WORKING PAPERS

N° 17-798

April 2017

"Asset pricing and risk-sharing in a complete market: An experimental investigation"

Bruno Biais (HEC & TSE), Thomas Mariotti (TSE),
Sophie Moinas (TSE) and Sébastien Pouget (TSE)

Toulouse School of Economics

Asset pricing and risk–sharing in a complete market: An experimental investigation

Bruno Biais (HEC & TSE), Thomas Mariotti (TSE), Sophie Moinas (TSE) and Sébastien Pouget (TSE)

April 2017

Many thanks for helpful comments to seminar participants at City University, Zurich University, the Max Planck Institut, the 2014 Conference of the Financial Engineering & Banking Society in Guilford, the 2014 Summer Symposium in Financial Markets at Gerzensee, the 2015 Barcelona GSE Summer Forum, the 2015 Conference of the Society for Experimental Finance, the 2015 Workshop for the Promotion of Experimental Validation of the Theory of Asset Pricing, and the Conference on the Experimental and Behavioral Aspects of Financial Markets at Chapman University in 2016, especially Elena Asparouhova, Peter Bossaerts, George Constantinides, John Duffy, Frank Heinemann, Debrah Meloso, Charles Plott, Bill Zame, Stefan Zeisberger and Martin Weber. Biais acknowledges financial support from the European Research Council under the European Community's Seventh Framework Program FP7/2007-2013 grant agreement N$\frac{1}{2}$ 295484. Moinas acknowledges financial support from the Agence Nationale de la Recherche (ANR-16-CE26-0008-01, project PIMS). Part of this research was carried out while Sophie Moinas was Judith C. and William G. Bollinger Visiting Professor at the Wharton School, University of Pennsylvania, and benefited from the support from the CNRS.

Abstract

We design an experiment that closely emulates and tests the standard model of complete competitive markets, without imposing parametric restrictions on preferences. Consistent with theory, aggregated elicited supply and demand curves cross at the expected dividend when there is no aggregate risk, and at a lower price when there is aggregate risk. In contradiction with theory, individual participants frequently make choices that violate first order stochastic dominance. We propose a random choice model which reconciles the above mentioned findings and is also consistent with additional features of the data, such as, e.g., large mistakes being less frequent than smaller ones.

1 Introduction

Since the seminal works of Debreu (1959) and Arrow (1964), the general equilibrium theory of asset pricing and risk sharing in perfect and complete markets has offered an elegant framework and sharp implications: agents should share risk perfectly, which, in turn, implies that only aggregate risk should be priced (Borch, 1962). Unfortunately, these implications are rejected by field data. Is it because human cognition and preferences do not conform to the standard rational choice paradigm? Or is it because, in practice, markets are imperfect and incomplete? These two potential explanations have very different implications for further research. The former calls for new models of human decision–making, whereas the latter emphasizes the need to model market imperfections. While it is difficult to disentangle these two explanations with field data, in the lab it is possible to do so. In a controlled experimental setting one can make sure the market is perfect and complete. Any deviation from the implications of rational choice and competitive equilibrium can then be attributed to human cognition.

To study this issue, we conduct an experimental analysis of the simplest possible setting in which the basic tenets of the theory can be tested. The state of the world ω can take only two values (u and d) and there are two non-redundant assets (a stock and a bond), so the market is complete. At the beginning of each of 8 trading rounds, participants start with heterogeneous endowments (stocks, bonds and other state–contingent income). Individual demand and supply functions (specifying how many shares the participant wants to buy or sell at different prices) are elicited. Participants are asked to choose the quantity they want to buy or sell at all prices on a grid. For half the rounds of the experiment, the market price is randomly drawn and thus participants have no opportunity to manipulate the price.[1] During the other rounds, in contrast, participants are informed that the price is set to minimize the difference between aggregate supply and aggregate demand. Empirically, we observe similar behaviour within these two different price–setting mechanisms. This suggests that our experimental design is able to generate a situation in which agents behave competitively, as in the standard competitive equilibrium model.[2]

The simplicity of our experimental setting enables us to pin down precisely the implications of rational choice for individual behaviour and market outcomes for a large class of preferences.[3]

- At the individual level, our experimental design enables us to test the hypothesis that par-

[1]This is in the same spirit as Becker, DeGroot, Marschak (1964). They i) ask a seller to quote a reservation price, ii) randomly draw a buying price, and iii) let the seller trade 1 unit at the buying price if the latter is above the reservation selling price. While in that mechanism trades are only for one unit, our random price mechanism can be seen as an extension of the Becker, DeGroot, Marschak (1964) to general supply or demand curves.

[2]Our market environment is similar to the call market in Smith, Williams, Bratton, and Vannoni (1982), McCabe, Rassenti and Smith (1993) or Plott and Pogolreslskyi (2015). The main differences are that i) they let participants submit limit orders, while we elicit demand functions, and ii) our design enables us to test if participants are competitive. Bohm, Linden and Sonnegard (1997) compare the performance of a market mechanism with that of a Becker, DeGroot, Marschak (1964) mechanism in eliciting the private value of participants for one unit of a commodity.

[3]That class includes risk averse, risk loving or risk neutral expected utility, as well as rank dependent expected utility.

ticipants are rational. We show that some actions are first order stochastically dominated. Rational choice, therefore, implies they should not be observed in the lab.

- At the market level, our experimental design enables us to test the hypothesis that participants are risk averse and share risk efficiently. We consider two treatments. In the first treatment there is no aggregate risk. So participants can perfectly hedge their risky endowments and the price of the stock should be equal to the expected dividend. In the second treatment there is aggregate risk. So, with risk-averse agents, the stock price should be lower than the expected dividend.[4]

We ran the experiment with 141 students in Toulouse University. There were 8 cohorts. Each participated in 8 replications of the experimental market. Participants' compensation was a linear function of their gains in two randomly drawn replications (with an average of 85.84 euros per participant).

Aggregate outcomes, in our experimental market, are consistent with the implications of rational choice. When there is no aggregate risk, market clearing prices are close to expected dividends, and average trading volume is close to that requested by perfect risk-sharing. With aggregate risk, there is a risk premium, consistent with a relative risk aversion coefficient close

[4] The two treatments alternate from one round to the next. While market participants know their own endowment, they have no information about the endowments of the others or the alternance between the two treatments. This is in line with competitive Walrasian equilibrium, where agents only need to contingent decisions on the price, and do not engage in strategic interactions with one another.

to .5 for the representative agent.

Observed individual behaviour, however, markedly differs from that predicted by rational choice. Around 30% of observed actions are first order stochastically dominated. The frequency of dominated actions, however, declines with experience, from approximately 34% in the first replication to 24% in the last replication. Probit regressions, in which the dependent variable is the indicator that the action is dominated, confirm that there is significant learning. They also show that, when an agent opted for a dominated action in a given round, he/she is likely to continue to do so at the next round, except if the action led to a loss-making trade. The latter suggests that agents learn from costly mistakes.

To reconcile aggregate and individual findings, we propose a model of bounded rationality. In line with random choice models (Luce, 1959, McKelvey and Palfrey, 1995, 1998), we posit that participants play randomly but put more weight on actions delivering larger value. We show that the distribution of individual actions predicted by this model is consistent with that observed in the lab: As mentioned above, first order stochastically dominated actions are less frequent than undominated ones, which is in line with the predictions of the random choice model. Furthermore, in our simple experimental setting we are able to identify "large" and "small" mistakes, such that, while both are dominated by rational actions, large mistakes are second order stochastically dominated by small mistakes. In line with the predictions of the random choice model for risk averse agents, we find that large mistakes are less frequent than small ones. We also show (analytically) that the actions generated by the random choice model aggregate to supply and demand curves which, like those we observe in the lab, generate market pricing and average trading consistent with rational choice.

Our work is directly in line with the insightful analyses of Bossaerts and Plott (2004) and Bossaerts, Plott and Zame (2007).[5] Like them we study the consequences of risk aversion in complete experimental markets and find prices involve risk–premia. Differences between the present paper and theirs include the following:

- The trading protocol we consider is different from Bossaerts and Plott (2004) and Bossaerts, Plott and Zame (2007). They consider a continuous double–auction, in which traders can dynamically post and hit quotes. This is similar to the workings of electronic limit order books in financial markets during the day, and this is also in line with the seminal experimental studies of Smith, Suchanek and Williams (1988) and Plott and Sunder (1982, 1988). In contrast, we consider a competitive market environment in which participants submit supply and demand functions which are aggregated and crossed to determine the equilibrium price.

