row)" has been added to the table for expositional purposes and was on the form observed by subjects in the study. Numbers beside each payoff amount indicate the die rolls that would result in that payoff being realized. Subjects indicated their choice (A or B) for each of the 11 rows on a separate Scantron form.

Cooper and Kagel (2009) report that groups are no better than individuals at eliminating dominated strategies.

Experimental economists have also found mixed results as to how much risk groups are willing to accept as compared to the risk tolerance of individuals. Like strategic games, risk measurement tasks are in part judgmental because they deal with the idiosyncratic personal preferences of the decision maker in the sense that there is no objectively right amount of risk to take.[3] The judgmental nature of risk-taking tasks makes reaching a joint decision more complicated than in other settings where an optimal choice can be demonstrated. One common method for measuring risk attitudes is the Holt and Laury (2002; hereafter H&L) task. In the H&L task, a decision maker faces a series of choices between two binary lotteries (see Table 1). Masclet et al. (2009) use an H&L task to measure risk aversion in individuals and groups. They find that groups make fewer safe choices than individuals, and the researchers attribute this to the fact that the most risk-averse members of the group tend to concede to the other, more risk-loving group members. Baker, Laury, and Williams (2008) conduct a similar study with groups, but they report that the groups choose the safe lottery more frequently than the individual group members. Harrison et al. (2005) also conducted group experiments with the H&L procedure. Rather than the group reaching a decision, each member indicated one's own

[3] Laughlin (1980) defines intellective tasks as those with definite objective criteria and judgmental tasks as those for which there are no objective criteria, such as political or behavioral judgments. Intellective tasks can be further classified on the basis of their demonstrability, the ability to demonstrate why one choice is superior to another (see Laughlin and Ellis 1986).

preference for the decision of the group with the majority opinion being implemented. They report that group decisions closely follow that of the individuals in the group. Using willingness-to-pay data, Shupp and Williams (2008) found that the comparison of groups and individuals varies with the riskiness of the choice such that groups are more risk averse with riskier choices and vice versa.

The risky-choice papers discussed above focus on groups of three or more people, but there is an ongoing debate in the management and psychology literatures as to whether pairs are groups (for opposing perspectives, see Moreland 2010; Williams 2010). Therefore, it is not clear that these results shed much light on dyad behavior. To the degree that choices by pairs are important, one needs to study such behavior directly. Unfortunately, there is less research in economics on saliently motivated risky decision making by pairs than by larger groups. Sheremeta and Zhang (2010) find that pairs choose a safe lottery more often than individuals, a similar result to Baker, Laury, and Williams (2008). Bone, Hey, and Suckling (1999) find that pairs violate the expected utility axioms in a similar manner to that observed for individuals. Bateman and Munro (2005) move beyond *ad hoc* groups, looking at risk taking by established couples, and also find behavior similar to that of individuals.

An appealing aspect of pair decision making with judgmental tasks is that one can evaluate the choice as the outcome of a bargaining process. A natural starting point is to assume that pairs will simply meet in the middle, or "split the difference." Deviations from the middle ground suggest that one party has relatively greater bargaining power.[4] Our *ad hoc* pairs are an ideal place to explore bargaining power because the parties will not face future interactions (nor do they have a more complicated multidimensional problem) and there is no predetermined power structure.

What could explain bargaining power in *ad hoc* face-to-face pairs? One possibility is demographic variation. Another potential source of explanatory power is innate dispositions, or personality traits. One of the most enduring and popular taxonomies of human personality is Costa and McCrae's (1992) Five Factor Model (FFM), which identifies five traits that can be used to describe one's personality. These traits are extraversion, agreeableness, conscientiousness, neuroticism, and openness, often referred to as the "Big Five." The FFM is widely used in management and psychology because it is a parsimonious way of summarizing traits that have been identified in almost all other models of personality (Costa and McCrae 1992; Gill and Hodgkinson 2007). Moreover, research has consistently shown that the Big Five traits are stable across adulthood (McCrae and Costa 1990) and predict a variety of work (e.g., task performance, citizenship behaviors, job satisfaction, and training proficiency) and nonwork (e.g., creativity, life satisfaction, smoking, personality disorders, and decision making) attitudes, behaviors, and phenomena (see Barrick and Mount 1991; Saulsman and Page 2004; Malouff, Thorsteinsson, and Schutte 2006).

Formal hypotheses regarding the effects of each individual personality trait are not presented. However, we do suggest that personality plays a role in decision-making processes, as there is evidence that Big Five traits are related to judgment and decision making across a variety of contexts, including jury decisions (Clark et al. 2007), entrepreneurial business

[4] Of course, either party may prefer to exaggerate their own starting position so that meeting in the "middle" results in an outcome closer to their own true private preference. The ability to implement such a strategy is dependent on one's ability to identify the preference of the other party, as exaggerating preferences in the direction of one's rival would lead to a worse outcome. Our design cannot separate between this form of bargaining power and the more traditional notion of being able to obtain the lion's share given the starting point of the parties.

ventures (Wooten, Timmerman, and Folger 1998), and decisions to engage in risky health-related behaviors (Trobst et al. 2000). Moreover, traits such as extraversion, conscientiousness, and agreeableness are linked to persuasion attempts (i.e., people who are more outgoing are more likely to try to persuade others to see things their way), source credibility (i.e., individuals who are more organized and disciplined are likely to be viewed in a more favorable way), and acquiescence (i.e., by definition, someone who is more agreeable is more likely to acquiesce to others), all of which have the potential to influence decisions involving multiple decision makers. Big Five traits are also purported to affect decision making by influencing (over)confidence in decisions, sensitivity to information from the environment (McElroy and Dowd 2007), and heuristic biases (Trobst et al. 2000). In economics, personality tests have been used with varying degrees of success to explain behavioral variation in social dilemmas (Boone, De Brabander, and van Witteloostuijn 1999; Perugini, Tan, and Zizzo 2010), ultimatum and dictator games (Schmitt, et al. 2008), and bargaining (Gunnthorsdottir, McCabe, and Smith 2002).[5] Thus, we broadly consider what role, if any, the Big Five personality traits play in decision-making processes examined in the current study.

3. Experimental Design

The experiments in the current study relied on the H&L risk measurement tool.[6] The payoffs are shown in Table 1. In each row, decision makers choose between option A and option B. The payoffs associated with option A are always $10.00 and $8.00, and the payoffs associated with option B are always $19.25 and $0.50. These numbers represent the actual $U.S. payments that the subjects could receive. What varies between rows is the likelihood of receiving the larger amount. In the first row, there is a 0% chance; in the second row, there is a 10% chance; and so on. Under standard assumptions, an individual who prefers more money to less should prefer option A in row 1 up to row X and then prefer option B for row X + 1 to row 11 with $1 \leq X < 11$. For a risk-neutral person, $X = 6$. A risk-averse person would select the safe option more than six times, and a greater number of safe choices implies a greater degree of risk aversion.[7]

Each subject completed the risk measurement task twice: once individually and once jointly with a randomly selected partner. Some pairs completed the task together before completing the task individually, and some pairs completed the task as individuals first. Subjects knew that they would complete the task twice but did not know how the second task differed from the first when they were completing the first task. To control for wealth effects, as in H&L, subjects were told that only one row from one task would be randomly selected for payment. Which task (individual or pair) would determine the subject's payoff was determined by a coin flip. The specific row used for payment was then selected by the roll of a die. Subjects were given the opportunity to inspect the randomization devices prior to making their decisions.

[5] For a discussion of incorporating personality research into laboratory experiments, see Swope et al. (2008). Some of these studies used alternative personality tests, such as the Myers-Briggs Personality Inventory. McCrae and Costa (1989) discuss the overlap between the Big Five and the Myers-Briggs approaches to personality.

[6] The task is slightly modified by inclusion of the first row, in which the outcome is certain to be the lower amount.

[7] Because of the coarseness of the instrument, individuals with slight risk aversion, or risk-preferring preferences would also choose $X = 6$. For the model of constant relative risk aversion where $U(w) = w^{1-r}/(1-r)$ and the parameter values of r such that $-0.15 < r < 0.15$ are consistent with switching at $X = 6$.

In addition to the two tasks, each subject completed a survey involving demographic information (i.e., age and gender) as well as personality characteristics.[8] Consistent with previous research, an established measure of the FFM from the International Personality Item Pool was used to assess the Big Five personality trait markers (Goldberg 1999). This measure uses 10 statements to which the respondents can strongly disagree, disagree, be neutral, agree, or strongly agree using a five-point Likert scale. For example, statements regarding neuroticism include "I get stressed out easily," "I seldom feel blue," and "I get irritated easily." Participants' responses were averaged to provide a mean score for each of the five personality traits.

In each laboratory session, eight subjects arrived at the lab at the prespecified time.[9] The subjects then drew ID numbers that would determine the task order and their pairings.[10] After drawing ID numbers, subjects proceeded to the lab's computer room, where they completed the online survey. Each station in the computer lab is separated by a system of privacy dividers. After all the subjects had completed the survey, those in the pair task were taken across the hall to one of the lab's group decision-making rooms (a small room with a table that can seat eight people). These subjects were given directions for the pair task and a Scantron form on which to indicate their responses and were then left in the room to make their decision. Video and audio recordings were made of the pair interaction. After the pairs completed the task, they returned to the computer lab, sat at separate work stations, received the individual task directions and Scantrons, and completed the task in private. The subjects who completed the individual task first remained in the computer initially and then moved to one of the lab's group rooms after completing the individual task. Copies of the directions for the pair and individual tasks are included in the Appendix. After all the subjects had completed both tasks, a coin was flipped to determine if the pair or individual task would be used to determine the payoff. If the pair task was selected, the pair rolled a single 12-sided die to determine which of the 11 rows would be selected (the subjects simply rerolled if the die landed on 12). Finally, a 10-sided die was used to determine the actual payoff that both subjects received. Once the subjects were paid, they were dismissed from the experiment. If the individual task was selected for payment, then each person privately rolled the two dice to determine their own payoff before being dismissed.

A total of 204 (102 dyads) undergraduate students at the University of Arkansas participated in the experiment. By random assignment, 96 (48 pairs) completed the individual task first, and 108 (54 pairs) completed the pair task first. Some subjects had participated in previous economics experiments, but none had participated in any studies measuring risk attitudes. The experiments lasted approximately 30 minutes. In addition to the salient earnings, subjects also received a $5 participation payment.

4. Data

The usable data consist of the H&L choices and complete survey information for 199 individuals and 97 pairs. We excluded one individual because this person never selected option

[8] A copy of the survey is available on request.

[9] Some sessions were run with four or six subjects because of other subjects' absenteeism.

[10] Subjects were not informed at this point that they would be engaging in a pair task, nor were they informed as to how their specific ID number would be used. Another advantage of drawing ID numbers is that no identifying information was collected from the subjects that could be connected to their responses, some of which are potentially sensitive.

Table 2. Descriptive Statistics of Sample Variables

	Mean	Standard Deviation	Minimum	Maximum
Male	0.648	0.479	0	1
Age	21.261	2.800	17	39
Big Five personality				
Agreeableness	3.980	0.527	1.9	5
Extraversion	3.447	0.722	1.8	5
Conscientiousness	3.210	0.441	1.4	4.2
Openness	3.670	0.452	2.5	4.6
Neuroticism	2.578	0.770	1	4.6
Video				
Justifications	1.035	0.907	0	3
Conversational leadership	0.910	0.780	0	2
Physical superiority	0.548	0.547	0	2
Social activity	0.151	0.359	0	1
No. of subjects	199			

B, perhaps indicating confusion. The other four subjects were excluded because of incomplete demographic information. Of the usable observations, 94 subjects performed the individual task first. There were more males (65%) in our sample, and the average age was 21, with a minimum age of 17 and a maximum age of 39. Table 2 contains summary statistics for each variable in our sample.

Following Baker, Laury, and Williams (2008), we use the number of safe choices as our measure of risk aversion. Theoretically, a subject should exhibit a single switching point from preferring the safe lottery to preferring the risky lottery, but it is common for some subjects to not behave consistently. Interestingly, no individual and no pair in our sample ever switched from the risky lottery back to the safe lottery.[11]

Our data also include the audio and video recordings of the pair decision-making process. While not common in economics, this type of information is commonly studied in other disciplines, such as psychology. Transforming the recordings into usable data requires a coding process that identifies the set of possible outcomes (for an overview, see Bakeman 2000). The same three observers watched each pair interaction and recorded occurrences of the following behaviors: justifications for choice (logical and emotional), conversational leadership (talking first and talking most), physical superiority (standing and completing the response form), and social activity (non–task-related interactions).[12] Conversational leadership is of particular interest, as psychologists have argued that there is a relationship between group leadership and

[11] As stated above, one subject selected the safe option every time and was eliminated from our sample. One of the subjects who was eliminated because of incomplete demographic information did not have a single switching point. We speculate that the high level of consistent behavior was due to our use of a Scantron response form.

[12] To assess interrater reliability among our three coders, we calculated Fleiss's kappa. This measure took on the following values for our coders (with higher numbers indicating greater consistency): 0.405 for the logical value, 0.244 for emotional, 0.716 for talking first, 0.460 for talking most, 0.736 for standing, 0.856 for completing the sheet, and 0.573 for engaging in non–task-related social activity. There is an ongoing debate as to how to objectively evaluate Fleiss's kappa, and therefore we also evaluated Cohen's kappa for each possible set of two coders for each of the behaviors. In the interest of brevity, rather than reporting the 3 combinations of two coders $\times$ 7 behaviors = 21 values for Cohen's kappa, we simply note that the probability of observing a Cohen's kappa as high as was observed by chance is less than 0.001 for each of the 21 values. Given the high degree of agreement between our coders, we rely on the average rater coding as the measure of subject behavior.

taking control of the conversation. Specifically, Van Knippenberg, Van Knippenberg, and Van Dijk (2000) report that risk-seeking individuals are more likely to take the lead of a group. A variable was created for each subject indicating the number of actions of a particular behavior the person did (i.e., someone who talked first and talked the most would receive a score of two for conversational leadership out of a maximum of two; someone who filled in the response sheet but did not stand would have a score of one out of a maximum of two for physical superiority).

5. A Model of Bargaining

In order to provide a framework for exploring differences in behavior across conditions, we present a simple conceptual model. Let $x^*_{i,p}$ denote the H&L choice revealing the true preference of person i randomly assigned to pair p. Let $x_{i,p}$ denote the actual individual choice made by the person. There are four different types of choices that we observe in our experiment: individual choices before or after the pair task and pair choices before or after the individual task. First, it seems reasonable to assume that individuals, when they take the individual task first, will reveal their true risk aversion. That is, $x_{i,p} = x^*_{i,p}$. This assumption hinges on the validity of the H&L method but is not testable. Second, it is also reasonable to assume that the pair choice is not influenced by whether the individual task is done first or second, as it is determined by a collective decision-making procedure based on two paired people's true risk preferences. In other words, the pair choice should not differ on the basis of task ordering.[13] This is testable and forms the basis of Hypothesis 1.