- While Bossaerts and Plott (2004) and Bossaerts, Plott and Zame (2007) consider a three–state model, we consider a two–state model. The advantage of a three–state complete market is that there are two risky assets, so that portfolios of risky assets can be analyzed. This enables Bossaerts and Plott (2004) and Bossaerts, Plott and Zame (2007) to test the hypothesis that agents hold the market portfolio. The problem with that design,

[5] While our theoretical framework is static, Bossaerts, Meloso and Zame (2013) and Crockett and Duffy (2013) offer interesting analyses of equilibrium market dynamics.

however, is that participants must trade several risky assets. Such a market is difficult to organize and cognitively challenging for participants. Our design is simpler since we only have two assets: the risk–free bond and the risky stock. Thus agents need to trade in only one market, where the risky asset is exchanged for the risk–free bond. Moreover, as mentioned above, the simplicity of our setting enables us to obtain precise implications for a large class of preferences.

- While Bossaerts and Plott (2004) and Bossaerts, Plott and Zame (2007) consider markets in which there is always aggregate risk, we also consider a treatment in which there is no aggregate risk, while individual endowments are risky. This enables us to test a very basic implication of theory that does not arise in Bossaerts and Plott (2004) and Bossaerts, Plott and Zame (2007): In complete markets without aggregate risk, i) agents should bear no risk and ii) this should give rise to risk-neutral pricing.

Our finding that aggregate behaviour is well behaved, while individual behaviour is noisy, echoes that of Bossaerts, Plott and Zame (2007). Their CAPM+ϵ model assumes that an individual trader's demand function is the sum of a mean-variance optimal demand and a noise term. Our random choice model allows for a large class of preferences (expected utility with any curvature, rank dependent expected utility). Moroever, in our simple experimental design, our random choice model imposes structure on the distribution of individual actions, which we confront to the experimental data.

The present paper is also related to the experimental literature studying the consequences of risk–aversion for economic and financial decisions. Holt and Laury (2002) observe lottery choices consistent with risk-aversion. Experimental findings are also suggestive of risk aversion for private value auctions (Goeree, Holt and Palfrey, 2002), asymmetric matching penny games (Goeree, Holt and Palfrey, 2003), and one–shot matrix games (Goeree and Holt, 2004).[6] In Bossaerts and Plott (2004) market pricing is consistent with risk aversion, since assets with large betas have greater expected returns than assets with low beta. Our experimental finding that with no aggregate risk prices are equal to expected dividend, while with aggregate risk prices are below aggregate dividend, offers additional evidence consistent with risk aversion.

Section 2 presents our experimental design. Section 3 presents testable implications of rational choice in our complete market experimental setting. Section 4 confronts these implications with experimental evidence. Section 5 presents a random choice model and shows it rationalises experimental findings. Section 6 concludes.

2 Experimental Design

The purpose of our experimental design is to closely emulate and test the standard model of complete competitive markets, without imposing parametric restrictions on preferences.

[6]The majority of participants in Holt and Laury (2002) have estimated relative risk–aversion between .15 and .97. Estimates of the same order of magnitude have been found by Goeree et al (2002, 2003) and Goeree and Holt (2004).

Assets: There are two equally likely states of nature, $\omega = u, d$, and two non-redundant assets, a bond and a stock. Hence markets are complete. One unit of the bond pays 1 in each state of nature. One share of the stock pays 120 in state u and 0 in state d. To implement complete markets in our two-state experimental setting, it is enough to consider one market, in which the stock is traded against the bond; that is, we take the bond as the numï¿½raire.[7] We denote by S the price of the stock relative to the bond.

Although considering only two assets precludes analyzing portfolios of risky assets, it sidesteps the difficulties associated with simultaneous trading in several experimental markets that have been emphasized by Bossaerts and Plott (2004) or Bossaerts, Plott, and Zame (2007). Moreover, having only one market in which the stock can be traded against the bond simplifies the cognitive task faced by participants. If, by contrast, we had considered two markets in which the stock and the bond could be traded against a perishable unit of account, participants would have faced a more difficult task: they would have had to factor their trades on the two markets into their budgets constraints, and to compare prices on the two markets.

Endowments: Participants' endowments come from their initial holdings of bond and stock, and from state-contingent additional income. We will consider two treatments. In Treatment 1, there is no aggregate risk: the sum of individual endowments is constant across states of nature. In Treatment 2, there is aggregate risk: the sum of individual endowments

[7] The mapping between the complete-market environment and our experimental setting is explicited in the online appendix.

is larger in state u than in state d. We chose this design because, as explained in the next subsection, it enables us to test sharp predictions of the theory.

For simplicity, participants can receive only three types of endowments:

1. Type 1 participants initially receive 5 shares of the stock and no bond. Their additional income is 0 in state u and 360 in state d. Thus their endowments are 600 in state u and 360 in state d.

2. Type 2 participants initially receive no stock and 310 bonds. Their additional income is 0 in state u and 240 in state d. Thus their endowments are 310 in state u and 550 in state d.[8]

3. Type 3 participants initially receive no stock and 310 bonds. They have no additional income. Thus their endowment is 310 in both states u and d.

In Treatment 1, if the number of participants is even, there are only Type 1 and Type 2 participants, in equal numbers; whereas if the number of participants is odd, there is one additional Type 3 participant. This treatment corresponds to a situation with no aggregate risk.

In Treatment 2, if the number of participants is even, there are Type 1 and Type 3 participants, in equal numbers; whereas if the number of participants is odd, there is one additional Type 3 participant. This treatment corresponds to a situation with aggregate risk.

[8] The endowments for Type 1 and Type 2 participants have been chosen so that both are fully hedged if they trade 2 shares, as discussed in the next section.

Participants do not know that there are two treatments. At each replication of the experiment, they are only informed about their own endowments and the distribution of the dividend. This is in line with the standard competitive equilibrium model, in which agents only rely on prices and rational expectations about the distribution of future values, as emphasized in Bossaerts and Plott (2004.)

Supply, demand and prices: Historically, experimental economics relied on the continuous double auction to study market equilibrium in the lab. As noted by Bossaerts (2008):

> "The very quest for an institution capable of generating the (Walrasian) competitive equilibrium ended up defining experimental economics... What established economics as an experimental science was proof that the continuous auction did the job (Smith 1962, Plott and Smith 1978)."

Yet, Bossaerts (2008) also notes:

> "The continuous double auction ... defies formal analysis in game theoretic terms."

To avoid this pitfall, we design our experimental market to closely emulate the standard competitive equilibrium model in which each participant states how much he/she wants to sell or buy at each price, and the price is set to clear the market.

As explained below, if all agents are competitive and risk averse, in equilibrium, Type 1 agents never buy the stock, while Type 2 and Type 3 agents never sell it. To simplify the task of participants in the experiment, we therefore restrict Type 1 participants to supply, and Type 2 and 3 participants to demand nonnegative quantities of the stock. A further simplification is that participants are restricted to trade quantities no greater than 4.[9]

Participants are asked which quantity they want to sell or buy at each point of a price grid ranging from 52 to 62 with unit increments. Whereas the price grid is discrete, at each price S, we allow participants to supply or demand any quantity of the stock in the interval $[0, 4]$. Thus participants can fine-tune their supply or demand, which enables them to equate marginal utility and price when they select trades in $[0, 4]$.

Once supply and demand curves have been elicited, it would seem natural to simply aggregate and cross them to determine the market clearing price. Two difficulties arise, however. The first difficulty is that discreteness of the price grid could prevent market clearing. The second difficulty is that participants could behave strategically and try to manipulate the price. To address these issues, we consider the two following pricing mechanisms:

1. In the *call* mechanism, the price is set to minimize the gap between supply and demand.[10] This is in line with the call market mechanism in Smith, Williams, Bratton, and Vannoni (1982), McCabe, Rassenti and Smith (1993) or Plott and Pogolreslskyi (2015). The main difference is that in these papers participants submit schedules of limit orders, while in

[9] These restrictions, in particular, forbid short sales, which would anyhow not arise in equilibrium with risk averse agents.

[10] In the very few instances in which several prices minimized the gap between supply and demand, we drew one of them randomly.

ours they submit demand functions. While limit order schedules give rise to monotonic demand curves, the demand curves we elicit can be non–monotonic.[11] As mentioned above, because we consider a discrete pricing grid, at the price minimizing the gap between supply and demand, there is typically rationing. To make sure this consideration does not affect the choices of the participants, we supply or demand the quantity of stock needed to clear the market. Thus participants' orders are fully executed. In practice the additional supply or demand injected in the market by the experimenter averaged to 6.52% of the minimum between supply and demand during the call sessions. This way to handle potential mismatch between supply and demand ensures that participants behave as predicted by the competitive model.[12]

2. In the *random* mechanism, the price is drawn from the price grid, each price being equally likely. As in the call mechanism, participants' orders are fully executed. This random mechanism is in line with the Becker, Degroot and Marshack (1964) mechanism. While in Becker, Degroot and Marshack (1964) each participant trades one or zero unit, in our mechanism, participants can trade any amount between 0 and 4.

At the beginning of each replication of the experiment, all participants are told which mechanism (random or call) will be used to set the price. Strategic considerations could affect

[11] In the the standard competitive equilibrium model too, demand curves may be non monotonic, due to the tradeoff between substitution and wealth effects.