HYPOTHESIS 1: There is no difference in the distribution of pair choices regardless of whether the pair task is done before or after the individual task.

The remaining question is how individual choices will be influenced by the pair choice when the pair task is done first. To address this question, let us assume that the individual choice made after the pair choice is the weighted average of the person's true choice and the predetermined pair choice. Formally, we conjecture that

$$x_{i,p} = (1 - \alpha_{i,p})x^*_{i,p} + \alpha_{i,p}x_p, \tag{1}$$

where x_p is the preceding pair choice and the parameter $\alpha_{i,p} \in [0,1]$ represents the extent to which the person is influenced by the pair choice. This parameter may depend on the person's characteristics and also the characteristics of his or her partner. For example, it may be that young people are more likely to be influenced by the pair choice than are older people.

Assume, without loss of generality, that person i is the more risk-averse member in the pairing. That is, $x^*_{i,p} \geq x^*_{-i,p}$, where the subscript $-i$ represents the partner of subject i. We assume that the pair choice is determined by a Nash bargaining procedure, consistent with the observed behavior shown in Figure 1. Specifically, we assume that the pair choice is determined by

[13] It is possible that there could be a "confirmation" effect where people who make their individual choices first hold on more strongly to their individual preferences when bargaining. However, our data suggest that this is not the case, as discussed later.

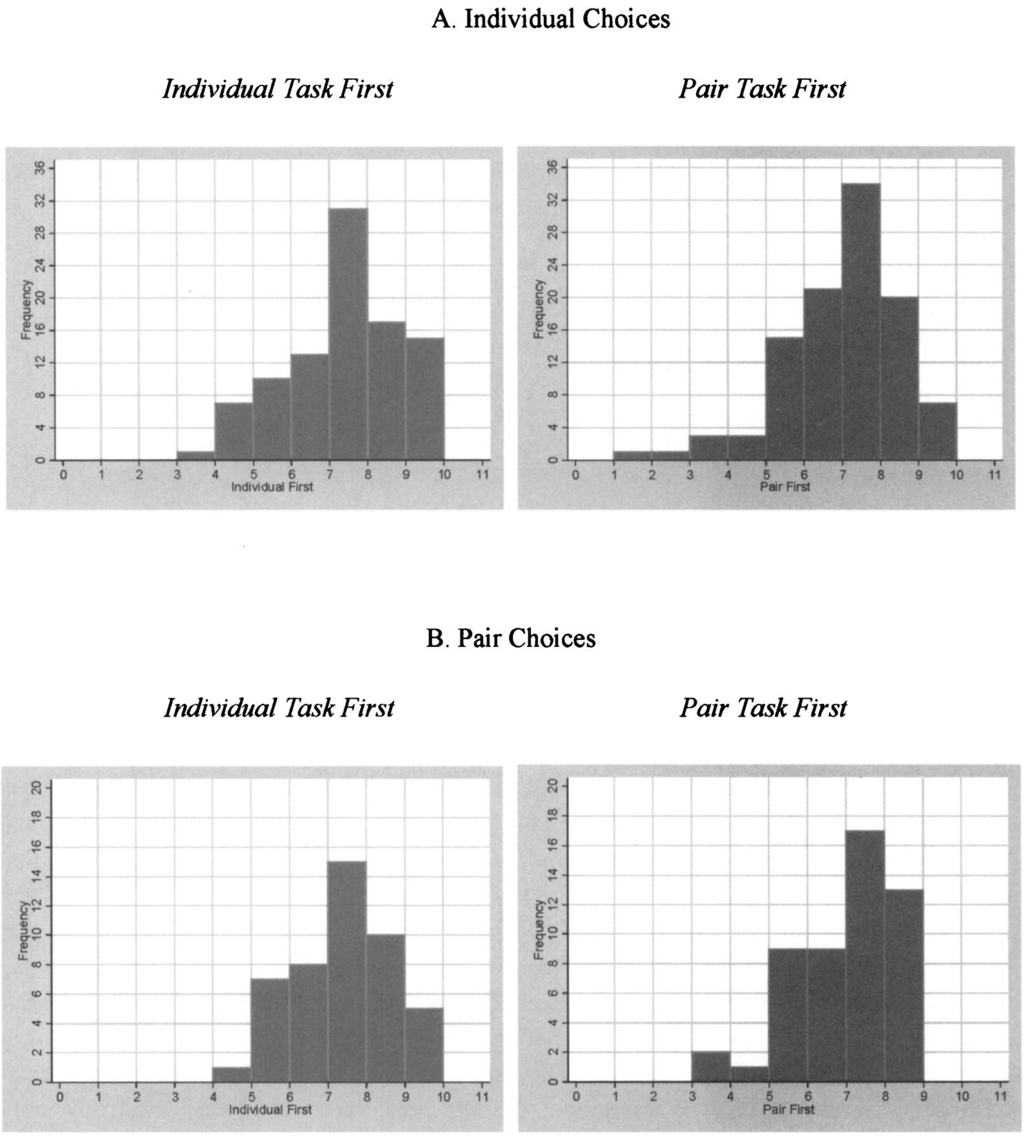


Figure 1. Histograms for Number of Safe Choices by Individuals and Pairs

$$x_p = w_p x^*_{i,p} + (1 - w_p) x^*_{-i,p}, \tag{2}$$

where $w_p \in [0,1]$ is the bargaining weight for the more risk-averse member of the pair.

The bargaining model in Equation 2 combined with Equation 1 implies that, when the individual task is completed second, the more risk-averse person's observed choice is weakly lower than his or her true choice and that the less risk-averse person's choice is weakly greater than his or her true choice. Thus, the variance of observed individual choices after the pair task should be (weakly) smaller than that of his or her true choices. This is testable and forms the basis of Hypothesis 2.

Hypothesis 2: The variance of individual choices after the pair task is weakly lower than the variance of individual choices before the pair task.

We can also compare the average of individual choices made before and after the pair task. Given that there are P dyads, the average of *all* individual choices made following the pair task is the simple average of the dyad average over the P dyads, which is $1/P \times \Sigma_p\ (0.5x_{i,p} + 0.5x_{-i,p})$.

However, the average over *all* individual choices made before the pair task is $1/P \times \Sigma_p$ $(0.5x^*_{i,p} + 0.5x^*_{-i,p})$ since these individuals reveal their true preferences. To compare the two averages, first substitute Equation 2 into Equation 1. From this, the individual choice of person i is

$$\begin{aligned} x_{i,p} &= (1-\alpha_{i,p})x^*_{i,p} + \alpha_{i,p}x_p = (1-\alpha_{i,p})x^*_{i,p} + \alpha_{i,p}(w_p x^*_{i,p} + (1-w_p)x^*_{-i,p}) \\ &= (1-\alpha_{i,p}(1-w_p))x^*_{i,p} + \alpha_{i,p}(1-w_p)x^*_{-i,p}. \end{aligned}$$

Likewise, the partner's choice is

$$\begin{aligned} x_{-i,p} &= (1-\alpha_{-i,p})x^*_{-i,p} + \alpha_{-i,p}x_p = (1-\alpha_{-i,p})x^*_{-i,p} + \alpha_{-i,p}(w_p x^*_{i,p} + (1-w_p)x^*_{-i,p}) \\ &= (1-\alpha_{-i,p}w_p)x^*_{-i,p} + \alpha_{-i,p}w_p x^*_{i,p}. \end{aligned}$$

Therefore, the average of the two individual choices of a single pair made after the pair task is

$$\begin{aligned} 0.5x_{i,p} + 0.5x_{-i,p} &= 0.5[(1-\alpha_{i,p}(1-w_p))x^*_{i,p} + \alpha_{i,p}(1-w_p)x^*_{-i,p} + (1-\alpha_{-i,p}w_p)x^*_{-i,p} + \alpha_{-i,p}w_p x^*_{i,p}] \\ &= 0.5[(1-\alpha_{i,p}(1-w_p) + \alpha_{-i,p}w_p)x^*_{i,p} + (1+\alpha_{i,p}(1-w_p) - \alpha_{-i,p}w_p)x^*_{-i,p}] \\ &= (0.5 - 0.5\alpha_{i,p}(1-w_p) + 0.5\alpha_{-i,p}w_p)x^*_{i,p} + (0.5 + 0.5\alpha_{i,p}(1-w_p) - 0.5\alpha_{-i,p}w_p)x^*_{-i,p}. \end{aligned}$$

The previous line shows that the average of the individual choices after the pair task is a weighted average of their *true* choices, which in general differs from a simple average. Thus, the average of individual choices after the pair task may differ from the average of individual choices before the pair task. Specifically, the two differ if $\alpha_{-i,p}w_p - \alpha_{i,p}(1 - w_p) \neq 0$. The direction of the change depends on which person is stronger in terms of bargaining power as well as which person is relatively more sensitive to the preceding pair choice. If $\alpha_{-i,p}w_p - \alpha_{i,p}(1 - w_p) > 0$, the average of individual choices after the pair task should be higher than that of *true* choices. If $\alpha_{-i,p}w_p - \alpha_{i,p}(1 - w_p) < 0$, the average of individual choices after the pair task should be lower than that of *true* choices. This case is likely to occur if the bargaining power of the more risk-averse member is lower or if the more risk-averse member is relatively more influenced by the pair choice. In a particular case where members have the same bargaining power, the comparison depends only on which person is more influenced by the preceding pair choice.[14]

6. Empirical Results

Observed behavior is summarized in Table 3 and presented graphically in Figure 1. It appears that individuals who have already made a pair decision take more risk than those individuals making their first decision. When pair members have individual experience, the pair

[14] Our data on bargaining outcomes confirm this assumption. If we regress x_p on $x^*_{i,p}$ and $x^*_{-i,p}$ for the sample of those who performed the individual task first, then we cannot reject that the bargaining weight is equal to 0.5.

Table 3. Tests Comparing Individual and Pair Choices

	Individual Choices	
	Individual Task First	Pair Task First
No. of observations	94	105
Mean	6.968	6.543
Variance	1.636	1.544
Normal distribution test	Jarque-Bera = 0.502 (>0.50)	Jarque-Bera = 18.715 (0.03)
Skewness	0.730	0.001
Kurtosis	0.645	0.025
Equal central tendency tests	t = 1.886 (0.061)	
	z approximation for U = 1.604 (0.1087)	
Equal variance test	F = 1.122 (0.566)	
Identical distribution test	Kolmogorov–Smirnov (0.735)	
	Pair Choices	
	Individual Task First	Pair Task First
No. of observations	46	51
Mean	6.913	6.529
Variance	1.330	1.347
Normal distribution test	JB = 0.385 (>0.05)	JB = 3.976 (0.070)
Skewness	0.972	0.038
Kurtosis	0.674	0.594
Equal central tendency tests	t = 1.409 (0.162)	
	z approximation for U = 1.113 (0.266)	
Equal variance test	F = 0.976 (0.937)	
Identical distribution test	Kolmogorov–Smirnov (0.984)	

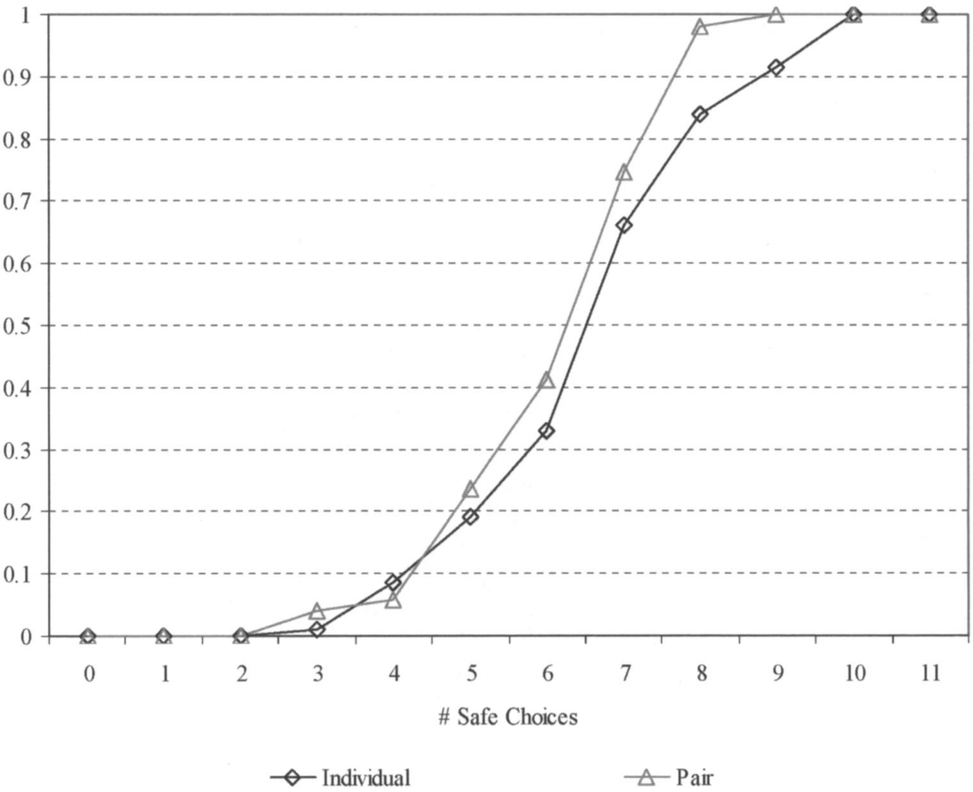
The p-value is presented in parentheses. Among those individual choices when the individual task is implemented first, one person's choice is excluded because the choice is unreasonably zero.

tends to take nominally less risk than pairs of inexperienced people. This suggests a potential ordering effect. Therefore, to compare individual and pair, we rely on the individual observations for those who did the individual task first and the pair observations for those completing the pair task first (a comparison of the top left and bottom right distributions shown in Figure 1). Figure 2 shows the cumulative density functions of the number of safe choices under these two inexperienced conditions. Figure 2 suggests that the distributions of choices are similar across the two treatments for riskier bets, but they appear to differ for safer bets, similar to the results of Shupp and Williams (2008). However, the distributions are not statistically different based on a Kolmogorov–Smirnov test, p-value of 0.505.

The following section ("Distribution Analysis and Test of Hypotheses") discusses the treatment effects in more detail and formally tests the hypotheses developed in the previous section. A separate section ("Bargaining Outcomes") treats pair decisions as the result of a bargaining process between the individuals and considers relative bargaining power. A final section ("Influence of Pair Choice on Individual Choice") examines the lasting impact that pair decisions and partner characteristics have on subsequent individual behavior.