[12] If instead we had used a rationing scheme, it could have affected behaviour. For example, participants anticipating prorata rationing may be tempted to inflate their demand.

participants' behaviour when they know their orders will be processed in the call mechanism, but not when they know the random mechanism will be used to set the price. Therefore, comparing the two enables us to test whether strategic considerations significantly affected participants' behaviour. If they do not, then the call mechanism provides a good approximation of a competitive market.

3 Testable Implications of Rational Choice

In this section we outline the testable implications arising in our simple experimental setting.

3.1 Individual Behaviour

Denote by $W^i(u, S)$ and $W^i(d, S)$ the final wealth of agent i, for a given price S, in states u and d respectively. For each price, $W^i(u, S)$ and $W^i(d, S)$ are pinned down by the quantity chosen by the agent at that price. Because the two states are equally likely, the expectation of the final wealth of i is

$$\mu_i(S) = \frac{W^i(u, S) + W^i(d, S)}{2},$$

while its standard deviation is

$$\sigma_i(S) = \frac{|W^i(u, S) - W^i(d, S)|}{2}.$$

When $W^i(u, S) > W^i(d, S)$,

$$W^i(u, S) = \mu_i(S) + \sigma_i(S),\ W^i(d, S) = \mu_i(S) - \sigma_i(S).$$

When $W^i(u,S) < W^i(d,S)$,

$$W^i(u,S) = \mu_i(S) - \sigma_i(S),\ W^i(d,S) = \mu_i(S) + \sigma_i(S).$$

Putting the two cases together, because the two states are equiprobable, the lottery faced by agent i is equivalent to the lottery $(\mu_i(S) + \sigma_i(S), \frac{1}{2}; \mu_i(S) - \sigma_i(S), \frac{1}{2})$ (irrespective of whether $W^i(u,S) > W^i(d,S)$ or $W^i(u,S) < W^i(d,S)$). Hence, in our simple empirical design, if $\mu_i'(S) > \mu_i(S)$, then $(\mu_i'(S) + \sigma_i(S), \frac{1}{2}; \mu_i'(S) - \sigma_i(S), \frac{1}{2})$ first-order stochastically dominates $(\mu_i(S) + \sigma_i(S), \frac{1}{2}; \mu_i(S) - \sigma_i(S), \frac{1}{2})$, as illustrated in Figure 1.

The class of preferences for which agents exhibit preference for first order stochastic dominant shifts is very large. It includes, for instance, expected utility with the only requirement that utility increase in wealth, Frechet-differentiable utility with increasing local utility functions (see Machina, 1982), or rank–dependent expected utility (Quiggin, 1982). We hereafter refer to agents with preferences in this class as FOSD agents. As explained below, our symmetric two state setup generates clear predictions for FOSD agents.

Our experimental design also enables us to test the implications of risk aversion. An agent is risk averse if she prefers any lottery than a mean–preserving spread of that lottery. Symmetrically, an agent is risk loving if she prefers a mean preserving spread of any lottery than that lottery. Following Segal and Spivak (1990), we will say that an agent exhibits second-order risk aversion whenever the risk premium she requires for a small bet is proportional to the square of the size of that bet. Intuitively, this means the agent is almost risk neutral for small bets. By contrast, an agent exhibits first-order risk aversion whenever the risk premium is proportional to the size of the bet. In that case the agent may be willing to pay a premium to buy insurance against a small risk. We spell out below the consequences of first- and second-order risk aversion in our setting.

Type 1 If a Type 1 participant sells Q shares of the stock at price S, her final wealth in states u and d is

$$W^i(u,S) = 600 + (S-120) \times Q, \tag{1}$$

$$W^i(d,S) = 360 + S \times Q. \tag{2}$$

(1) and (2) imply that the mean and standard deviation of i's final wealth are $\mu^i(S) = (S-60) \times Q + 480$ and $\sigma^i(S) = 60 \times |Q-2|$, respectively. Therefore, two possible trades $2+x$ and $2-x$ lead to equally volatile final wealth, for all $x \in (0,2]$. However, if the price is strictly lower than the expected dividend, i.e., $S < 60$, selling $2+x$ shares leads to a lower expected final wealth than selling $2-x$ shares, while for $S > 60$, the converse holds. Hence, trading $2+x$ shares is first order stochastically dominated by trading $2-x$ shares when $S < 60$, while the converse holds when $S > 60$.[13]

[13] The cdfs for the lotteries generated by selling $2+x$ shares and $2-x$ are identical to those represented in Figure 1, with $W^i(u,S)|_{2+x} = \mu_i - \sigma_i$, $W^i(d,S)|_{2-x} = \mu_i' - \sigma_i$, $W^i(d,S)|_{2+x} = \mu_i + \sigma_i$ and $W^i(u,S)|_{2-x} = \mu_i' + \sigma_i$. Note that $W^i(u,S)|_{2-x} \geq W^i(d,S)|_{2-x}$, while $W^i(u,S)|_{2+x} \leq W^i(d,S)|_{2+x}$. Yet, the first step in the cdf is the same for both lotteries (going from 0 to .5). This property of the cdf's, which delivers the first-order dominance result, obtains because the two states are equiprobable. For arbitary probabilities, first order stochastic dominance does not always obtain, as illustrated in the figure in the online appendix. Hence equiprobability is instrumental in yielding sharp testable implications.

Moreover, if an agent is risk–loving or risk neutral, it is optimal for her to supply 4 shares at any price $S > 60$ (and 0 shares at any price $S < 60$). To see that selling 4 shares is preferred to selling $4 - x$ shares for all $x \in (0, 4]$ when $S > 60$, compare the lottery generated by selling $4 - x$ to that generated by selling 4 shares. The latter is equal to the former, plus a positive constant ($x(S - 60)$), plus a mean preserving spread ($(-60x, \frac{1}{2}; 60x, \frac{1}{2})$). Indeed:

$$W^i(u, S)|_{Q=4} - W^i(u, S)|_{Q=4-x} = [600 + (S - 120) \times 4] - [600 + (S - 120) \times (4 - x)] = x(S-60) - 60x$$

and

$$W^i(d, S)|_{Q=4} - W^i(d, S)|_{Q=4-x} = [360 + S \times 4] - [360 + S \times (4 - x)] = x(S - 60) + 60x.$$

What if $S = 60$? Type 1 participants are perfectly hedged if they sell 2 shares. Any other trade generates a mean preserving spread around that safe outcome. Hence, when $S = 60$, a risk-averse Type 1 participant finds it optimal to sell exactly two shares. Symmetrically, a risk loving participant prefers to take as much risk as possible at that price, which can done, equivalently, by selling 0 or 4 shares. Risk neutral agents are indifferent between any quantity in $[0, 4]$. The remarks above yield the following implication.

Implication 1

- *If a Type 1 participant is an FOSD agent, she does not sell more than* 2 *shares at* $S < 60$, *nor sell less than* 2 *shares at* $S > 60$. *If she is risk–loving or risk–neutral, she sells 0 shares at* $S < 60$, *and* 4 *shares at* $S > 60$.

- *If a Type 1 participant is risk averse (resp. risk–loving), she supplies* 2 *shares (resp. 0 or 4 shares) at* $S = 60$. *If she is risk-neutral she is indifferent between all trades at* $S = 60$.

To illustrate this implication, Figure 2, Panel A, depicts the choices of Type 1 agents in the price–quantity plane. The relevant prices are between 52 and 62, the relevant quantities are between 0 and 4, and there are four quadrants determined by the horizontal line $Q = 2$ and the vertical line $S = 60$. If follows from Implication 1 that the North-West and South-East quadrants are dominated for an FOSD–agent. Moreover, a risk-averse Type 1 agent's supply function goes through the point $(60, 2)$.

Type 2 If a Type 2 participant buys Q shares of the stock at price S, her final wealth in states u and d is

$$W^i(u, S) = 310 + (120 - S) \times Q, \tag{3}$$

$$W^i(d, S) = 550 - S \times Q. \tag{4}$$

Following similar steps as for Type 1, we obtain the following implication.

Implication 2 • *If a Type 2 participant is an FOSD–agent, she does not buy less than* 2 *shares at* $S < 60$, *or more than* 2 *shares at* $S > 60$. *If this participant is risk–loving, she buys 4 shares at* $S < 60$, *and* 0 *shares at* $S > 60$.

- *If a Type 2 participant is risk averse (resp. risk–loving), she buys* 2 *shares (resp. 0 or 4 shares) at* $S = 60$.

This implication is illustrated in Panel B of Figure 2, which is the mirror image of Panel A.