Distribution Analysis and Test of Hypotheses

Supporting Hypothesis 1, the results indicate that the distributions of pair choices (see panel B of Figure 2) do not differ based on whether the task is completed first or second. That



Notes: N = 47 individuals who performed the individual treatment first and 27 pairs who performed the pair treatment first.

Figure 2. Cumulative Distribution Function for Selecting Risky Option by Individuals and Pairs

is, none of the four statistical tests reported in the bottom panel of Table 3 can reject the null hypotheses of there being no difference between the samples. Thus, as expected, there is no evidence of an ordering effect for the pair task.[15]

We do not find support for Hypothesis 2 that individual choices made after the pair task are less dispersed. The *F*-test of equal variances cannot be rejected, as reported in the top panel of Table 3. The result of the Kolmogorov–Smirnov test for identical individual choice distributions yields a similar conclusion (p-value = 0.984).

From panel A of Figure 1, it appears that subjects who completed the individual task first (left graph) make more safe choices as individuals than those who performed the pair task first (right graph). This result is weakly supported statistically in Table 3, where the t-test for equal

[15] One possible explanation for an ordering effect could be learning, and it is possible that learning differs from an individual and a pair task. While learning may occur between the first and the second task, the results of testing Hypotheses 1 and 3 provide some evidence that behavior is not being driven by a simple learning effect. For the results to hold, one would have to argue that no learning occurs from completing the individual task (Hypothesis 1 result) and that people systematically overstate their degree of risk aversion without learning. However, it is possible that some more complicated learning story explains the order effect that we observe.

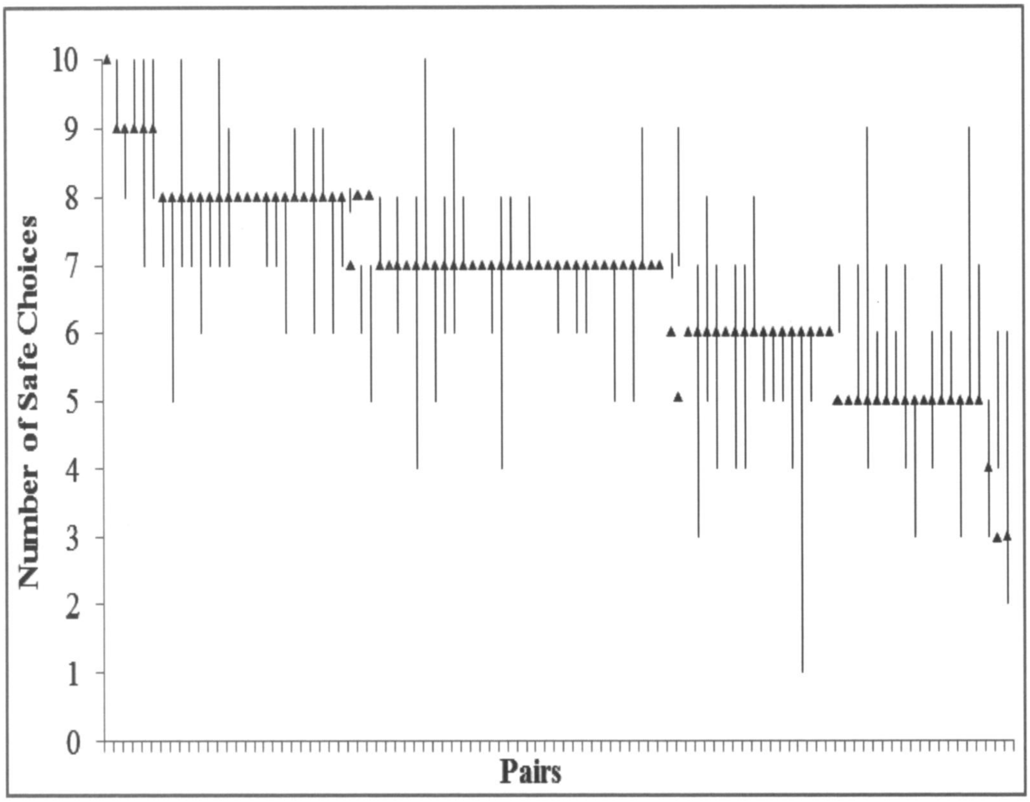


Figure 3. Pair Choices and the Range of Individual Choices

mean has a p-value of 0.061 and the Mann–Whitney test has a p-value of 0.109. These results suggest that the treatment order affects behavior.

Bargaining Outcomes

We first present graphical evidence in Figure 3 that pair outcomes are the result of bargaining. The solid markers in Figure 3 indicate the number of safe choices selected by the pair, while the top (bottom) of the solid vertical lines indicate the number of safe choices selected by the less (more) risk-averse individual in the pair. For 92.8% of the pairings, the pair choice falls weakly inside the range defined by the individual choices (i.e., only seven of the markers in Figure 3 lie off the vertical lines). Thus, we find little evidence that dyads make choices more extreme in either direction than both individuals in the pair, consistent with pair decision being the result of a bargaining process.

Given the potential ordering effect discussed in the previous section, we focus on those subjects who performed the individual task first since their individual choices should represent their true risk aversion ($x_{i,p} = x^*_{i,p}$). Let us continue to assume that individual i is the more risk-averse member of the pair. Under the hypothesis that people bargain in the pair task, $x_{-i,p} \leq x_p \leq x_{i,p}$. We define the pair-specific Nash bargaining weight as the following:

$$w_p = \frac{x_p - x_{-i,p}}{x_{i,p} - x_{-i,p}} = \frac{x_p - x^*_{-i,p}}{x^*_{i,p} - x^*_{-i,p}}, \tag{3}$$

where $w_p = 1$ indicates that the more risk-averse person gets his or her preferred outcome and $w_p = 0$ indicates that the more risk-loving person gets his or her preferred outcome. We use this relative weight variable as the dependent variable in a regression that controls for various individual characteristics with the objective of identifying which characteristics determine the bargaining weight of the more risk-averse person. In the spirit of the literature on collective bargaining models (Browning et al., 1994), we assume that intersubject *differences* in individual characteristics, such as extraversion and agreeableness, determine the bargaining weight. Formally, we use the following regression:

$$w_p = \beta_0 + \beta_1 \Delta Age + \beta_2 \Delta Personality + \beta_3 \Delta Conversation + \beta_4 GenderComposition + u_p, \quad (4)$$

where u_p is the random component of the bargaining weight, ΔAge is the age difference between the more and less risk-averse individuals ($Age_i - Age_{-i}$), $\Delta Personality$ represents a vector of the differences in Big Five personality measures, $\Delta Conversation$ is the difference in the measure of conversational leadership constructed from the video recordings, and *GenderComposition* is a vector of dummy variables that controls for the gender composition of the pair. Specifically, we include the following dummies: male, male partner, and both male. The reference group is a female with a female partner.

The sample we used for this analysis includes 34 pairs who performed the pair task second and made different individual choices.[16] The regression results are presented in Table 4. Columns 1 and 2 of Table 4 use different specifications, with column 2 including personality characteristics. First, we note that the constant terms are similar to 0.5, suggesting that two people in the reference group (both women) and who are similar in age and other characteristics will have equal bargaining power and essentially agree on an outcome that is halfway between their individual preferences. We find that two variables, age difference and both male, are significant for the case where the personality measures are omitted from the regression analysis, suggesting that these two factors influence bargaining power. However, only the age difference remains significant when controls for personality are included. The negative coefficient for age difference indicates that the greater the age gap, the lower the bargaining weight of the older person. That is, the pair's decision will be more similar to the preference of the younger person than to that of the older person. The negative sign and the statistical significance of the gender composition variable for both people being male indicates a peer pressure effect between males such that the more risk-averse person is given less bargaining power. That is, when two males are making a decision, their choice will be more similar to the more risk-loving individual than to the more risk-averse individual.

None of the Big Five personality traits demonstrated statistically significant effects, when these are all included in the regression (see column 2 of Table 4). We also tried alternative specifications in which only one of the personality traits is included in the regression, but none resulted in statistically significant results. Thus, there was no evidence that personality impacts bargaining power. Similarly, none of the variables from the audio-video recordings were significant when included separately. That is, the number of justifications, conversational leadership, physical superiority, and social activity do not significantly impact bargaining power. Because none of these variables are significant and including them does not

[16] Nine pairs were dropped because both people in the pair made the same individual choice. When the two individuals have the same preferences, there is no bargaining, and the dependent variable is infinite. Three additional pairs were dropped because the pair choice lay outside the range defined by the individual choices.

Table 4. Determinants for High Risk-Aversion Person's Bargaining Weight

	(1)	(2)	Placebo (3)	Placebo (4)
ΔAge	−0.083**	−0.083**	0.013	0.018
	(0.023)	(0.029)	(0.013)	(0.015)
Male	0.240	0.236	−0.047	−0.107
	(0.167)	(0.221)	(0.270)	(0.312)
Male partner	0.096	0.033	−0.142	−0.174
	(0.205)	(0.208)	(0.240)	(0.244)
Both male	−0.486*	−0.442	0.245	0.142
	(0.226)	(0.271)	(0.340)	(0.415)
ΔAgreeableness		0.115		−0.203*
		(0.098)		(0.075)
ΔExtraversion		0.023		0.029
		(0.051)		(0.117)
ΔConscientiousness		0.039		0.012
		(0.112)		(0.145)
ΔOpenness		−0.111		−0.225
		(0.088)		(0.148)
ΔNeuroticism		−0.018		−0.031
		(0.074)		(0.143)
Constant	0.500**	0.512**	0.639**	0.701**
	(0.143)	(0.161)	(0.173)	(0.144)
Observations	34	34	32	32
Adjusted R^2	0.317	0.279	−0.100	−0.010
Joint significance of all variables	$F = 15.68$ (< 0.001)	$F = 7.14$ (< 0.001)	$F = 0.67$ (0.618)	$F = 1.61$ (0.173)

The dependent variable is w_p except in columns 3, and 4 where it is $\tilde{w}_p$, which is defined in Equation 4. The linear equation is estimated by ordinary least squares. Tobit results are qualitatively the same. The sample for columns 1–2 includes 34 pairs who performed the individual task first. Columns 3 and 4 present placebo test results where the sample includes 32 pairs who performed the pair task first. Robust standard errors are reported in parentheses. Asterisks (** and *) denote statistical significance at the 1% and 5% levels, respectively.

substantially change other estimated coefficients, these results are omitted from Table 4 for brevity.

The last two columns of Table 4 test for the validity of our bargaining model using a placebo test. Here we use only dyads that performed the pair task first. For these pairs, the bargaining weights created on the basis of observed individual choices should be tainted with the ordering effect, as shown in Equation 1 and tested in the following section. For those dyads who performed the pair task first, the bargaining weight is

$$\tilde{w}_p = \frac{x_p - x_{-i,p}}{x_{i,p} - x_{-i,p}} = \frac{(1-\alpha_{i,p})x_p - (1-\alpha_{-i,p})x^*_{-i,p}}{(1-\alpha_{i,p})x^*_{i,p} - (1-\alpha_{-i,p})x^*_{-i,p} + (\alpha_{i,p} - \alpha_{-i,p})} \neq \frac{x_p - x^*_{-i,p}}{x^*_{i,p} - x^*_{-i,p}}. \quad (5)$$

This fictitious bargaining weight is constructed on the basis of individual choices that are influenced by the preceding pair choice (and not the true individual preference). The last inequality means that the pseudo–bargaining weight should deviate from the true bargaining weight unless $\alpha_{i,p} = \alpha_{-i,p}$.[17] The results (see columns 3 and 4 of Table 4) show that all the variables

[17] In fact, $\alpha_{i,p}$ and $\alpha_{-i,p}$ are likely to be negatively correlated because the group choice is made by bargaining where it is likely that one person wins and the other person loses.

included in the analysis lack explanatory power, except agreeableness. The most important result here is the fact that the joint significance of all variables is always rejected for the placebo (columns 3 and 4), but it is not for the correctly specified bargaining model (columns 1 and 2).

Influence of Pair Choice on Individual Choice

In this section, we explore how the preceding pair choice influences the subsequent individual choice. First, we examine what affects the likelihood that a person will deviate from the dyad choice when making a subsequent individual choice. For this purpose, we construct a dummy variable (i.e., *Follow*) that takes the value of one if the individual choice coincides with the group choice and zero otherwise. We used this to estimate the following Probit regression:

$$Follow_{i,p} = F(\gamma_1 + \gamma_2 x_p + \gamma_3 Age_{i,p} + \gamma_4 Age^2_{i,p} + \gamma_5 Personality_{i,p} + \gamma_6 GenderComposition_p), \quad (6)$$

where F is the standard normal cumulative distribution function and, as before, x_p denotes the decision of individual i's pair. The pair choice is included in order to control for initial conditions for each individual.[18]

Table 5 presents the results for the regression analysis based on Equation 6 using the data for the individual choices of those subjects who went through the pair task first.[19] We find strong statistical evidence that agreeable individuals are more likely to deviate from the pair choice. This result is intuitive because agreeable individuals would presumably be more willing to sacrifice their individual preferences (i.e., acquiesce) when participating in the pair task. No other control variables are significant in the regression.

To further investigate whether the pair choice influences the individual choices when the pair task precedes the individual task, we ideally want to estimate Equation 1. However, this is not possible because we do not observe the true choice. Thus, instead, we estimate the following reduced-form equation:

$$x_{i,p} = \delta_0 + \delta_1 x_p + \delta_2 X_{i,p} + \delta_3 X_{-i,p} + v_{i,p}. \quad (7)$$

In Equation 1, the individual choice depends on the predetermined pair choice (x_p), one's true choice, and $\alpha_{i,p}$ This last parameter is likely to depend on an individual's own characteristics as well of those of his or her partner.[20] Therefore, $X_{i,p}$ is the vector of individual i's characteristics, including demographic characteristics, conversational leadership, and Big Five personality traits. $X_{-i,p}$ is a similar vector of partner characteristics.[21] Finally, $v_{i,p}$ is an error term.

Equation 7 is useful for testing if the pair choice influences the subsequent individual choice. In terms of Equation 1, we want to test if $\alpha_{i,p} = 0$. Suppose that the individual choice is

[18] Because of the structure of the task, there is an upper bound on the number of safe choices. The bargaining framework creates a lower bound on the number of safe choices for the more risk-averse person. Therefore, the more risk-averse individual has less room to deviate from the pair choice the more safe choices the pair made. The same effect would restrict the more risk-loving person's ability to deviate from the pair if the pair made few safe choices, suggesting that x_p should enter Equation 6 nonlinearly; however, none of our pair observations is close to the lower boundary.

[19] One of the 105 observations was lost because the partner's gender was missing.

[20] One might include interaction terms between the pair choice and individual characteristics because the pair choice and the pair choice influence factor are included in Equation 1 in the multiplicative form. Even if we include all possible interaction terms, the results remain qualitatively the same.