Type 3 In contrast with Type 1 and Type 2 participants, Type 3 participants start with a riskless initial endowment. Thus, trading increases their risk–exposure. More precisely, if a Type 3 participant buys Q shares of the stock at price S, her final wealth is $W^i(u, S) = 310 + (120 - S) \times Q$, and $W^i(d, S) = 310 - S \times Q$, in states u and d respectively. This can be rewritten as

$$W^i(u, S) = 310 + (60 - S) \times Q + 60 \times Q, \quad (5)$$

$$W^i(d, S) = 310 + (60 - S) \times Q - 60 \times Q. \quad (6)$$

(5) and (6) show that, relative to the safe endowment of Type 3, trading Q adds a constant ($(60 - S) \times Q$) plus a mean preserving spread ($(-60Q, \frac{1}{2}; 60Q, \frac{1}{2})$). Therefore, we obtain the following implication, illustrated in Figure 2, Panel C.

Implication 3 *When $S > 60$, a risk averse Type 3 agent strictly prefers not to buy any share. When $S < 60$, a risk loving Type 3 agent strictly prefers to buy 4 shares*

The restrictions imposed by theory for Type 3 are weaker than for Type 1 and Type 2, since there are no robust theoretical predictions for risk averse participants when $S \leq 60$, or risk loving participants when $S \geq 60$.

3.2 Equilibrium Outcomes

Treatment 1 In Treatment 1, there are N Type 1 participants, N Type 2 participants, and possibly one Type 3 participant. By Implications 1, 2 and 3, if all participants are risk averse, at $S = 60$ Type 1 should sell two shares, Type 2 buy shares two and Type 3 trade 0 share. Thus, if all participants are risk averse, $S = 60$ is an equilibrium, with trading volume $2N$.

Is it the only equilibrium? At price $S > 60$, by Implication 1, supply is at least $2N$. Thus, by Implications 2 and 3, $S > 60$ can be an equilibrium only if at that price all Type 2 participants buy 2 shares. Now, if Type 2 participants exhibit finite second-order risk aversion, for any ε small enough, buying 2 shares is dominated by trading $2-\varepsilon$ shares (see Segal and Spivak, 1990). Indeed, relative to the safe lottery obtained when buying two shares, trading $2 - \varepsilon$ shares adds a small gamble, with positive expected profit, a gamble that finitely second-order risk averse participants will accept. Consequently, there cannot exist an equilibrium with price strictly above 60 (and symmetric arguments rule out an equilibrium with price strictly below 60). The above remarks yield our fourth implication.

Implication 4 *In Treatment 1, if all participants are finitely risk-averse, there exists an equilibrium such that $S = 60$, Type 1 participants sell 2 shares, Type 2 participants buy 2 shares, and Type 3 participants do not trade. If participants are finitely second-order risk averse, the equilibrium is unique.*

The intuition is the following. In Treatment 1, aggregate wealth is constant across states. Thus, in equilibrium risk averse agents perfectly hedge their risk exposure and the price does not reflect any risk–premium, it is simply equal to the expected dividend.[14] With expected

[14] While the equilibrium analysis underlying Implication 4 is conducted in a perfect market, with continuous prices, it also applies in our experimental setting with a discrete price grid, because the equilibrium price $S = 60$

utility maximization, this is an instance of the mutuality principle (Borch, 1962). In our simple experimental design, it obtains as soon as agents exhibit second-order risk aversion.

Treatment 2 In Treatment 2, there are only Type 1 and Type 3 participants. By Implication 1, if they exhibit FOSD preference, Type 1 agents sell at least 2 shares at any price above 60. By Implication 3, if Type 3 participants are risk averse they don't buy any share at any price above 60. If participants are FOSD and risk averse, their preferences are monotone and convex in state–contingent consumption, hence there exists an equilibrium. This yields the following implication.

Implication 5 *In Treatment 2, if participants are FOSD and risk averse, there exists an equilibrium and the equilibrium price is lower than or equal to 60.*

As FOSD risk averse Type 1 participants sell less than 2 shares at any price lower than 60, equations (1) and (2) imply that their equilibrium consumption is larger in state u than in state d. Moreover, FOSD risk averse Type 3 participants prefer to trade non negative quantities at any price lower than 60. Hence equations (5) and (6) imply that their equilibrium consumption is larger in state u than in state d. Thus, individual equilibrium consumptions comove with aggregate endowment. This extends, in our simple context, the mutuality principle (Borch, 1962) slightly beyond the expected utility case.

The gap between the expected dividend of the stock (60) and its equilibrium price (S) is the risk premium requested by risk averse agents to bear aggregate risk. Uniqueness of equilibrium

belongs to the pricing grid.

is not guaranteed in general, but if participants are risk averse expected utility maximizers with relative risk aversion $-wv^{i\prime\prime}(w)/v^{i\prime}(w) \leq 1, \forall w$ and i, then substitution effects dominate income effects and the equilibrium is unique.[15]

4 Is the Evidence Consistent with Rational Choice and Competitive Behaviour?

We ran the experiment with 141 participants. All of them were students enrolled in the first year of Toulouse University's master in finance.[16] There were 8 cohorts. All participated in 8 replications of the experiment, lasting, overall, one hour and a half (but for one of the cohorts an operational problem prevented us from collecting the data for the last replication). Treatment 1 (corresponding to no aggregate risk) and Treatment 2 (corresponding to aggregate risk alternated during the 8 replications. Also, for half the cohorts, the first four replications involved random prices, while the last four replications involved a call auction. For the other cohorts, the first four replications involved a call auction, while the last four replications involved random prices. The details of the experimental sessions are documented in the table in the

[15]The condition ensures that the aggregate excess demand function for state-contingent wealth satisfies the gross substitute property, see Varian (1985) or Mas-Colell, Whinston, and Green (1995, Example 17.F.2 and Proposition 17.F.3, pp. 612–613).

[16]These students come from different backgrounds. The majority studied management, and had very little exposure to micro-economic or finance theory. Some studied economics, and had greater exposure to microeconomics, but not to finance theory. Others come from engineering or maths.

online appendix.

For each group, two of the eight rounds were randomly drawn at the end of the experiment, one from sessions with random pricing and the other from sessions with a call auction. As announced at the start of the experiment, participants received the sum of their final wealths in these two rounds, divided by ten. The average individual payment was € 85.84, the minimum was € 5 and the maximum € 120. We ran an anonymous survey among the participants, asking them their monthly budget (including all expenses: housing, food, leisure, etc.). The average was € 646. Thus, the amount participants could make in the lab, and its variability, were significant relative to their budget in the field.[17] Overall, 68 participants were assigned the role of Type 1 sellers, 68 participants were switching from Type 2 to Type 3 buyers depending on the treatment, and 5 participants were assigned the role of Type 3 buyers.

4.1 Aggregate Outcomes

In this section we study aggregate supply and demand and market pricing. First, we report the evidence on whether supply and demand reflect strategic behaviour. In half the replications of the experiment, the price at which participants traded was randomly drawn. For the other replications the price was set to minimize the gap between supply and demand. To test

[17] Bossaerts and Plott (2004) compare participants' behaviour in asset market in the US and in Bulgaria. The experimental design was the same in the two countries, but monetary incentives were much stronger in Bulgaria. Bossaerts and Plott (2004) find qualitatively similar results. In both countries there is a risk premium. The main difference is that the risk premium is larger in Bulgaria.

the null hypothesis of competitive behaviour, we ran Wilcoxon rank–sum tests comparing the distribution of quantities offered or demanded at each price in the two mechanisms.[18] As illustrated in Figure 3, the p–value was above 10% in all cases but 4 (out of 44). This does not suggest rejection of the null hypothesis of competitive behaviour. Accordingly, when analyzing aggregate outcomes we hereafter pool the observations from the two pricing mechanisms (except in Figure 5, panel C and D, below, which documents further the similarity in participants' behaviour between call and random price treatments.)

Figure 4 plots the average across cohorts of the price minimizing the distance between aggregate supply and aggregate demand.[19] For each of the two treatments (Treatment 1 without aggregate risk, and Treatment 2 with aggregate risk), each of the eight cohorts played four replications. The solid line in Figure 4 plots the price (averaged across 8 cohorts) for the first, second, third and fourth Treatment 1 replications. The dashed line plots the average prices

[18] At each price level, we first computed for each individual the average quantity in four cases: random mechanism/aggregate risk, random mechanism/no aggregate risk, call/aggregate risk and call/no aggregate risk. So, since there are 8 replications, we computed an average quantity across 2 replications for each subject. Then, for each of the 11 prices level, we ran separate Wilcoxon rank–sum tests to examine if the quantity traded differs in the random mechanism and the call market. We ran these tests separately for buyers and sellers, and for Treatment 1 and Treatment 2. So we computed 4 times 11 statistics. For each, the null hypothesis is that the distribution across subjects of quantities for the random price mechanisms is equal to its counterpart for the call mechanism.