[21] Unlike the estimation in Table 2, which relies on difference in personality characteristics to determine group outcome, the estimation in Table 4 requires each individual personality scores to determine the influence of one's partner controlling for one's own personality.

Table 5. Probit Regression for Likelihood Individual Follows Pair Choice

	(1)
	Probit
Pair Choice	0.139*
	(0.041)
Age	−0.001
	(0.123)
Age squared	0.000
	(0.002)
Agreeableness	−0.251*
	(0.097)
Extraversion	0.042
	(0.086)
Conscientiousness	0.152
	(0.103)
Openness	−0.156
	(0.134)
Neuroticism	0.018
	(0.070)
Male	0.198
	(0.210)
Male partner	0.189
	(0.220)
Both male	−0.121
	(0.238)
Observations	104
Pseudo-R^2	0.164

Robust standard errors reported in parentheses. Marginal effects evaluated at sample means are presented. Asterisk (*) denotes statistical significance at the 5% level.

not affected by the pair choice. In this case, the individual choice should be the true choice regardless of the treatment ordering. Since the true choice is based on one's own risk preference, the partner's characteristics should not affect the individual choice.[22] If individual choices are not impacted by partner characteristics, the estimates of the $\delta_3 s$ should be jointly insignificant.

We estimate Equation 7 by ordinary least squares. Robust standard errors are adjusted for clustering by pairs. We also estimate the equation after including pair fixed effects. In this case, we assume that $v_{i,p} = a_p + \varepsilon_{i,p}$, where a_p represents the pair-specific effects and $\varepsilon_{i,p}$ is the error term. The regression sample includes 102 subjects because three observations are lost because of missing data for the partner's birth year. Table 6 presents the regression results. There are four columns that employ different specifications and different samples. In column 1, we include only the subject's own demographic and personality characteristics. In column 2, we add the dyad choice. Column 3 includes partner characteristics as well as the measure of conversational leadership observed from video recording. In the last column, we implement a specification test that is based on a similar idea to the placebo test in Table 4, namely, applying the same estimation to those who did the tasks in the opposite order.

[22] It is still likely that the pair choice might be significant when we estimate Equation 7. This is because the pair choice also reflects the true choice.

Table 6. Impacts of Pair Choice on Subsequent Individual Choice

				Placebo
	(1)	(2)	(3)	(4)
Own characteristics				
Gender (male = 1)	−0.033	0.165	0.017	0.233
	(0.387)	(0.248)	(0.258)	(0.376)
Age	−0.091	−0.085	−0.085	0.611
	(0.331)	(0.272)	(0.234)	(0.399)
Age squared	0.001	0.002	0.002	−0.012
	(0.006)	(0.006)	(0.005)	(0.007)
Agreeableness	0.337	0.234	0.138	0.757
	(0.328)	(0.223)	(0.238)	(0.465)
Extraversion	0.087	−0.042	−0.041	−0.339
	(0.228)	(0.197)	(0.244)	(0.339)
Conscientiousness	0.681*	0.619**	0.596**	−0.381
	(0.359)	(0.246)	(0.224)	(0.416)
Openness	−0.438	−0.455*	−0.540	−0.069
	(0.312)	(0.255)	(0.327)	(0.327)
Neuroticism	0.213	0.263*	0.135	−0.140
	(0.207)	(0.134)	(0.155)	(0.245)
Pair choice		0.805***	0.735***	0.612***
		(0.067)	(0.088)	(0.144)
Justifications (video)			0.208	0.436
			(0.205)	(0.261)
Conversational leadership (video)			−0.165	0.200
			(0.217)	(0.208)
Physical superiority (video)			−0.134	−0.394
			(0.454)	(0.316)
Social activity (video)			−0.650***	−0.481
			(0.242)	(1.508)
Partner's characteristics				
Gender (male = 1)			−0.053	0.231
			(0.268)	(0.382)
Age			0.050	0.059
			(0.233)	(0.420)
Age squared			−0.002	−0.002
			(0.004)	(0.008)
Agreeableness			−0.064	−0.157
			(0.263)	(0.545)
Extraversion			0.359*	−0.260
			(0.188)	(0.362)
Conscientiousness			0.311	−0.596
			(0.245)	(0.470)
Openness			−0.167	0.422
			(0.253)	(0.419)
Neuroticism			0.153	−0.412
			(0.153)	(0.251)
Justifications (video)			−0.123	−0.373
			(0.200)	(0.272)
Conversational leadership (video)			−0.027	0.138
			(0.243)	(0.218)

Table 6. Continued

	(1)	(2)	(3)	Placebo (4)
Physical superiority (video)			−0.363	−0.372
			(0.494)	(0.235)
Social activity (video)			0.106	−0.086
			(0.382)	(1.573)
Constant	5.371	0.155	0.211	−2.158
	(4.496)	(3.826)	(4.809)	(4.940)
Observations	102	102	102	93
Adjusted R^2	0.016	0.502	0.482	0.201
Joint significance of partner's characteristics			F = 1.89 (0.059)	F = 0.97 (0.489)

In columns 1–3, the sample includes 102 individuals who performed the pair task first. In column 4, the sample includes 93 individuals who performed the individual task first. Robust standard errors, adjusted for clustering by pairs, are reported in parentheses. Asterisks (***, **, and *) denote statistical significance at the 1%, 5%, and 10% levels, respectively.

For the purpose of this section, the main result is the joint significance of the partner's characteristics in column 3. It turns out that the variables are jointly weakly significant (p-value = 0.059). In particular, the findings suggest that the partner's personality matters. The partner's extraversion is significant for one's individual choice after the dyad choice. These results provide evidence that the preceding pair choice influences subsequent individual choice.

The findings regarding partner characteristics do not hold in the last column, where we use data from those individuals who performed the individual task before the pair task.[23] For these people, the dyad choice should not influence the individual choice in the way implied by Equation 1. On the contrary, the pair choice is the consequence of a bargaining process that is governed by the bargaining weight. That is, $x_p = w_p x^*_{i,p} + (1 - w_p) x^*_{-i,p}$. Since the dependent variable in column 4 is one's true choice ($x^*_{i,p}$), the results that are statistically significant should originate from the reverse of the bargaining model; that is,

$$x^*_{i,p} = \frac{1}{w_p} x_p - \frac{w_p}{1 - w_p} x^*_{-i,p}. \tag{8}$$

Indeed, the results in column 4 show that the pair choice is significant. On the other hand, most of the other explanatory variables are insignificant. This is, however, not surprising because the equation is (intentionally) misspecified. In particular, the partner's characteristics are jointly insignificant (p-value = 0.489). By construction, the specification has little explanatory power; the adjusted R^2 of 0.201 is low compared to 0.482 for the comparable specification in column 3.

Some other findings in Table 6 are worth noting. First, we find that it is difficult to explain one's individual choice after the pair choice by one's own characteristics alone. In column 1, all characteristics except conscientiousness are insignificant. Second, in column 3, it is difficult to interpret why a certain personality trait has a positive or a negative effect. This is because our estimation equation is not a structural equation that identifies determinants for one's choice. For instance, we find that one's conscientiousness has a significant effect on the individual

[23] The regression analysis in column 4 of Table 5 is based on 93 individuals since one observation is lost because of missing data for the partner's age.

choice. However, we do not know if that personality matters because it affects one's psychological tendency to follow the preceding pair choice or because it affects one's bargaining power. It is also possible that personality directly affects the degree of one's risk aversion, as psychologists have described personality traits as reflecting fundamental dispositions that influence broad and specific attitudes, perceptions, and behavior (Barrick and Mount 1991; Johnson, Rosen, and Djurdjevic 2011).

Finally, we find that, although it is again difficult to interpret the direction of the effect, non–task-related social activity while bargaining does affect subsequent individual choice. Somewhat surprisingly, the arguments made by one's partners do not have any lasting effect on the individual choice, nor do any of the other variables from the audio-video recordings.

7. Conclusions

Many risky decisions are undertaken by pairs, but relatively few studies have systematically looked at jointly made choices. Studies that have considered pairs or small groups have been focused primarily on determining if group choices differ from individual choices. The results have been mixed in terms of risk tolerance and seem to suggest that the decision structure (negotiation, voting, and so on) may matter. Our research focuses on negotiation and goes a step further by considering how an individual's demographic and personality characteristics affect bargaining weight and the willingness of an individual to deviate from the group.

We use the common risk attitude elicitation tool of Holt and Laury (2002) with prizes ranging from \$0.50 to \$19.25 to directly compare the risk-taking behavior of individuals and dyads. Our subjects completed the task twice, once individually and once in a pair (video and audio recorded) with the ordering determined randomly. We find no difference in the number of safe choices made by individuals and pairs when considering only subjects' initial task, which constitutes a between-subjects comparison. However, we find some evidence suggesting an ordering effect. The behavior of pairs does not depend on the task order, but subjects who had previously made decisions in a pair took on more risk individually than those making an individual choice first.

In only a few cases did the pair make a choice more extreme than the individual choices of the members. Instead, pair behavior appears to be consistent with a more traditional economic approach where the pair decision is the result of a bargaining process between the members. We find evidence that an individual's bargaining weight is dependent on relative age and dyad gender composition. The relatively older the more risk-averse person is, the less bargaining power the person will have, suggesting that the younger person drives risk-taking behavior. The more risk-averse person has less power in all male pairs, indicating evidence of a peer pressure effect between males and leading to more risk-taking behavior. We find no evidence to suggest that relative personality differences, as captured by the Big Five, have an effect on the bargaining process. Of course, this lack of significance may be due to our sample size, which is relatively large for a lab study on risk taking, but is small in comparison to the large-scale individual experiments, field studies, and meta-analyses that are often used to assess personality. We also find no evidence that conversational leadership or physical superiority, as measured from the audio and video recordings of the pair task, has any significant impact on bargaining power.

We do find that individual behavior is influenced by a preceding dyad decision and the personality characteristics of one's partner. One's partner's extraversion significantly impacts one's own subsequent individual choice. We also find that more agreeable people are less likely to simply follow the pair choice when making a subsequent individual choice. These results are interesting in that they suggest that risk-taking behavior can be influenced by factors not typically considered germane to the task. Thus, this article adds to a growing literature on how personality affects economic decision making. More broadly, our results demonstrate the need to consider not only "individual" versus "group" behavioral outcomes but also within-group dynamics and how group member personalities affect economic choices.

References

Bakeman, R. 2000. Behavioral observation and coding. In *Handbook of research methods in social and personality psychology*, edited by H. T. Reis and C. M. Judd. Cambridge: Cambridge University Press, pp. 138–59.

Baker, R., S. Laury, and A. Williams. 2008. Comparing small-group and individual behavior in lottery-choice experiments. *Southern Economic Journal* 75(2):367–82.

Baltussen, G., T. Post, M. J. Van den Assem, and R. H. Thaler. 2008. Deal or no deal? Decision making under risk in a large-payoff game show. *American Economic Review* 98(1):38–71.

Barrick, M. R., and K. Mount. 1991. The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology* 44:1–26.

Barsky, R., F. Juster, M. Kimball, and M. Shapiro. 1997. Preference parameters and behavioral heterogeneity: An experimental approach in the Health and Retirement Study. *Quarterly Journal of Economics* 112(2):537–79.

Bateman, I., and A. Munro. 2005. An experiment on risky choice amongst households. *Economic Journal* 115:C176–89.

Bone, J., J. Hey, and J. Suckling. 1999. Are groups more (or less) consistent than individuals? *Journal of Risk and Uncertainty* 18:63–81.

Boone, C., B. De Brabander, and A. van Witteloostuijn. 1999. The impact of personality on behavior in five prisoner's dilemma games. *Journal of Economic Psychology* 20:343–77.

Bornstein, G., T. Kugler, and A. Ziegelmeyer. 2004. Individual and group decisions in the centipede game: Are groups more "rational" players? *Journal of Experimental Social Psychology* 40(5):599–605.

Browning, M., F. Bourguignon, P. Chiappori, and V. Lechene. 1994. Income and outcomes: A structural model of intrahousehold allocation. *Journal of Political Economy* 102(6):1067–96.

Cason, T. N., and V. L. Mui. 1997. A laboratory study of group polarisation in the team dictator game. *Economic Journal* 107(September):1465–83.

Clark, J., M. T. Boccaccini, B. Caillouet, and W. F. Chaplin. 2007. Big Five personality traits, jury selection, and case outcomes in real criminal and civil cases. *Criminal Justice and Behavior* 34:641–60.

Cooper, D. J., and J. H. Kagel. 2005. Are two heads better than one? Team versus individual play in signaling games. *American Economic Review* 95:477–509.

Cooper, D. J., and J. H. Kagel. 2009. Equilibrium selection in signaling games with teams: Forward induction of faster adaptive learning? *Research in Economics* 63:216–24.

Costa, P. T., and R. R. McCrae. 1992. *The NEO PI-R professional manual*. Odessa, FL: Psychological Assessment Resources.

Cox, J. C. 2002. Trust, reciprocity, and other-regarding preferences: Groups vs. individuals and males vs. females. In *Advances in experimental business research*, edited by R. Zwick and A. Rapoport. Dordrecht: Kluwer Academic Publishers, pp. 331–50.

Cox, J. C., and S. C. Hayne. 2006. Barking up the right tree: Are small groups rational agents? *Experiential Economics* 9:209–22.

Deck, C., J. Lee, and J. Reyes. 2008. Risk attitudes in large stake gambles: Evidence from a game show. *Applied Economics* 40(1):41–52.

De Palma, A., N. Picard, and A. Ziegelmeyer. 2011. Individual and couple decision behavior under risk: Evidence on the dynamics of power balance. *Theory and Decision*. 70(1): 45–64.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. Wagner. 2005. Individual risk attitudes: New evidence from a large, representative, experimentally-validated survey. IZA Discussion Paper No. 1730.

Gill, C. M., and G. P. Hodgkinson. 2007. Development and validation of the Five-Factor Model Questionnaire (FFMQ): An adjectival-based personality inventory for use in occupational settings. *Personnel Psychology* 60:731–66.

Goldberg, L. R. 1999. A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In *Personality psychology in Europe*, volume 7, edited by I. Mervielde, I. Deary, F. De Fruyt, and F. Ostendorf. Tilburg: Tilburg University Press.

Gunnthorsdottir, A., K. McCabe, and V. Smith. 2002. Using the Machiavellianism instrument to predict trustworthiness in a bargaining game. *Journal of Economic Psychology* 23:49–66.

Harrison, G. W., M. I. Lau, E. E. Rutsrom, and M. Tarazona-Gomez. 2005. Preferences over social risk. Working Paper, University of Durham.

Holt, C., and S. Laury. 2002. Risk aversion and incentive effects. *American Economic Review* 92:1644–55.

Isenberg, D. 1986. Group polarization: A critical review and meta-analysis *Journal of Personality and Social Psychology* 50:1141–51.

Johnson, R., C. Rosen, and E. Djurdjevic. 2011. Assessing the impact of common method variance on higher order multidimensional constructs. *Journal of Applied Psychology* 96(4):744–61.

Kocher, M., S. Strauss, and M. Sutter. 2006. Individual or team decision-making—Causes and consequences of self-selection. *Games and Economic Behavior* 56(2):259–70.

Kocher, M. G., and M. Sutter. 2005. The decision maker matters: Individual versus group behaviour in experimental beauty contest games. *The Economic Journal* 115(500):200–23.

Laughlin, P. R. 1980. Social combination process of cooperative problem-solving groups at verbal intellective tasks. In *Progress in social psychology*, volume 1, edited by M. Fishbein. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 127–55.

Laughlin, P. R., and A. L. Ellis. 1986. Demonstrability and social combination processes on mathematical intellective tasks. *Journal of Experimental Social Psychology* 22:170–89.

Luhan, W. J., M. G. Kocher, and M. Sutter. 2009. Group polarization in the team dictator game reconsidered. *Experimental Economics* 12(1):26–41.

Malouff, J. M., E. B. Thorsteinsson, and N. S. Schutte. 2006. The five-factor model of personality and smoking: A meta-analysis. *Journal of Drug Education* 36:47–58.

Masclet, M., N. Colombier, L. Denant-Boemont, and Y. Loheac. 2009. Group and individual risk preferences: A lottery-choice experiment with self-employed and salaried workers. *Journal of Economic Behavior and Organization* 70(3):470–84.

Mazzocco, M. 2004. Saving, risk sharing, and preferences for risk. *American Economic Review* 94(4):1169–82.

McCrae, R. R., and P. T. Costa. 1989. Reinterpreting the Myers-Briggs type indicator from the perspective of the five-factor model of personality. *Journal of Personality* 57:17–40.

McCrae, R. R., and P. T. Costa. 1990. *Personality in adulthood.* New York: Guilford Press.

McElroy, T., K. Dowd, and K. 2007. Susceptibility to anchoring effects: How openness-to experience influences responses to anchoring cues. *Judgment and Decision Making* 2:48–53.

Moreland, R. 2010. Are dyads really groups? *Small Group Research* 41:251–67.

Perugini, M., J. Tan, and D. Zizzo. 2010. Which is the more predictable gender? Public good contribution and personality. *Economic Issues* 15:83–110.

Saulsman, L. M., and A. C. Page. 2004. The five-factor model and personality disorder empirical literature: A meta-analytic review. *Clinical Psychology Review* 23:1055–85.

Schmitt, P., R. Shupp, K. Swope, and J. Mayer. 2008. Pre-commitment and personality: Behavioral explanations in ultimatum games. *Journal of Economic Behavior and Organization* 3–4:597–605.

Sheremeta, R., and J. Zhang. 2010. Can groups solve the problem of overbidding in contests?" *Social Choice and Welfare*. 35(2):175–97.

Shupp, R., and A. Williams. 2008. Risk preference differentials of small groups and individuals. *Economic Journal* 118:258–83.

Stoner, J. 1968. Risky and cautious shifts in group decisions: The influence of widely held values. *Journal of Experimental Social Psychology* 4:442–59.

Sutter, M., M. G. Kocher, and S. Strauss. 2009. Individuals and teams in auctions. *Oxford Economic Papers* 61(2):380–94.

Swope, K., J. Cadigan, P. Schmitt, and R. Schupp. 2008. Personality preferences in laboratory economics experiments. *Journal of Socio-Economics* 37:998–1009.

Trobst, K. K., J. S. Wiggins, P. T. Costa, J. H. Herbst, R. R. McCrae, and H. L. Masters. 2000. Personality psychology and problem behaviors: HIV risk and the five-factor model. *Journal of Personality* 68:1233–52.

Van Knippenberg, D., B. Van Knippenberg, and E. Van Dijk. 2000. Who takes the lead in risky decision making? Effects of group members' risk preferences and prototypicality. *Organizational Behavior and Human Decision Processes* 83(2):213–34.

Williams, K. 2010. Dyads can be groups (and often are). *Small Group Research* 41:268–74.

Wooten, K. C., T. A. Timmerman, and R. Folger. 1998. The use of personality and the five-factor model to predict new business ventures: From outplacement to start-up. *Journal of Vocational Behavior* 54:82–101.

Journal of Economic Psychology 53 (2016) 54–66

Contents lists available at ScienceDirect

Journal of Economic Psychology

journal homepage: www.elsevier.com/locate/joep

Influence and choice shifts in households: An experimental investigation

Xiaojun Yang [a,*], Fredrik Carlsson [b,1]

[a] National School of Development, Peking University, 100871 Beijing, China
[b] Department of Economics, University of Gothenburg, Box 640, 405 30 Gothenburg, Sweden

A R T I C L E I N F O

Article history:
Received 24 March 2014
Received in revised form 13 November 2015
Accepted 16 November 2015
Available online 2 December 2015

JEL classification:
C91
C92
C93
D10

PsycINFO classification:
2100
2260

Keywords:
Individual decisions
Joint decisions
Patience
Choice shifts
Rural China

A B S T R A C T

In this paper, we examine the relative influence of individual decisions on joint household decisions, and whether and to what extent joint choices are more or less patient than individual choices in households. We find that both spouses have a significant influence on joint decisions, whereas husbands on average have a stronger influence than wives. Moreover, we find a substantial share of choice shifts from individual to joint household decisions, i.e. joint decisions are either more patient or more impatient than both individual choices. A number of observable characteristics are significantly correlated with these shifts in preferences from individual decisions to joint decisions.

1. Introduction

The empirical literature on household decision-making is by now extensive. Previous studies looking at actual decisions in the household suggests that the outcomes of household decisions depend on who in the household has control over the resources (Thomas, 1990; Browning, Bourguignon, Chiappori, & Lechene, 1994; Duflo, 2003; Lundberg, Pollak, & Wales, 1997; Namoro & Roushdy, 2008; Phipps & Burton 1998). In addition, by changing the control of income or access to financial assets, a set of field experiments reveal the importance of financial control in improving women's decision power and the decisions of the household (Ashraf, 2009; Ashraf, Karlan, & Yin, 2010; De Mel, McKenzie, & Woodruff, 2009; Mani, 2010;

* Corresponding author. Tel./fax: +86 10 62767657.
E-mail addresses: xjyang@nsd.pku.edu.cn (X. Yang), fredrik.carlsson@economics.gu.se (F. Carlsson).
[1] Tel.: +46 31 786 41 74; fax: +46 31 773 10 43.

http://dx.doi.org/10.1016/j.joep.2015.11.002
0167-4870/

Robinson, 2012). Recently, controlled experiments have also been used to investigate the influence of spouses on joint decisions (Abdellaoui, Haridon, & Paraschiv, 2011; Bateman & Munro, 2005; Carlsson, He, Martinsson, Qin, & Sutter 2012; Carlsson, Martinsson, Qin, & Sutter 2013; de Palma, Picard, & Ziegelmeyer, 2011). Apart from having control over the decision environment, the perhaps main advantage of controlled experiments is that both individual and joint decisions can be observed and related to each other. Controlled experiments have allowed researchers to directly estimate the spouses' respective influences and relate them to the characteristics of the households and the individual decision makers.

Intertemporal choices are of great importance to households since they often concern decisions such as savings, investments, and education. With the exceptions of Abdellaoui et al. (2011) and Carlsson et al. (2012), the literature on households' intertemporal decisions is relatively scarce. In this paper we study households' and spouses' intertemporal decisions in a controlled experiment where decisions are made both individually and jointly.[2] We investigate two aspects of household decisions making. The first is the influence on joint decisions by the husband and wife. The second is to what extent joint household decisions are more extreme or polarized than individual decisions.

Using the framework in Carlsson et al. (2012), we relate the individual choices to the choices made jointly and investigate to what extent the husband and wife influence joint decisions, and compare their relative influence. There is evidence that group decisions can become more extreme or polarized than individual decisions (Ambrus, Greiner, & Pathak, 2009; Cason & Mui, 1997; Eliaz, Raj, & Razin, 2006; Moscovici & Zavalloni, 1969; Shapiro, 2010; Stoner, 1968; Sunstein, 2000; Sunstein, 2002). Theoretically, there are a number of factors that can explain the difference between group and individual decisions as well as shifts in decisions, such as social comparison concerns (Levinger & Schneider, 1969), diffusion of responsibility (Eliaz et al., 2006), and altruistic concerns (Shapiro, 2010). Similar to group decisions, many household decisions are discussed and reflect, to varying extents, individual members' preferences. Studying to what extent joint household decisions are shifted is therefore of particular interest since the "diffusion of responsibility" and altruism play potentially important roles in household decision-making.

The type of shift we investigate is whether the joint decision on how much to allocate to the sooner date is outside the range of allocations to the sooner date given by the individual decisions. If a joint choice is more patient than the individual ones, we refer to it as a *patient shift*. The opposite case, where the joint choice is more impatient than the individual ones, is referred to as an *impatient shift*. Since a household is a group where individuals know their partners well, household joint intertemporal decisions could be useful in helping some individuals overcome for example self-control problem (Kono, Matsuda, Murooka, & Tanaka, 2011). In this sense, individual spouses could make more patient decisions in a joint setting than they would have made the decisions separately. A plausible explanation for why the joint choices are shifted to be more patient is that the spouses care about each other's preferences, and apply patient preferences when they know that the outcome will affect their spouse (Shapiro, 2010). Thus, even if, say, the husband is a hyperbolic discounter he might think it is better if the joint decision is more patient and is therefore willing to shift the decision.

We conduct an artefactual field experiment with 164 couples in rural China. In this experiment, couples made both separate and joint decisions on how much money to allocate to an early date and a later date. In addition, instead of the widely used multiple price list elicitation method in time preference literature (Andersen, Harrison, Lau, & Rutström, 2006; Andersen, Harrison, Lau, & Rutström, 2008; Coller & Williams, 1999; Harrison, Lau, & Williams, 2002; Tanaka, Camerer, & Nguyen, 2010), we employ the Convex Time Budget experimental method suggested by Andreoni and Sprenger (2012) to elicit individual and couple's intertemporal allocation decisions. The main advantage of this method is to account for the curvature of utility function.[3] The subjects can thus continuously allocate a certain amount of money between a sooner date and a later date. In the experiment, the subjects were asked to make ten different decisions where the interest rate and whether the early date is immediate or not are varied. With this approach we obtain detailed information about the characteristics of individual and joint choices, including the relationship between individual and joint decisions and to what degree joint choices can be shifted outside the range of individual choices.

The rest of this paper is organized as follows. In Section 2 we introduce the details about experimental design and procedure. Section 3 presents the analytical framework. We present and discuss descriptive results and regression results in Section 4. Finally, Section 5 concludes the paper.

2. Experimental design and procedure

2.1. Location of the experiment and description of the sample

The experiment was conducted in two counties of the Gansu province, which is located in the northwest of China. The two counties are Linxia and Jingning, which are located in the southwestern and southeastern parts of the province, respectively. As can be seen in Table 1A in Appendix A, in each county, we randomly chose three townships, and in total 13 villages were randomly selected.

[2] Here and henceforth, the respondents indicate husbands, wives or couples.

[3] There have been some concerns about the Convex Time Budget, in particular regarding the amount of information actually obtained from the experiments (Harrison, Lau, & Rutström, 2013). The main concern is the potentially large amount of corner observations, i.e. subject that either allocates all the money to the early date or the late date alternative.

In eight of the villages we randomly chose 10–25 households in each village with official marital status from the village registration list provided by the village leaders. In the other five villages, we randomly selected around five households in each village, also with official marital status.[4] With the assistance of one village cadre, two randomly matched enumerators (always one male and one female) approached the selected households. If both the husband and wife voluntarily agreed to be interviewed after our welcome announcement, the village cadre left. If one of the spouses was not home when the enumerators arrived at their house, the enumerators waited for a while or made an appointment to come back later. We had to make sure to interview the selected households in each village within one day in order to keep information about the experiment from spreading. If an appointment could not be made or if one spouse in a couple refused to be interviewed,[5] the enumerators visited the neighbor instead. Finally, 164 couples agreed to voluntarily participate in the experiment.

Table 1 describes the summary statistics of the sampled households. The average ages of the husbands and wives are 49 and 46 years, respectively. On average, the husbands have 5 years of education and the wives 2.5 years. As regards individual questions, husbands and wives have surprisingly similar responses. For example, the average income contribution to the households of the wives is around 40%. Husbands are the main decision makers in everyday life, but wives have more decision power when it comes to daily expenses such as food and clothes. As for the common household characteristics, the average household has five members, and the average length of marriage is 26 years. In 2010, the average household's gross income per capita was 7064 yuan.[6]

2.2. Experimental design

We apply the Convex Time Budget method suggested by Andreoni and Sprenger (2012) to investigate subjects' intertemporal choices.[7] In Table 2, the 10 intertemporal choice sets for each respondent are described. There are only two timeframes with the same delay time of one month: the near period between today and one month from today and the far period between two months and three months from today. To investigate whether respondents have present-biased preferences, we use "today" not "tomorrow" in the experimental design. However, this could imply different transaction costs between payments today and future payments (Andersen et al., 2008). As can be seen in the Table 1, both husband and wife expressed quite high level of trust that they would receive the future payments. In addition, respondents were told that they would get two vouchers, one for sooner payments and one for later payments, signed by the project coordinator. The voucher indicated the amount of cash and corresponding date the respondent could redeem the money. The five interest rates we used in the experiment were tested and decided upon based on the results of the pilot experiment.[8] Respondents needed to allocate the given 20 tokens between a sooner and a later date with increasing interest rates.

As described in detail below, subjects were presented with two plates. The red plate represented the sooner date (today or two months from today) and the orange plate represented the later date (one month from today or three months from today). Their task was to decide how many tokens to put on each plate, where in all choices each token was worth 2 yuan if it was allocated to the red plate. One token was worth $2 \times (1 + r)$ yuan if it was allocated to the orange plate. r is the rate of return for waiting, and it increased from the first choice to the fifth choice.

The respondents made both individual and joint decisions. As described below the order was randomly determined. When they made the individual choices they were clearly told that the money was theirs, and when they made the joint choices they would receive the same amount each. Thus, even when the decisions were made jointly, each spouse would receive their own individual money. The rationale to have this design is to make the individual decisions and joint decisions comparable. It is of course possible that the individual choices were made taking into consideration the preferences of the spouse, and we have no way to control for that. However, what we did was to stress that the choices would not be revealed to the spouse and that the money was individual and not to be paid to the household.