[19] For the replications relying on a call market, but not for those with random pricing, this was the actual price in the market.

for Treatment 2. The figure shows that, during the first two replications, average prices are similar in the two treatments (close to 58). During the last two replications, in Treatment 1, the average market price increases and converges to the expected dividend (60) as predicted by theory (see Implication 4 above). In contrast in Treatment 2, the average market price decreases and converges close to 57. The difference between the expected dividend (60) and the average price (57) in Treatment 2 can be interpreted in terms of risk premium, and is consistent with the prediction from theory for risk averse participants (see Implication 5 above). With a price of 57 and an expected final value of 60, the risk–premium is 5.26%. For a representative investor endowed with the aggregate wealth and with power utility function, this corresponds to a relative risk aversion coefficient a bit lower than .5, in line with evidence from other experiments infering risk–aversion from participants' behaviour.[20]

To shed light on aggregate demand and supply in our experimental market, we aggregate all the supply and demand curves from all Type 1 and Type 2 participants for Treatment 1, and from all Type 1 and Type 3 participants for Treatment 2. It is legitimate to do so because, in our competitive setting, individual demands and supplies are independent of the number and characteristics of the other market participants.

Figure 5, Panel A, depicts aggregate demand and supply in Treatment 1 (divided by the number of participants–replications, to facilitate interpretations). Demand is approximately decreasing and supply approximately increasing. Quite strikingly, the price minimizing the

[20] See Holt and Laury (2002) for lottery choices, Goeree et al (2002) for private value auctions, Goeree et al (2003) for asymmetric matching penny games, and Goeree and Holt (2004) for one–shot matrix games.

gap between supply and demand is 60, which is the equilibrium price predicted by theory (see Implication 4 above). At that price, trading volume per participant is very close to the two shares predicted by theory for risk averse agents (see Implication 4 above).

Figure 5, Panel B, depicts aggregate demand and supply in Treatment 2. As predicted by theory, supply is not very different from its counterpart in Panel A. In contrast, as predicted by theory for risk averse agents, demand is lower than in Panel A. Correspondingly, the price minimizing the gap between supply and demand is lower than the expected dividend, as predicted by theory for risk averse agents (see Implication 5 above).

The next two panels of Figure 5 depict aggregate supply and demand in the call and random price mechanisms, in Treatment 1 and 2, respectively. They illustrate that aggregate supply and demand are very similar in the call and the random price mechanisms.

4.2 Individual Outcomes

The evidence on aggregate outcomes suggests our experimental market conforms rather well to the predictions of theory. We now turn to the analysis of individual outcomes, to study whether well behaved aggregate outcomes stem from well behaved individual actions.

Inspecting individual supply and demand curves reveals significant heterogeneity in individual behaviour. To illustrate this, Figure 6 plots the supply functions of four different Type 1 participants in Treatment 1. The behaviour depicted in Panel A is quite noisy, and involves several dominated actions, in contradiction with Implication 1. In contrast, the behaviour depicted in Panel B is very steady and conforms to the prediction of theory for an infinitely risk averse expected utility maximizer. Two other examples are in Panel C, where the participant's behaviour conforms to the prediction of theory for finite risk aversion, and Panel D, where the participant's behaviour is optimal for a risk loving or risk neutral agent (both in line with Implication 1). Which one of these examples is more representative of the data? We hereafter quantify the extent to which individual actions conform to the predictions of rational choice.

Implications 1 and 2 offer sharp predictions for rational supply and demand: For example, supplying more than two shares at $S < 60$, or less than two shares at $S > 60$ is a dominated action for an FOSD Type 1 agent. Symmetrically, demanding less than two shares at $S < 60$, or more than two shares at $S > 60$ is a dominated action for an FOSD Type 2 agent.

To test the predictions from theory, we compute the frequency of dominated and undominated actions at prices other than 60 for Type 1 or 2 agents.[21] As mentioned above, 68 participants, in our experimental market, were of Type 1 (i.e., they were sellers). Each participated in 8 replications of the market. For each of these participants, and each replication, we computed the proportion of dominated actions across the 10 possible prices.[22] We then computed, for each replication, the average across 68 participants of this proportion. The solid line in Figure 7, Panel A, plots this average, while the dashed line are the 10% confidence interval. Figure 7, Panel B, offers a similar plot for the 68 Type 2 (buyers) participants. For these participants there are only 4 replications. During the 4 other replications these participants

[21] At price 60, or for Type 3 agents, there are no dominated actions.

[22] There are 11 possible prices, but at price 60 there is no dominated action.

were of Type 3, for which we are not able to define dominated actions without making any parametric restriction on preferences.

Inspecting Panel A, one can see that, for sellers, the average proportion of dominated actions starts around 35% at the first replication, and declines to 25% at the last replication. It is always significantly lower than 50%. Hence we reject the pure noise hypothesis that participants simply choose randomly any quantity between 0 and 4. The proportion of dominated actions is also always significantly larger than 0. Hence we reject the fully rational choice hypothesis. Similarly, Panel B shows that, for buyers, the frequency of dominated actions starts from slightly less than 35% at the first replication and then steadily declines to less than 25% at the last replication. Again, both pure noise and fully rational choice are rejected. Thus, participants are not fully rational, but learn to play more rationally as they become more experienced.

Further statistical evidence is offered in Table 1. The table displays estimates for probit regressions in which the dependent variable is the indicator that an action is first-order stochastically dominated. Our data include a total of 12,287 observations: one per participant, per period and per price. We focus on Type 1 and Type 2 participants and on prices different from 60, because these are the cases in which we can identify first order stochastically dominated actions. This leaves us with 8,110 observations. We further drop 520 observations that correspond to the 8 participants who never played any dominated action. We thus end up with 7,590 observations.

We control for individual and price fixed effects. The main regressor of interest is a proxy for the participant's experience, equal to the number of periods during which he/she has already participated. In all the specifications, this variable is significantly negative, suggesting that experience reduces the propensity to choose dominated actions. Another regressor of interest is the indicator that the replication involved random pricing. Its estimated coefficient is not significantly different from zero, consistent with the hypothesis that participants behave in the same way in the random pricing and the market clearing environments. To further analyse learning in our experiment, the third specification (in column 3) involves three additional indicator variables: i) the indicator that the price has previously been selected as transaction price, ii) the indicator that the agent has already played a dominated action at that price and iii) the indicator that this dominated action led to a loss-making trade. The indicator that the price has been previously selected as transaction price is not significant. This suggests that participants do not consider more carefully prices at which they have previously transacted. The indicator that the agent has already played a dominated action at that price is significantly positive. This suggests there is an element of inertia in participants' actions: they tend to repeat previous choices. However, the indicator that the agent has previously chosen a dominated action at that price and correspondingly incurred a loss is significantly negative. This suggests participants learn from their mistakes when the latter are costly.

5 Reconciling the Evidence with Quasi-rational Choice

The above reported findings are puzzling. On the one hand, aggregate outcomes are very much in line with the implications of rational choice. On the other hand, a large fraction of individual actions are inconsistent with rational choice. The latter suggests bounded rationality, the former that individual boundedly rational choices add up to "rational" aggregate outcomes. Thus, to reconcile the different aspects of our experimental findings, we offer a bounded rationality model of individual choice in line with quantal response models (McKelvey and Palfrey, 1995 and 1998), we derive its empirical implications and we confront them to the data.

5.1 Bounded Rationality and Random Actions

As in Luce (1959) we assume that, instead of selecting the deterministic optimal action, players choose randomly among actions, putting larger probability on actions generating higher value.[23]

At a given price (S), each possible quantity $q_i \in [0,4]$, offered or demanded at that price, delivers a given rational value function for agent i: $V_i(q_i, S)$. We assume that the probability that agent i opts for quantity q_i at price S is increasing in $V_i(q_i, S)$. More precisely, for each agent i we assume there exists an increasing function ϕ_i such that the density of q_i is

$$f_i(q_i; S) = \frac{\phi_i(V_i(q_i, S))}{\int_0^4 \phi_i(V_i(q, S))dq}. \tag{7}$$

Because ϕ_i is increasing, $V_i(q, S) > V_i(\hat{q}, S)$ is equivalent to $f_i(q; S) > f_i(\hat{q}; S)$, for all $(q, \hat{q}) \in [0,4]^2$. That is, actions giving a larger rational value function are more likely to be chosen. A special case is $\phi_i(V_i(q_i, S)) = \exp(\lambda_i V_i(q_i, S))$ in which the parameter λ_i measures the respon-

[23] In our competitive setting, agents only need to condition on the price, and don't need to form expectations about the actions of the others. Hence, our framework is simpler than that of quantal response equilibrium, in which "players choose among strategies... based on their relative utility... and assume other players do so as well." (McKelvey and Palfrey, 1995).

siveness of the agent to differences in valuation, as in McKelvey and Palfrey (1995) and Luce (1959).

5.2 Implications of First Order Stochastic Dominance

5.2.1 At price $S = 60$

Consider the choice of the amount q_i sold by Type 1 agent i at price $S = 60$. As stated in Implication 1, if i is risk averse, he finds it optimal to choose $q_i = 2$. Otherwise stated,

$$V_i(2, S = 60) \geq V_i(q, S = 60), \forall q \in [0, 4].$$

By (7), this implies

$$f_i(2, S = 60) \geq f_i(q, S = 60), \forall q \in [0, 4].$$

That is, for a risk averse agent, f_i reaches its maximum at 2, which is the mode of the distribution of q_i.