2.3. Experimental procedure

We employed and trained 10 interviewers, from now on called experimenters, to conduct the experiment. Once a couple had agreed to participate in the whole survey, one of the experimenters gave a brief introduction of the tasks. Then the

4 The survey mainly consisted of two separate experiments. In the eight villages, we mainly conducted the experiments with individual and joint decisions as used in current paper. In the other five villages, we mainly conducted another experiments about respondents' own decisions, predictions and decisions for spouse. Since we control for the effect of who has initial control over the tokens in the joint decisions, we increase sample size by randomly selecting additional respondents in the other five villages (see the sample distribution in Table 1A).

5 Three households refused to be interviewed. Among them, two households could not participate in the survey mainly because the wives stated they were too busy. One household refused to continue the experiment when the enumerators told them they could obtain some payments from our experiment. They did not tell us the concrete reason—they just did not want to continue.

6 At the time of the experiment, 1 USD = 6.59 CNY.

7 An alternative method to account for utility function curvature is to elicit both risk and time preferences using multiple price list method, which is referred as Double Multiple Price List (DMPL) method (Andersen et al., 2008). Andreoni, Kuhn, and Sprenger (2015) argue that the large proportion of corner observations with the CTB method improves the out-of-sample predictive validity compared to DMPL method. In addition, Angerer, Lergetporer, Glätzle-Rützler, and Sutter (2015) use both multiple price list method and time-investment-exercise (simplified version of CTB) to elicit children' time preferences, and find that both methods yield similar results and the behaviors can be explained by the same factors.

8 During the pilot studies, we first used the interest rates used by Gine, Goldberg, Silverman, and Yang (2012), i.e., 0.1, 0.25, 0.5, 0.75, and 1. However, especially at the high interest rates, there were almost no trade-offs; hence we reduced the rates to 0.05, 0.1, 0.25, 0.4, and 0.6.

Table 1
Descriptive statistics of individual and household characteristics ($N = 164$ households).

	Husband		Wife	
	Mean	Std. dev.	Mean	Std. dev.
Individual characteristics				
Age (years)	48.78	9.34	46.26	9.11
Higher than primary school (1 = yes)	0.50		0.19	
Communist party member (1 = yes)	0.12		0.01	
Individual attitudes				
General decision maker (1 = husband; 2 = joint; 3 = wife)	1.24	0.46	1.38	0.59
Wife income contribution share	0.40	0.17	0.39	0.17
Husband income contribution share	0.60	0.17	0.61	0.17
Decision maker on savings (1 = husband; 2 = joint; 3 = wife)	1.31	0.49	1.34	0.51
Decision maker on daily expense (1 = husband; 2 = joint; 3 = wife)	2.36	0.78	2.18	0.81
Decision maker on durable goods (1 = husband; 2 = joint; 3 = wife)	1.55	0.53	1.55	0.61
Decision maker on expensive fixed asset (1 = husband; 2 = joint; 3 = wife)	1.55	0.52	1.50	0.54
Trustiness on the future payments (1 = totally do not trust; 2 = do not trust; 3 = neither trust nor distrust; 4 = trust; 5 = totally trust)	4.56	0.82	4.49	0.77
Household characteristics				
Household is minority (1 = yes)	0.15			
Household population (persons)	4.98	1.50		
The length of marriage (years)	26.06	9.80		
Log of equivalence scaled total gross income (yuan); Equivalence = $(\text{adults} + 0.5 * \text{children})^{0.75}$	9.03	0.68		

Table 2
Description of the 10 decisions in the time preference experiment.

Sooner date	Later date	Token budget	Interest rate	Sooner value of one token	Later value of one token
0	30	20	0.05	2	2.1
0	30	20	0.1	2	2.2
0	30	20	0.25	2	2.5
0	30	20	0.4	2	2.8
0	30	20	0.6	2	3.2
60	90	20	0.05	2	2.1
60	90	20	0.1	2	2.2
60	90	20	0.25	2	2.5
60	90	20	0.4	2	2.8
60	90	20	0.6	2	3.2

couple together answered a set of questions about the household. The rest of the procedure depended on the order of the parts of the experiment (see Section 2.4). However, we will for simplicity only describe in detail one of the orders used.

In the version where individual decisions were made before the joint decisions, the respondents were (following the first initial questions) physically separated into two rooms where they could not hear each other; one experimenter followed the wife and one followed the husband. The experimenter read out the experimental instructions to the respondent, and the respondent was told that s/he could earn some money and that the amount earned depended on his/her decisions in the experiment. The respondent needed to make 10 separate decisions, and one of these decisions would be randomly chosen to be paid out by rolling a 10-sided dice at the end of the survey. The number that came up on the dice decided which choice would determine the respondent's earnings. Thus, each decision had an equal chance of being used in the end.

To help the respondents understand the experiment, they first made two trial decisions.[9] The purpose of the trial decisions is to help respondents make more informed decisions and avoid misunderstandings of the experimental tasks. The drawback with trial decisions is that the experiment takes too long and hence causes respondents to be fatigued. However, our experience from the pilot experiment was that the trial tasks were crucial for the understanding of the experiment. Once the experimenter was certain that the respondent had understood, s/he was asked to make the first five independent decisions about how to allocate 20 tokens between today and one month from today. To help the respondent remember which dates the two plates represented, the experimenter put a sign in front of each plate with the corresponding date and the value of a token. When the respondent finished her/his allocations, the experimenter translated the total tokens on each plate into Chinese yuan and wrote the decision on the whiteboard. The respondent also had the chance to revise the allocations before moving to the next choice.

[9] The trial decisions were about how to allocate 10 tokens between one month from today and two months from today. Before the respondent did this, the experimenter asked some control questions about the meaning of the plates and the tokens. The respondent started to make the trial decisions only when s/he had understood the meaning of the plates and the tokens. The trial decisions were the same regardless of the order between individual decisions and joint decisions.

The elicitation procedure was similar for the second five independent choices. Yet the respondent was reminded that s/he needed to wait for both the sooner payment (two months from today) and the later payment (three months from today).

When both the husband and the wife finished, they were brought together for the joint decisions. However, they did not have any prior information about the joint decisions when they finished their individual decisions. The couple was told that they would make 10 intertemporal choices similar to the individual decisions they had just made. The main difference was that both of them would obtain the same amount of payments according to one of the joint decisions, which would be randomly selected by rolling a 10-sided dice. Before each decision was made, they were encouraged to speak to each other and discuss the decision, as they needed to agree on how to allocate the money between the sooner and later dates. The couple followed the same elicitation method as the individual decisions. When the respondents finished the joint decisions, one of them was randomly selected to roll the 10-sided dice to determine the earnings. They then were separated to roll the 10-sided dice for one of individual decisions. The aim of this design is to mitigate the endowment effect of earnings from individual decisions on joint decisions.

On average, the whole survey lasted for one and a half hours for each household. The average experimental payment for each household was around 180 yuan, which equals three days of non-farm wages for a local full-time worker.

2.4. Order effects and initial control over the tokens

In the design, we control for two important order effects. The first one is about the order of individual and joint decisions. Half of the households made the individual decisions first and then the joint decisions. The other half of the households made the joint decisions first and then the individual decisions. While the natural order would be to first conduct the individual experiment and then the joint, we want to test if the ordering affects the behavior in the joint decision experiment. There could, for example, be learning effects, or the respondents may try to smooth out the earnings over time and the two parts of the experiment.

The second order effect concerns the order of the two parts of the time preference experiment. Half of the households answered the five questions regarding money allocated between today and one month from today first, while the other half started with the five questions regarding money allocated between two months and three months from today.

In addition, we control for experimenter effects by interchanging their interviewing subjects in each household. For example, if the male experimenter interviewed the husband and the couple in one household, then the female experimenter needed to interview the husband and couple in the next household.

Finally, in the joint experiment, to control for the effects of who had the initial control over the tokens on the joint decisions, we had four alternatives for how the tokens were initially distributed. The first reference situation was that the experimenter just put the 20 tokens between the husband and the wife, but did not say anything else about who was responsible to put tokens on the plates. The second situation was that the experimenter gave the 20 tokens to the wife, making her in charge of putting the tokens on the plates. In the third situation, the experimenter gave the 20 tokens to the husband, who was initially responsible to put the tokens on the plates. The fourth situation was that the experimenter gave 10 tokens to the wife and 10 tokens to the husband, making both of them in charge of putting the tokens on the plates. For all cases, both spouses could adjust the amount of tokens on the plates until they had reached an agreement, i.e., they were not told that only one or both should put the tokens on the plates.

3. Analytical framework

3.1. Underlying preferences

In the experiment, for a given interest rate, r, the subjects had to decide how much of a given initial amount of money to allocate to a sooner date, c_t, and a later date, $c_{t+\tau}$, where t indicates the sooner dates, i.e., $t = 0$ or $t = 60$ days; τ is the delay time, i.e., $\tau = 30$ days. c_t and $c_{t+\tau}$ represents the monetary allocations to the sooner date and later date, respectively, and $c_{t+\tau}$ includes the rates of return, r, in different choices. Since the experiment was fairly complex and we could not ask subjects to make too many decisions, we chose to keep the delay time constant. The data still allows us to estimate time preferences, present bias, and the curvature of the utility function, given a set of assumptions.

In our analysis we assume a constant absolute risk aversion (CARA) utility function $\left(u = -\frac{1}{\rho}e^{-\rho c_t}\right)$, where ρ is the coefficient of absolute risk aversion.[10] Following Andreoni and Sprenger (2012), the utility function and budget restriction can be written as:

$$U = u(c_t) + \beta\delta^{\tau}u(c_{t+\tau}) \quad (1)$$

$$\text{s.t.} \ \frac{c_{t+\tau}}{1+r} + c_t = m \quad (2)$$

[10] The main advantage of CARA utility function is that the background parameters can be dropped out from the marginal condition. Since we do not have background consumption information, we prefer the CARA utility function in the following model estimation. Actually, we also get the similar estimated parameters based on constant relative risk aversion (CRRA) utility function.

where β is the present bias parameter ($\beta > 0$), δ is the discount factor, τ is the delay time, and m is the experimental budget. The marginal condition for a utility maximizing individual can be written:

$$\exp(-\rho(c_t - c_{t+\tau})) = \left\{ \begin{array}{ll} \beta\delta^{\tau}(1+r) & \text{if } t = 0 \\ \delta^{\tau}(1+r) & \text{if } t > 0 \\ & \end{array} \right\} \tag{3}$$

Take logs and rearrange, Eq. (3) can be rewritten as:

$$c_t - c_{t+\tau} = \frac{\ln\beta}{-\rho}P + \frac{\ln\delta}{-\rho}\tau + \frac{1}{-\rho}\ln(1+r) \tag{4}$$

where P is a dummy variable equal to one if $t = 0$. Note that in our case, τ is constant, so the second term is the intercept. We estimate the following model

$$(c_t - c_{t+\tau})_{ik} = \frac{\ln\beta}{-\rho}P + \frac{\ln\delta}{-\rho}\tau + \frac{1}{-\rho}\ln(1+r) + \varepsilon_{ik} \tag{5}$$

where k is the choice situation and ε_{ik} is an additive mean-zero error term. Because of the issue with corner solution, we estimate a two-limit Tobit model. In reduced form the model is expressed as:

$$(c_t - c_{t+\tau})_{ik} = \gamma_1 P + \gamma_2\tau + \gamma_3\ln(1+r) + \varepsilon_{ik} \tag{6}$$

The parameters of interest can be obtained by non-linear combination:

$$\widehat{\rho} = -\frac{1}{\widehat{\gamma}_3} \quad \widehat{\delta} = \exp\left(\frac{\widehat{\gamma}_2}{\widehat{\gamma}_3}\right) \quad \widehat{\beta} = \exp\left(\frac{\widehat{\gamma}_1}{\widehat{\gamma}_3}\right) \tag{7}$$

3.2. Relative influence of the spouses

We follow the approach outlined in Carlsson et al. (2012) and estimate the influence of each spouse by explaining the joint decisions by the individual decisions. Thus, our focus is on choices and not the underlying preference parameters. The joint sooner allocation decision for household i in choice situation k is specified as

$$c^J_{t_{ik}} = \alpha + \mu^H c^H_{t_{ik}} + \mu^W c^W_{t_{ik}} + \varepsilon_{ik} \tag{8}$$

where J, H, and W denote decisions made jointly, by the husband, and by the wife respectively, and ε_{ik} is an error terms reflecting unobservable factors that influence the joint decisions. Since $c^J_{t_{ik}}$ (allocation to the sooner date) is censored by 0 and 40, we will employ a Tobit model to estimate Eq. (8). The parameters μ^H and μ^W are measures of the husband's and wife's influence on the joint decision. From this also follows that we do not impose any restrictions on the influence parameters, i.e. we do not restrict the sum to be equal to unity. Finally, our specification does not rule out aspects such as altruistic concern for the spouse's consumption (or envy for that matter). This will be revealed by a relatively stronger influence of the spouse's preferences. The ratio between the two influence parameters, $\lambda = \widehat{\frac{\mu^W}{\mu^H}}$, is then a measure of the relative influence of the wife and the husband. If the ratio is above one, then the wife has a stronger influence on the joint decision.

3.3. Choice shifts

Next we investigate to what extent the joint choices are more extreme than the individual choices. Again, we focus on choices and not the underlying preference parameters. Since we focus on the comparison between individual and joint decisions, we use the sooner allocations to classify the joint decisions into three categories for household i in choice situation k: impatient shift, in between, and patient shift.[11] We employ a multinomial logit model using these three categories as the dependent variable, and investigate the factors that could explain the likelihood of a household joint decision ending up in a certain category.

$$P_{1ik}\left[c^J_{tik} > \text{Max}\{c^H_{tik}, c^W_{tik}\}\right] = f\left(|c^H_{tik} - c^W_{tik}|, r, t, x\right) + \varepsilon_{1ik} \tag{9}$$

$$P_{2ik}\left[c^J_{tik} \in [c^H_{tik}, c^W_{tik}]\right] = f\left(|c^H_{tik} - c^W_{tik}|, r, t, x\right) + \varepsilon_{2ik} \tag{10}$$

$$P_{3ik}\left[c^J_{tik} < \text{Min}\{c^H_{tik}, c^W_{tik}\}\right] = f\left(|c^H_{tik} - c^W_{tik}|, r, t, x\right) + \varepsilon_{3ik} \tag{11}$$

[11] It actually does not affect the categories if we instead use the difference between the later and sooner allocations to measure the patience. But it is much easier to implement if we only use the sooner allocation.

The first probability in Eq. (9) above is the probability of an impatient shift, i.e., the amount of money allocated to the sooner date in the joint decisions is larger than both the husband's and the wife's individual allocations. The second probability is that the joint decision is in between the spouses' individual decisions (or exactly the same). The third is the probability of a patient shift, i.e. the amount of money allocated to the sooner date in the joint decisions is smaller than both the husband's and the wife's individual allocations.

To investigate how the potential conflicts between husband's and wife's preferences affect the likelihood of a shift, we include the absolute value of the difference between husband and wife individual decisions. We control for the interest rate, the sooner date (in particular if the sooner date is today or not), a vector additional characteristics of the households and spouses, and an error term.

The likelihood of a joint shift is thus assumed to be a function of the absolute value of the difference between husband and wife individual decisions. If all three outcomes are feasible we have that $\sum_{i=1}^{3} P_i$. However, all three outcomes will not always be feasible, mainly due to corner decisions of the spouses. For example, if one spouse makes an extremely patient individual decision, then there is no possibility of a joint patient shift. In those cases the number of potential outcomes will be reduced.

The error term captures unobservable factors such as unobserved preference heterogeneity and mistakes in the decisions. However, the standard formulation with iid errors is very restrictive. In order to allow for individual heterogeneity we estimate a so-called scaled multinomial logit model, which is a restricted version of the generalized mixed logit model (Fiebig, Keane, Louviere, & Wasi, 2010). The scale factor for households i is $\sigma_i = \exp\left(-\frac{\gamma^2}{2} + \gamma w_i\right)$, where w_i is a random variation across households, and γ is a structural parameter; the model assumes that w_i is normally distributed. In addition we allow for observed heterogeneity by including the absolute value of the difference between husband and wife individual decisions as an independent variable. The intuition is that if there is small difference in preferences, then the variance will increase and thus the randomness of the outcome.

4. Results

4.1. Individual and joint allocations

In Table 3, we summarize the average allocations, in Chinese yuan, made to the sooner dates by the husbands, wives, and couples for all the decisions. As can be seen, the allocation to the sooner date decreases when the rate of return increases, which is an indication of that the subjects are aware of the basic trade-offs they face in the choice tasks. We also observe that wives' and joint decisions are on average more patient than husbands' decisions. Although most joint decisions are in between the husbands' and wives' decisions, husbands' decisions are much closer to joint decisions compared with wives' decisions. It indicates husbands could have more influence on the joint decisions at this aggregate level.

The table also reports the share of allocations that are corner allocations, i.e., when the subject allocates zero yuan to the sooner date and thus allocates everything to the later date. As expected, the share of corner allocations increases when the rate of return increases. More specifically, at the individual and couple level, 142 husbands, 149 wives and 152 couples have corner allocations (either zero or 40 yuan allocated to the sooner date), respectively.[12] In addition, at choice level, 59 percent of husbands' decisions, 63 percent of wives' decisions and 62 percent of joint decisions are allocated at corners. This result is quite comparable to Andreoni and Sprenger (2012).

As discussed in Section 3, we first employ the two-limit Tobit model to estimate Eq. (6). We estimate separate models for the husbands', wives', and joint decisions, and cluster the standard error at the household level. The estimated results are summarized in Table 4 (see column (1), (3), and (5)). As expected, the coefficient of the rate of return is negative and highly significant. The significant and positive sign of the present time dummy variable indicates that subjects on average have present-biased time preferences. Regarding the estimated preference parameters, the estimated coefficient of absolute risk aversion is statistically significantly different from zero, and very close to unity. In addition, different from Andreoni and Sprenger (2012), we find that the estimated present bias parameter $(\widehat{\beta})$ is statistically significantly different from one, whereas the daily discount factor is greater than one. Thus, both individual and joint decisions are present-biased. Note that since the present bias parameter is lower than one, the subjects still discount future payments.

To investigate what factors influence individual and joint decisions, we again employ the two-limit Tobit model to estimate Eq. (6) by including a set of individual and household characteristics. The estimated results are presented in Table 4 (see column (2), (4), and (6)). There are actually very few observable characteristics that have a statistically significant effect on the allocation decisions. For husband decisions, wife decisions and joint decisions, in minority households, subjects allocate more money to the sooner date, i.e., are more impatient. For husband decisions, the allocations to the sooner date decreases with wife's income contribution. Wife's income contribution could increase wife's relative influence in household, and wives are in general more patient than husbands as described in Table 3. In addition, older wives are more patient, whereas highly educated husband significantly decreases the allocations to later date in joint decisions. Nonetheless,

[12] Among the corner solutions, 52 husbands (32%), 46 wives (28%), and 55 couples (34%) do not have interior allocations, respectively (i.e. they allocate zero or 40 yuan in all 10 choices). This is quite similar to the share in Andreoni and Sprenger (2012): 37%.

Table 3
Husband's, wife's and joint allocations to the sooner dates in Chinese yuan.

Sooner date	Interest rate	Husband		Wife		Joint		|Husband − joint|	|Wife − joint|
		Mean	Share corner (%)	Mean	Share corner (%)	Mean	Share corner (%)		
0	0.05	22.5 (16.2)	23	24.1 (15.9)	17	20.8 (15.8)	24	8.6 (11.6)	11.4 (12.5)
0	0.1	18.4 (15.9)	28	17.9 (15.1)	26	16.5 (15.2)	30	8.6 (11.5)	10.7 (11.6)
0	0.25	12.7 (14.3)	39	10.0 (12.9)	48	10.3 (12.6)	44	7.1 (9.9)	9.4 (11.1)
0	0.4	9.7 (13.2)	49	7.0 (11.3)	59	7.8 (11.7)	52	6.1 (9.5)	7.1 (9.8)
0	0.6	7.1 (12.2)	62	4.3 (9.5)	76	4.9 (9.9)	70	4.4 (7.6)	5.4 (9.9)
60	0.05	16.8 (14.9)	30	12.7 (13.5)	38	14.7 (15.0)	37	10.5 (12.7)	11.0 (13.0)
60	0.1	11.9 (13.1)	40	8.9 (11.9)	47	9.6 (12.0)	47	8.2 (11.0)	8.3 (10.9)
60	0.25	8.2 (11.5)	51	4.8 (8.4)	64	6.2 (9.6)	58	5.3 (8.4)	5.6 (9.0)
60	0.4	5.8 (10.0)	60	2.9 (6.7)	77	4.1 (8.1)	68	4.2 (7.2)	3.5 (6.8)
60	0.6	3.8 (8.7)	73	2.0 (5.6)	84	2.5 (7.1)	83	2.8 (6.4)	2.5 (6.2)

Note:
1. Figures in the parentheses are standard deviation.
2. Share corner is the percentage of zero allocation to the sooner date.
3. The last two columns are the absolute difference between husband decisions and joint decisions, and between wife decisions and joint decisions.

Table 4
The determinants of husband's, wife's, and joint decisions; dependent variable is the difference between sooner and later allocations ($c_t - c_{t+\tau}$).

	Husband		Wife		Joint	
	(1)	(2)	(3)	(4)	(5)	(6)
Present time dummy (1 = today)	23.685*** (4.943)	23.060*** (4.886)	33.690*** (4.871)	33.479*** (4.770)	25.191*** (4.530)	24.916*** (4.458)
Time delay (30 days)	−0.314** (0.160)	0.923 (2.015)	−0.561*** (0.150)	0.584 (1.907)	−0.554*** (0.154)	1.558 (2.469)
Interest rate (log(1 + r))	−199.916*** (15.855)	−199.026*** (15.748)	−245.201*** (19.547)	−244.385*** (19.092)	−212.963*** (15.200)	−212.711*** (15.003)
Husband age (years)		−0.085 (0.443)				−0.627 (1.033)
Husband higher than primary school (1 = yes)		8.685 (9.015)				15.969* (8.763)
Wife age (years)				−0.912** (0.438)		−0.208 (1.092)
Wife higher than primary school (1 = yes)				10.640 (11.535)		−8.295 (11.765)
Wife's income contribution (%)		−49.163* (28.398)		−13.145 (26.609)		−38.501 (27.012)
Household is minority (1 = yes)		45.795*** (14.277)		46.355* (24.190)		36.884** (15.869)
Log of equivalence scaled total gross income (yuan)		−2.528 (6.218)		2.571 (5.835)		−2.078 (7.289)
If first separate then joint decision (1 = yes)		1.008 (9.840)		12.205 (7.954)		−5.709 (9.594)
If first five choices are between today and one month (1 = yes)		17.856* (9.229)		−4.577 (8.078)		11.188 (15.501)
Daily discount factor ($\widehat{\delta}$)	1.002 (0.001)		1.002 (0.001)		1.003 (0.001)	
Present bias ($\widehat{\beta}$)	0.888 (0.020)		0.872 (0.015)		0.888 (0.017)	
CARA curvature ($\widehat{\rho}$)	0.005 (0.000)		0.004 (0.000)		0.005 (0.000)	
Observations	1640	1640	1640	1640	1640	1640
Uncensored	671	671	600	600	630	630
Log pseudolikelihood	−4473.217	−4419.397	−3948.506	−3877.814	−4156.064	−4094.861
Prob > F	0.000	0.000	0.000	0.000	0.000	0.000

Notes:
1. Experimenter effects and village dummies are controlled in model (2), (4) and (6). Initial tokens control dummies are also controlled in model (6).
2. All regressions are clustered at household level. Figures in parentheses are robust standard errors.
* Represents significance at 10%.
** Represents significance at 5%.
*** Represents significance at 1%.

Table 5
The influence of individual decisions on joint decisions; dependent variable is the sooner allocations (c_t).

	(1)	(2)	(3)
Husband's decision	0.475*** (0.022)	0.449*** (0.030)	0.461*** (0.034)
Wife's decision	0.328*** (0.022)	0.326*** (0.032)	0.321*** (0.038)
Husband's decision × if first separate then joint decision (1 = yes)		0.052 (0.040)	
Wife's decision × if first separate then joint decision (1 = yes)		0.008 (0.042)	
Husband's decision × 20 tokens to wife			0.067 (0.048)
Husband's decision × 20 tokens to husband			−0.058 (0.048)
Husband's decision × 10 tokens to each			0.069 (0.052)
Wife's decision × 20 tokens to wife			−0.027 (0.053)
Wife's decision × 20 tokens to husband			0.085 (0.054)
Wife's decision × 10 tokens to each			−0.047 (0.056)
Observations	1640	1640	1640
Uncensored	630	630	630
Log pseudolikelihood	−3315.429	−3314.453	−3310.339
Prob > F	0.000	0.000	0.000

Notes:
1. The results reported in table are marginal effects. Two order dummies, experimenter effects and village dummies are controlled in all model specifications.
2. All the regressions are clustered at household level. Figures in the parentheses are robust standard errors.
3. * and ** represent the significant level at 10% and 5%.
*** Represents the significant level at1%.

Table 6
Size of observed shifts.

	Mean	Std. dev	Median	No. of obs.
Impatient shift	9.456	8.323	8	180
Patient shift	9.213	7.946	8	155

household income does not significantly impact individual decisions or joint decisions. There is no statistically significant effect on decisions by the order of the individual and joint decisions. Husbands however, are more impatient when they first make five choices between today and one month.

4.2. Relative influence of the spouses

We now move to the relationship between the individual decisions and the joint decisions. We estimate models where we explain the joint decisions with the husbands' and wives' decisions, as specified in Eq. (8). We employ a Tobit model to estimate all model specifications. The standard errors are clustered at household level, and the marginal effects are presented in Table 5.

In the first model, we only control for the husbands' and wives' individual decisions. As can be seen, both spouses have a significant impact on the joint decision in the sense that there is a positive and significant correlation between the individual decisions and the joint decision. However, both coefficients are well below one, suggesting that on average neither spouse has complete control over the joint decision. The relative influence of the two spouses can be estimated as the ratio between the wife's individual decision coefficient and the husband's individual decision coefficient. This parameter is 0.69, which means that the husband on average has a stronger influence on the joint decision than the wife. The value of the relative influence parameter has a clear and simple explanation. It is the ratio of marginal effects of the two spouses' influence on the joint decisions. The husbands' influence parameter is around 0.48. This means that if the husband allocates, say, 10 yuan more to the sooner date in the individual experiment, then the allocation to sooner date in the joint experiment increases by 4.8 yuan. For the wife, the increase in the joint experiment for the same change is 69% of this, i.e. 3.3 yuan. Moreover, we can reject the hypothesis that the relative influence parameter is equal to one (p-value = 0.000).

In the second model, we add interaction terms between husband's and wife's decisions and a dummy variable equal to one if the individual decisions were made before the joint decisions. The estimated results in column (2) show that there is no significant order effect on the influence of individual decisions on joint decisions. In the third model, we interact the spouses' individual decisions with the treatment dummy variables concerning who had initial control over the tokens. As can be seen, none of the interaction terms are statistically significant. This is different from the study by de Palma et al. (2011), where, in an experiment on risky choices, women who ultimately implement the joint decisions show more decision power.

4.3. Choice shifts

In this section we examine to what extent joint decisions are more patient or impatient than individual decisions at the choice level. Based on the classification of responses in Section 3, we find that 11% of the joint choices are more impatient than both the husbands' and wives' individual choices, while 9% of the joint choices are more patient. Thus, in 80% of the choice situations, the joint choice is in between, or equal to, the spouses' individual choices. At the same time, a majority of the households experience a shift. In 27% of the households there is at least one impatient shift, in 25% there is at least one patient shift, and in 15% there are both patient and impatient shifts.

Furthermore, the size of shifts is often sizeable. Table 6 reports the mean and standard deviations of the observed shifts, measured as the difference between the joint allocation on the early period and the corresponding lowest or highest individual allocation. The minimum size of a shift is 2 yuan (since each token is worth 2 yuan) and the maximum size is 40 yuan. The average size of both patient and impatient shifts is around 9 yuan, i.e., a little bit more than 4 out of 20 tokens.[13] This means that what we observe most likely is not simply due to noise in the joint decision as compared with the individual choices.

Next we estimate a multinomial logit model where the dependent variable is the three joint shift categories and the standard errors are clustered at the household level. The first model we report is a standard multinomial logit model that includes four dummy variables: the time order dummy (if the first five choices are made between today and one month), and three initial tokens control dummies (20 tokens given to wife, 20 tokens given to husband, 10 tokens given to each). We do this because the scaled multinomial logit model would not converge when we include the four dummy variables. The second model we report is thus the scaled multinomial logit model without including the dummy variables. Table 7 reports the model results with the patient shift as the reference category.

[13] We also find that among the impatient shifts, 49% are closer to the husband's individual choice, i.e. the husband is more impatient than the wife, and 17% are closer to the wife's individual choice. Among the patient shifts, 41% are closer to the husband's individual choice, and 41% are closer to the wife's individual choice.

Table 7
The determinants of the likelihoods of impatient and patient shifts.