Now consider two possible choices: $q_i = 2-x$ and $q_i = 2+x$, where $x \in (0, 2]$. When $S = 60$, they give rise to the same lottery. Hence, $V_i(2-x, S = 60) = V_i(2+x, S = 60), \forall x \in (0, 2]$, and correspondingly $f_i(2 - x, S = 60) = f_i(2 + x, S = 60), \forall x \in (0, 2]$. Hence, for a given function ϕ_i, at price $S = 60$, if agents are risk averse the distribution of q_i is symmetric around its mean, equal to 2:

$$E_{\phi_i}(q_i) = \int_0^4 q f_i(q; S) dq = \int_0^4 q \frac{\phi_i(V_i(q, S))}{\int_0^4 \phi_i(V_i(q, S)) dq} dq = 2, \tag{8}$$

where the expectation is parametrised by the function ϕ_i of agent i. Now suppose the ϕ_i of the participants are independent draws from the same distribution. Then we have that the q_i are i.i.d random variables, such that, at price $S = 60$, by the law of large numbers

$$\frac{\sum_{i=1}^{N} q_i}{N} \rightarrow 2,$$

with probability 1 as N goes to infinity. Thus, as stated in the next implication (whose proof is in appendix), aggregate market outcomes should go to rational competitive equilibrium when the number of participants goes to infinity.

Implication 6 *When participants behave according to (7), if they are risk averse, as the number of participants goes to infinity, the aggregate outcome in Treatment 1 goes to that arising with rational choice: The equilibrium price goes to the expected dividend, 60, and the average per–agent trade goes to the full hedge trade, 2.*

Implication 6 rationalizes the empirical findings presented in the previous section. When participants behave according to (7), individual behaviour is noisy and involves dominated actions. Yet, as the number of participants grows large, the aggregate outcome goes to the rational choice equilibrium outcome. In particular, in Treatment 1, in which there is no aggregate risk, the price minimizing the wedge between supply and demand is on average equal to the expected dividend and the average transaction at that price is 2.

5.2.2 At price $S \neq 60$

Now turn to participants' choices at other prices than 60. Consider the supply of a risk averse Type 1 agent at price $S < 60$. As shown in the analysis leading to Implication 1, for prices below 60, selling less than 2 first order stochastically dominates selling more than 2, hence $V_i(2-x,S) > V_i(2+x,S)$, $\forall\ x \in (0,2]$. By (7), this implies $f_i(2-x,S) > f_i(2+x,S)$. That is, agent i plays dominated actions less often than non-dominated ones, which implies

$$F_i(2+\bar{x},S) - F_i(2+\underline{x},S) < F_i(2-\underline{x},S) - F_i(2-\bar{x},S), \forall S < 60, \forall 0 \leq \underline{x} \leq \bar{x} \leq 2,$$

where $F_i(q,S)$ denotes the cdf of quantity q at price S, when i behaves according to (7). Symmetric arguments apply for Type 2. Our next implication follows from the Glivenko–Cantelli Theorem.

Implication 7 *When participants are FOSD and behave according to (7), as the number of Type 1 participants goes to infinity, the empirical distribution of their actions is such that*

$$\Pr\left(\frac{\#observations\ in\ [2+\underline{x},2+\bar{x}]}{N} < \frac{\#observations\ in\ [2-\bar{x},2-\underline{x}]}{N}\right) \tag{9}$$

converges to 1 for $S < 60$ *and to* 0 *for* $S > 60$. *Similarly, for Type 2 agents (9) converges to 0 for* $S < 60$ *and to* 1 *for* $S > 60$.

Implication 7 states that, in a large population of FOSD agents behaving according to (7), the distribution of demand and supply is asymmetric around $q = 2$. This implies that undominated actions (on one side of $q = 2$) should be more frequent than dominated actions (on the other side of $q = 2$), as is the case in the data. To examine an intuitive subset of additional restrictions imposed by Implication 7, it is helpful to define large and small risk-exposures. As explained in Subsection 3.1, when Type 1 or Type 2 participants trade exactly two shares, they are perfectly hedged, as their final wealth is the same in the two states. In contrast, whenever they trade $2+x$ or $2-x$, $x \in (0,2]$, they are exposed to risk. The larger x, the more volatile their final wealth, i.e., the larger their risk-exposure. Thus, we define trades with $x \in (0,1]$ as leading to small risk-exposure, and those with $x \in (1,2]$ as leading to large risk-exposure. Of course, as explained in Subsection 3.1, a large risk-exposure (or a small one) can correspond to a first-order stochastically (FOS) undominated trade (if, e.g., the agent sells more than 2 shares at price $S > 60$) or a FOS-dominated trade (e.g., if the agent sells strictly less than 2 shares at price $S > 60$).

Relying on this categorisation, Figure 8 offers more evidence on the extent to which the data are in line with Implication 7. The left column of Figure 8 depicts histograms of the distribution of quantities offered by Type 1 participants. The top row corresponds to quantities offered at prices strictly lower than 60, while the middle row corresponds to quantities offered at prices strictly higher than 60, and the bottom row corresponds to quantities offered at price 60. The right column of Figure 8 offers similar information for quantities demanded by Type 2 participants.

For the left column and the top row of Figure 8, corresponding to sales at prices lower than 60, in line with Implication 7, the empirical distribution of actions is asymmetric around 2, putting more weight on large FOS-undominated risk-exposures ($q \in [0,1)$) than on large FOS-dominated risk-exposures ($q \in (3,4]$), and also more weight on small FOS-undominated risk-exposures ($q \in [1,2)$) than on small FOS-dominated risk-exposures ($q \in (2,3]$). Symmetrically, for the middle row of the left column, corresponding to sales at prices larger than 60, the histogram is asymmetric around 2, again putting more weight on large FOS-undominated risk-exposures ($q \in (3,4]$) than on large FOS-dominated risk-exposures $q \in [0,1)$, and also more weight on small FOS-undominated risk-exposures ($q \in (2,3]$) than on small FOS-dominated risk exposures ($q \in [1,2)$). Also, for the top row of the right column, corresponding to purchases at prices lower than 60, in line with Implication 7, large FOS-undominated risk exposures ($q \in (3,4]$) are more frequent than large FOS-dominated risk-exposures ($q \in [0,1)$), and small FOS-undominated risk-exposures ($q \in (2,3]$) than small FOS-dominated risk-exposures ($q \in [1,2)$). For the middle row of the right column, while large FOS-undominated risk-exposures are more frequent than large FOS-dominated risk-exposures, observed frequencies contradict Implication 7 for small risk-exposures.

Overall, the empirical distribution of actions seems rather in line with Implication 7. To offer more statistical information on this point we ran non-parametric sign tests of the null hypothesis that frequencies of FOS-dominated and undominated actions were the same, controlling for the size of the corresponding risk-exposure. To do so as simply as possible, we pooled data across prices above and below 60. Thus, for example, for Type-1 (sellers), we pooled FOS-undominated sales at prices above 60 (with $q \in (2,4]$) with FOS-undominated sales a prices below 60 (with $q \in [0,2)$). The resulting pooled frequencies are in Table 2. The p-value for the null hypothesis that the frequency or large FOS-undominated risk-exposures is equal to that of large FOS-dominated ones is 1.36% for sellers and 0.01% for buyers. For the null hypothesis that the frequency or small FOS-undominated risk-exposures is equal to that of small FOS-dominated ones the p-value is .45% for sellers and .13% for buyers. Thus, in all cases, the null hypothesis is rejected, in line Implication 7.

Figure 9, describes the evolution of the frequencies of small and large FOS-dominated actions, as participants become more experienced. The two full lines plot the frequencies of FOS-dominated actions leading to large and small risk exposures respectively, while the dashed lines depict 90% confidence intervals.[24] In each category (large and small risk-exposures), the frequencies of dominated actions decline as participants get more experienced.

5.3 Implications of Second Order Stochastic Dominance

5.3.1 Individual actions

While first order dominance enables one to rank the frequency of dominated and undominated actions, second order stochastic dominance enables one to rank the frequency of different dominated actions, as stated in our next implication (whose proof is in appendix).