		MNL with dummies	Scaled MNL
Impatient shift		7.406*** (2.200)	9.631*** (2.522)
In between		4.860*** (1.820)	5.497*** (1.659)
Impatience shift	Interest rate (r)	−3.575*** (0.732)	−3.987*** (1.042)
	Present time dummy (1 = today)	0.248 (0.254)	0.358 (0.284)
	Husband age (years)	−0.102*** (0.039)	−0.121*** (0.045)
	Husband higher than primary school (1 = yes)	0.579** (0.279)	0.680** (0.318)
	Wife age (years)	0.049 (0.039)	0.060 (0.043)
	Wife higher than primary school (1 = yes)	−1.145*** (0.381)	−1.170*** (0.454)
	Wife's income contribution (%)	−1.200 (0.971)	−1.008 (1.058)
	Household is minority (1 = yes)	−2.139*** (0.429)	−2.198*** (0.560)
	Log of equivalence scaled total gross income (yuan)	0.491** (0.220)	−0.696*** (0.256)
	If first five choices are between today and one month (1 = yes)	0.318 (0.278)	
	If initial 20 tokens given to wife (1 = yes)	0.207 (0.380)	
	If initial 20 tokens given to husband (1 = yes)	0.337 (0.389)	
	If initial 10 tokens given to each (1 = yes)	−0.563 (0.367)	
In between	Interest rate (r)	−1.187*** (0.598)	−1.094* (0.616)
	Present time dummy (1 = today)	0.072 (0.203)	0.086 (0.212)
	Husband age (years)	−0.010 (0.029)	−0.005 (0.032)
	Husband higher than primary school (1 = yes)	0.182 (0.224)	0.191 (0.232)
	Wife age (years)	−0.004 (0.029)	−0.009 (0.032)
	Wife higher than primary school (1 = yes)	−0.179 (0.292)	−0.120 (0.310)
	Wife's income contribution (%)	−2.097*** (0.774)	−2.158*** (0.788)
	Household is minority (1 = yes)	−1.212*** (0.292)	−1.091*** (0.289)
	Log of equivalence scaled total gross income (yuan)	−0.257 (0.180)	−0.318* (0.173)
	If first five choices are between today and one month (1 = yes)	0.301 (0.226)	
	If initial 20 tokens given to wife (1 = yes)	0.225 (0.293)	
	If initial 20 tokens given to husband (1 = yes)	−0.004 (0.321)	
	If initial 10 tokens given to each (1 = yes)	−0.379 (0.281)	
Variance function			
Tau			0.401* (0.232)
Absolute difference between husband's and wife's sooner allocation			−3.025 (4.459)
Obs.		1526	1526

* Denotes statistical significance at the 10%, levels.
** Denotes statistical significance at the 5% levels.
*** Denotes statistical significance at the 1% level.

Note that the multinomial and scaled multinomial logit model results are similar in relative size and statistical significance. Moreover, all the four dummy variables are statistically insignificant. We therefore from now on focus on the scaled multinomial logit model. To begin with we find that the likelihood of a shift depends on the interest rate, but not on whether the early payment involves today or not. The likelihood of a patient shifts is higher with a higher interest rate, and the likelihood of an impatient shift or a joint choice in between individual choices is lower with a higher interest rate. That the likelihood of a shift does not depend on whether the sooner payment is today or not is interesting. It suggests that shifts are not driven by a change in the present bias concern, but instead by a change in the discount factor.

Regarding households characteristics, there are a number of interesting results. The likelihood of an impatient shift is higher in households where the husband has higher than primary education, whereas the likelihood of this shift decreases with wife's education level. In addition, the likelihood of shifts is lower in households where the wife has a higher income contribution. Furthermore, the likelihood of a patience shift is higher in minority households and in households with a relatively higher income. Given that we see patience as something advantageous for the households in the long run, it is thus more likely in these households that the joint decision is better than the individual decisions.

Regarding the variance function we find a statistically significant structural parameter, which indicates that there is an unobserved heterogeneity in variance across households. More importantly, the coefficient of the absolute difference in husband and wife allocation is not statistically significant, which means that we do not find any support for a hypothesis that the variance is larger for households with a smaller difference in individual decisions. Taken together with the fact that the size of the choice shifts is sizeable, it is not likely that the shifts are just random shifts.

5. Conclusions

In this paper we have investigated the relative influence of spouses' preferences on joint decisions, and the occurrence of choice shifts from individual decisions to household joint decisions regarding intertemporal choices.

We find that both spouses have a significant impact on the joint decision. Wife's relative influence parameter is equal to 0.69, which means that the husband has a stronger influence than wife. Furthermore, we find that there are substantial shifts between individual and joint decisions. At the choice level, 11% of the joint choices are more impatient than the individual choices, while 9% are more patient. In addition, by accounting for the preference heterogeneity and mistakes in the decisions making, we employ a scaled multinomial logit model to investigate the relevant factors that could affect the likelihood of choice shifts. We find that the interest rate and a set of individual and household characteristics impact the likelihood of choice shifts.

One obvious question is of course whether these shifts are good or bad. Patience is often seen as a virtue, and as shown by Becker (1980), based on a conjecture of Ramsey (1928), income distribution in a long-run steady state is determined by the lowest discount rate; i.e., the household with the lowest discount rate will own all the capital. This conclusion of course rests on a number of simplifying assumptions, but, taking these as given, a more patient shift would be beneficial for the household. This is also supported by an empirical literature on the relationship between patience and life outcomes (see e.g. Chabris, Laibson, Morris, Schuldt, & Taubinsky, 2008; Moffitt et al., 2011; Sutter, Kocher, Glätzle-Rüetzler, & Trautman, 2013). For example, Sutter et al. (2013) conducted experiments, on risk and time preferences, with children and adolescents and found that impatience is a significant predictor of health-related behavior, savings and conduct at school.

In this paper, we find both patient shifts and impatient shifts. It indicates that there is no clear pattern in the sense that joint household choices tend to generate beneficial shifts, i.e., patient shifts. In addition, our findings provide additional evidence on the efficiency and rationality of group decisions. As discussed in the introduction, there is evidence that group decisions are more in line with the standard game-theoretical predictions of rationality and selfishness than individuals (see Charness & Sutter, 2012; Kugler, Kausel, & Kocher, 2012). What we find in our experiment is that there are almost as many cases where the joint decisions are improved (patient shifts) as where the joint decisions are worse (impatient shifts) in a joint household decisions setting. This is consistent with for example Hertzberg (2012), who documents that a household could have hyperbolic discounting preferences even if the two spouses are time-consistent if the spouses have misaligned altruistic preferences over each other's outcomes. Clearly, more empirical studies are needed to examine in what types of households these shifts are more likely to occur.

Acknowledgments

Financial support from Sida to the Environmental Economics Unit at the University of Gothenburg and Formas COMMONS is gratefully acknowledged. We have received valuable comments from the associate editor of this journal, two anonymous reviewers, Francisco Alpizar, Håkan J Holm, Johan Stennek, and Conny Wollbrant. We also wish to thank the seminar and conference participants at the University of Gothenburg, the Choice Lab at NHH Norwegian School of Economics, the 6th annual Environment for Development (EfD) conference in Costa Rica, and the 28th annual congress of the European Economic Association and the 67th European meeting of the Econometric Society.

Appendix A

See Table 1A.

Table 1A
Distribution of participating households.

County	Township	Village	Household
Linxia	Yulin	Yufeng	20
		Yaowan	25
	Tuqiao	Dalu	16
		Chongtaiyuan	5
	Nanyuan	Xiaozhai	5
		Jiangjiazhai	10
Jingning	Bali	Guandaocha	20
		Dalv	15
		Jiping	10
	Siqiao	Mougou	25
		Yangchuan	5
	Weirong	Beiguan	5
		Ligou	3

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.joep.2015.11.002.

References

Abdellaoui, M., Haridon, O. L., & Paraschiv, C. (2011). Individual vs collective behavior an experimental investigation of risk and time preferences in couples. *Mimeo*.

Ambrus, A., Greiner, B., & Pathak, P. (2009). *Group versus individual decision-making: is there a shift?* Working paper, Harvard University.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2006). Elicitation using multiple price list formats. *Experimental Economics, 9*, 383–405.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Eliciting risk and time preferences. *Econometrica, 76*(3), 583–618.

Andreoni, J., Kuhn, M. A., & Sprenger, C. (2015). Measuring time preferences: A comparison of experimental methods. *Journal of Economic Behavior and Organization, 116*, 451–464.

Andreoni, J., & Sprenger, C. (2012). Estimating time preferences from convex budgets. *American Economic Review, 102*(7), 3333–3356.

Angerer, S., Lergetporer, P., Glätzle-Rützler, D., & Sutter, M. (2015). How to measure time preferences in children: A comparison of two methods. *Journal of the Economic Science Association*. http://dx.doi.org/10.1007/s40881-015-0016-0.

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

Ashraf, N., Karlan, D., & Yin, W. (2010). Female empowerment: Impact of a commitment savings product in the Philippines. *World Development, 38*(3), 333–344.

Bateman, I., & Munro, A. (2005). An experiment on risk choice amongst households. *The Economic Journal, 115*, 176–189.

Becker, G. (1980). On the long-run steady state in a simple dynamic model of equilibrium with heterogeneous households. *Quarterly Journal of Economics, 95*(2), 375–382.

Browning, M., Bourguignon, F., Chiappori, P.-A., & Lechene, V. (1994). Income and outcomes: A structural model of intra-household allocation. *Journal of Political Economy, 102*(6), 1067–1096.

Carlsson, F., He, H., Martinsson, P., Qin, P., & Sutter, M. (2012). Household decision making in rural China: Using experiments on intertemporal choice to estimate the relative influence of spouses. *Journal of Economic Behavior and Organization, 84*, 525–536.

Carlsson, F., Martinsson, P., Qin, P., & Sutter, M. (2013). The influence of spouses on household decision making under risk: An experiment in rural China. *Experimental Economics, 16*, 383–401.

Cason, T., & Mui, V. (1997). A laboratory study of group polarization in the team dictator game. *The Economic Journal, 107*, 1465–1483.

Chabris, C., Laibson, D., Morris, C., Schuldt, J., & Taubinsky, D. (2008). Individual laboratory-measures discount rates predict field behaviour. *Journal of Risk and Uncertainty, 37*, 237–269.

Charness, G., & Sutter, M. (2012). Groups make better self-interested decisions. *Journal of Economic Perspectives, 26*(3), 157–176.

Coller, M., & Williams, M. B. (1999). Eliciting individual discount rates. *Experimental Economics, 2*, 107–127.

De Mel, S., McKenzie, D., & Woodruff, C. (2009). Are women more credit constrained? Experimental evidence on gender and microenterprise returns. *Applied Economics, 1*(3), 1–32.

de Palma, A., Picard, N., & Ziegelmeyer, A. (2011). Individual and couple decision behavior under risk: Evidence on the dynamics of power balance. *Theory and Decision, 70*, 45–64.

Duflo, E. (2003). Grandmothers and granddaughters: Old age pension and intra-household allocation in South Africa. *World Bank Economic Review, 17*(1), 1–25.

Eliaz, K., Raj, D., & Razin, R. (2006). Choice shifts in groups: A decision-theoretic basis. *American Economic Review, 96*, 1321–1332.

Fiebig, D., Keane, M., Louviere, J., & Wasi, N. (2010). The generalized multinomial logit model: Accounting for scale and coefficient heterogeneity. *Marketing Science, 29*(3), 393–421.

Gine, X., Goldberg, J., Silverman, D., & Yang, D. (2012). *Revising commitments: Field evidence on the adjustment of prior choices*. NBER working paper no. 18065.

Harrison, G. W., Lau, M. I., & Rutström, E. E. (2013). *Identifying time preferences with experiments: Comment*. Working paper. Georgia State University.

Harrison, G. W., Lau, M. I., & Williams, M. B. (2002). Estimating individual discount rates in Denmark: A field experiment. *American Economic Review, 92*(5), 1606–1617.

Hertzberg, A. (2012). *Exponential individuals, hyperbolic households*. Working paper. Columbia Business School.

Kono, H., Matsuda, A., Murooka, T., & Tanaka, T. (2011). *Does marriage work as a savings commitment device? Experimental evidence from Vietnam*. Unpublished working paper.

Kugler, T., Kausel, E. E., & Kocher, M. G. (2012). Are groups more rational than individuals? A review of interactive decision making in groups. *Wiley Interdisciplinary Reviews – Cognitive Science, 3*, 471–482.

Levinger, G., & Schneider, D. J. (1969). Test of the "risk is value" hypothesis. *Journal of Personality and Social Psychology, 11*, 165–169.

Lundberg, S., Pollak, R. A., & Wales, T. J. (1997). Do husbands and wives pool their resources? Evidence from the U.K. child benefit. *Journal of Human Resources, 32*(3), 463–480.

Mani, A. (2010). *Mine, yours or ours? The efficiency of household investment decisions: An experimental approach*. Working paper series 63. University of Warwick.

Moffittt, T., Arsenault, L., Belsky, D., Dickson, N., Hancox, R., Harrington, H., ... Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *PNAS, 108*(7), 2683–2698.

Moscovici, S., & Zavalloni, M. (1969). The group as a polarizer of attitudes. *Journal of Personality and Social Psychology, 12*, 125–135.

Namoro, S., & Roushdy, R. (2008). Intra-household resource allocation in Egypt women empowerment and investment in children. *Middle East Development Journal*, 31–47 [demo issue].

Phipps, S., & Burton, P. (1998). What's mine is yours? The influence of male and female incomes on patterns of household expenditure. *Economica, 65*(260), 599–613.

Ramsey, F. (1928). A mathematical theory of saving. *The Economic Journal, 38*, 543–559.

Robinson, J. (2012). Limited insurance within the household: Evidence from a field experiment in Kenya. *American Economic Journal: Applied Economics, 4*(4), 140–164.

Shapiro, J. (2010). *Discounting for you, me and we: Time preference in groups and pairs*. Working paper. Yale Department of Economics.

Stoner, J. A. F. (1968). Risky and cautious shifts in group decisions: The influence of widely held values. *Journal of Experimental Social Psychology, 4*(4), 442–459.

Sunstein, C. R. (2000). Deliberative trouble? Why groups go to extremes. *Yale Law Journal, 110*, 71–119.

Sunstein, C. R. (2002). The law of group polarization. *The Journal of Political Philosophy, 10*, 175–195.

Sutter, M., Kocher, M., Glätzle-Rüetzler, D., & Trautman, S. (2013). Impatience and uncertainty: Experimental decisions predict adolescents field behaviour. *American Economic Review, 103*, 510–531.

Tanaka, T., Camerer, C., & Nguyen, Q. (2010). Risk and time preferences: Experimental and household survey data from Vietnam. *American Economic Review, 100*(1), 557–571.

Thomas, D. (1990). Intra-household resource allocation: An inferential approach. *Journal of Human Resources, 25*(4), 635–664.