Implication 8 *For Type 1, for $S < 60$, selling $2 + \bar{x}$ is second order stochastically dominated by selling $2 + \underline{x}$, $\forall 0 \leq \underline{x} \leq \bar{x} \leq 2$, while for $S > 60$, selling $\underline{x}$ is second order stochastically dominated by selling $\bar{x}$, $\forall 0 \leq \underline{x} \leq \bar{x} \leq 2$. Similarly, for Type 2, for $S > 60$, buying $2 + \bar{x}$ is second order stochastically dominated by buying $2 + \underline{x}$, $\forall 0 \leq \underline{x} \leq \bar{x} \leq 2$, while for $S < 60$ buying $\underline{x}$ is second order stochastically dominated by buying $\bar{x}$, $\forall 0 \leq \underline{x} \leq \bar{x} \leq 2$,*

Implication 8 predicts that the frequency of large FOS-dominated trades should be lower than that of small FOS-dominated trades (intuitively, large mistakes should be less frequent

[24] The full lines plot average across participants of individual frequencies. The confidence intervals plotted by the dashed lines reflect standard errors across participants.

than relatively small ones). Inspection of Figure 8 and Table 2 shows that this is the case in the data, both for Type 1 and Type 2 participants. To offer more statistical information on this point, again, we ran sign-tests. The p-value for the null hypothesis that the frequency of large dominated trades is equal to that of small dominated trades is 0.00% for Type 1 (sellers) and 1.13% for Type 2 (buyers). So we reject the null hypothesis, as requested by Implication 8.

5.3.2 Aggregate Outcomes

While first and second order stochastic dominance, combined with (7), yields implications for individual actions, as discussed above, it also yields implications for aggregate market outcomes, as stated in our next implication (whose proof is in appendix).

Implication 9 *If Participants are FOSD and SOSD, then the equilibrium price should be lower than 60 in Treatment 2.*

Implication 8 rationalizes the empirical findings presented in the previous section that, in Treatment 2, the price minimizing the wedge between supply and demand is lower than 60.

6 Conclusion

This paper offers an experimental test of the theory of competitive equilibrium in complete markets.

Our first methodological contribution is to design an experimental market that emulates as closely as possible the standard competitive equilibrium model: i) supply and demand curves are elicited, aggregated and crossed to set market clearing prices and ii) the hypothesis that participants are competitive can be tested. Our second methodological contribution is to rely on the simplicity of the experimental setting to pin down precise testable implications on individual behaviour, that hold for a very large class of preferences. Our third methodological contribution is to apply random choice models (in line with Luce (1959) and McKelvey and Palfrey (1995)) to our experimental market setting, and spell out their implications for the distribution of individual actions as well as for aggregate outcomes.

Our main experimental findings are the following:

- At the aggregate level, the experimental complete market conforms to theory. The hypothesis that participants are competitive cannot be rejected. Aggregate supply and demand cross at the expected dividend when there is no aggregate risk, and at a lower price when there is.

- Individual participants, however, frequently choose first order stochastically dominated actions. Yet, dominated actions become less frequent as participants become more experienced, and participants seem to learn from their mistakes.

- Our random choice model reconciles the apparently contradictory findings obtained at the aggregate and individual levels. It predicts that individual deviations will average out, leading to well behaved aggregate supply and demand. The random choice model also imposes further restrictions on the distribution of individual actions, consistent with experimental findings: Dominated actions are less frequent than undominated ones, and large mistakes are less frequent than small ones.

Our experimental findings suggest that, when markets are perfect and complete, individual irrationality does not preclude aggregate outcomes consistent with the predictions of competitive equilibrium. Thus deviations from those predictions, observed in the field, could stem from market imperfection and incompleteness, rather than from limited cognition. It will be interesting, in further work, to extend our methodology to study imperfect markets in the lab.

Appendix

Proof of Implication 6: By the law of large numbers, if Type 1 participants are risk averse, at price $S = 60$ their average supply goes to 2 as the number of Type 1 participants, N_1, goes to infinity, i.e., with probability

$$\lim_{N_1 \to \infty} \frac{\sum_{i=1}^{N_1} q_i^1(60)}{N_1} = 2, \tag{10}$$

where $q_i^1(60)$ denotes the supply of participant Type 1 participant i at price 60.

Symmetrically, the average demand from risk averse Type 2 participants also goes to 2, as the number of Type 2 participants, N_2, goes to infinity, i.e.,

$$\lim_{N_2 \to \infty} \frac{\sum_{i=1}^{N_2} q_i^2(60)}{N_2} = 2, \tag{11}$$

where $q_i^2(60)$ denotes the demand of participant Type 2 participant i at price 60. Now, for $S = 60$ to be an equilibrium price, we need that $\sum_{i=1}^{N_1} q_i^1(60) = \sum_{i=1}^{N_2} q_i^2(60)$. For finite $N_1 = N_2$ this is equivalent to

$$\frac{\sum_{i=1}^{N_1} q_i^1(60)}{N_1} = \frac{\sum_{i=1}^{N_2} q_i^2(60)}{N_2},$$

which, by (10) and (11), holds in the limit, as the number of participants goes to infinity, if they are all risk averse. QED

Proof of Implication 8: Consider Type 1 (the proof for Type 2 is symmetric), the lottery corresponding to the sale of $2 + \underline{x}$ has the following two equiprobable outcomes:

$$W(u, S)|_{Q=2+\underline{x}} = 600 + (S - 120)(2 + \underline{x}) = 360 + 2S + (S - 60)\underline{x} - 60\underline{x},$$

and

$$W(d,S)|_{Q=2+\underline{x}} = 360 + S(2+\underline{x}) = 360 + 2S + (S-60)\underline{x} + 60\underline{x}.$$

The lottery corresponding to the sale of $2+\bar{x}$ has the following two equiprobable outcomes:

$$W(u,S)|_{Q=2+\bar{x}} = W(d|2+\underline{x}) + (S-60)(\bar{x}-\underline{x}) - 60(\bar{x}-\underline{x}).$$

$$W(d,S)|_{Q=2+\bar{x}} = W(d|2+\underline{x}) + (S-60)(\bar{x}-\underline{x}) + 60(\bar{x}-\underline{x})$$

By construction $S-60<0$. Moreover, plus or minus $60(\bar{x}-\underline{x})$ with equal probability is a mean preserving spread. Hence $((W(u|2+\bar{x}),\frac{1}{2};W(d|2+\bar{x}),\frac{1}{2}))$ is second order stochastically dominated by $((W(u|2+\underline{x}),\frac{1}{2};W(d|2+\underline{x}),\frac{1}{2}))$. Note that, here again, equiprobability is key for the proof. If $S=60$, then $2+\bar{x}$ is a mean preserving spread of $2+\underline{x}$, since

$$W(u,S=60)|_{Q=2+\bar{x}} = 600 + (60-120)(2+\bar{x}) = 480 - 60\bar{x},$$

and

$$W(d,S=60)|_{Q=2+\bar{x}} = 360 + 60(2+\bar{x}) = 480 + 60\bar{x},$$

while

$$W(u,S=60)|_{Q=2+\underline{x}} = W(u,S=60)|_{Q=2+\bar{x}} - 60(\bar{x}-\underline{x}),$$

and

$$W(d,S=60)|_{Q=2+\underline{x}} = W(d,S=60)|_{Q=2+\bar{x}} + 60(\bar{x}-\underline{x}).$$

QED

Proof of Implication 9: In Treatment 2, for Type 1 agents, selling less than 2 at prices $S > 60$ is first order stochastically dominated. Hence, if participants are FOSD agents and behave according to (7), expected individual supply at $S > 60$ is larger than or equal to 2. Hence, by the law of large numbers

$$\lim_{N_1 \to \infty} \frac{\sum_{i=1}^{N_1} S_i(60)}{N_1} \geq 2, \forall S > 60.$$

For Type 3 agents, buying $2 + \bar{x}$ is second order stochastically dominated by $2 + \underline{x}$, $\forall 0 \leq \underline{x} \leq \bar{x} \leq 2$. Hence, if participants are FOSD and SOSD agents and behave according to (7), expected individual demand at $S > 60$ is strictly lower than 2. Hence, by the law of large numbers

$$\lim_{N_2 \to \infty} \frac{\sum_{i=1}^{N_2} D_i(60)}{N_2} < 2, \forall S > 60.$$

This implies the equilibrium price should be lower than 60. QED

References

Arrow, K., (1964). "The Role of Securities in the Optimal Allocation of Risk-bearing," *Review of Economic Studies* 31 (2): 91-96.

Elena Asparouhova & Peter Bossaerts & Nilanjan Roy & William Zame, (2013). "Lucas In The Laboratory," NBER Working Papers 19068, National Bureau of Economic Research.

Becker, G., M. DeGroot and J. Marschak (1964). "Measuring utility by a single–response sequential method." *Behavioral Science.* 9, 226–32.

Bohm, J., J. Linden and J. Sonnegard, (1997). "Eliciting Reservation Prices: Becker-DeGroot-Marschak Mechanisms vs. Markets." *Economic Journal.* 107, 1079–89.

Borch, K., (1962). "Equilibrium in a reinsurance market", *Econometrica*, 30 (3) 424–444.

Bossaerts. P., (2008). "The Experimental Study of Asset Pricing Theory," *Foundations and Trends in Finance*, Vol. 3, No. 4. 289–361

Bossaerts, P., and C. Plott (2004): "Basic Principles of Asset Pricing Theory: Evidence from Large-Scale Experimental Financial Markets," *Review of Finance*, 8(2), 135–169.

Bossaerts, P., C. Plott, and W.R. Zame (2007): "Prices and Portfolio Choices in Financial Markets: Theory, Econometrics, Experiments," *Econometrica*, 75(4), 993–1038.

Bossaerts, P., D. Meloso and B. Zame (2013). "Dynamically Complete Experimental Asset Markets." mimeo.

Crockett, S. and J. Duffy (2013). "An Experimental Test of the Lucas Asset Pricing Model," Working Paper.

Debreu, G. (1959). *The Theory of Value: An axiomatic analysis of economic equilibrium.*

Goeree, J. and C. Holt (2004). "A Model of Noisy Introspection." *Games and Economic Behavior*, 46, 365-382.

Goeree, J., C. Holt and T. Palfrey (2003). "Risk Averse Behavior in Generalized Matching Pennies Games." *Games and Economic Behavior*, 45, 97-113.

Goeree, J., C. Holt and T. Palfrey (2002). "Quantal Response Equilibrium and Overbidding in Private-Value Auctions" *Journal of Economic Theory*, 104, 247-272.

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

Luce, R. D. (1959). *Individual Choice Behavior: A Theoretical Analysis.* New York: Wiley.

McCabe, K., S. Rassenti and V. Smith (1992). "Designing Call Auction Institutions: Is Double Dutch The Best?" *The Economic Journal*, 102, 9-23.

McKelvey, R. and T. Palfrey (1995). "Quantal Response Equilibria for Normal Form Games." *Games and Economic Behavior*, 10, 6–38.

McKelvey, R. and T. Palfrey (1998), "Quantal Response Equilibria for Extensive Form Games." *Experimental Economics,* 1, 9–41.

Machina, M.J. (1982). ""Expected utility" analysis without the independence axiom." *Econometrica*, 50, 277-323.

Mas-Colell, A., M.D. Whinston, and J.R. Green (1995): *Microeconomic Theory*, New York, Oxford: Oxford University Press.

Plott, C. and S. Sunder (1982). "Efficiency of Experimental Security Markets with Insider

Information: An Application of Rational-Expectations Models." *Journal of Political Economy,* 90, 663-698.

Plott, C. and S. Sunder (1988). "Rational Expectations and the Aggregation of Diverse Information in Laboratory Security Markets." *Econometrica*, 56, 1085-1118.

Plott, C. and V. Smith (1978). "An experimental examination of two exchange institutions." *The Review of Economic Studies* 45, 133–153.

Plott, C. and K. Pogolresky (2015). "Call Market Experiments: Efficiency and Price Discovery Through Multiple Calls and Emergent Newton Adjustments." Working paper. Caltech.

Quiggin, J.C., "A theory of anticipated utility." *Journal of Economic Behavior and Organizations*, 3, 323–343.

Segal, U. and A. Spivak (1990). "First order versus second order risk aversion." *The Journal of Economic Theory* 51, 111–125.

Smith, V. (1962). "An experimental study of competitive market behaviour." *The Journal of Political Economy* 70, 111–137.

Smith, V, G. Suchanek and W. Williams (1988). "Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets." *Econometrica*, 56, 1119-1151.

Smith, V. L., W. Williams, W. K. Bratton, and M. Vannoni (1982). "Competitive Market Institutions: Double Auctions vs. Sealed Bid-Offer Auctions", *American Economic Review*, 72, 58–77.

Varian, H. (1985): "Additive Utility and Gross Substitutes," Unpublished Manuscript, University of Michigan, Ann Arbor.

Table 1: Probit regressions

The dependent variable is an indicator that the action is FOS dominated. Regressions are estimated over 7,590 actions, by the 128 participants, who chose at least once a dominated action.

	FOS dominated action		
	(1)	(2)	(3)
Number of Periods during which the subject participated	-0.0563***	-0.0558***	-0.0673***
	(-7.98)	(-7.91)	(-8.80)
Indicator that the replication involved Random Pricing		0.0455	0.0489
		(1.43)	(1.52)
Indicator that trading already took place at that price			-0.0735
			(-1.43)
Indicator that subject played dominated action at that price			0.703***
			(8.64)
Indicator that dominated action led to loss-making trade			-0.365***
			(-3.61)
Constant	-1.058***	-1.079***	-1.022***
	(-5.51)	(-5.60)	(-5.26)

t statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2: Empirical frequencies of FOS-undominated and dominated actions

For each of the 68 Type 1 participants, and 68 Type 2 participants, we computed the frequency of dominated and undominated actions, leading to large risk exposure, small risk exposure with trade size different from 2, or trade size equal to 2. We then computed the average across the 136 participants of these frequencies.

	Large undominated	Small undominated	$q = 2$	Small dominated	Large dominated
Type 1 (seller)	20%	29%	19%	21%	10%
Type 2 (buyer)	24%	26%	22%	14%	13%

Figure 1: First order stochastic dominance among lotteries

If $\mu_i' > \mu_i$ then lottery (μ_i',σ_i) (cdf depicted by dashed red line) 1st order stochastically dominates lottery $(\mu_i,\sigma i)$ (cdf depicted by full black line)

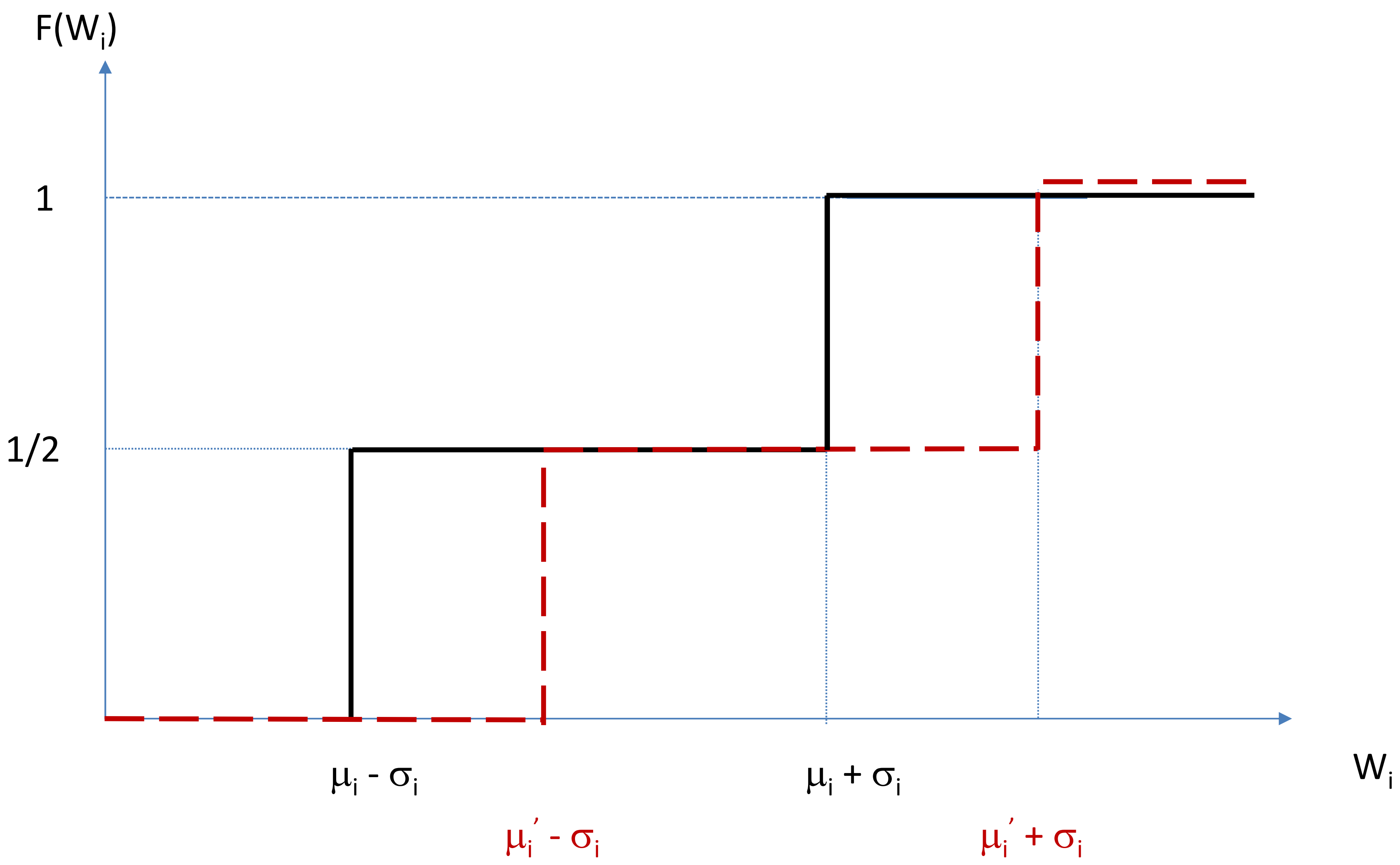
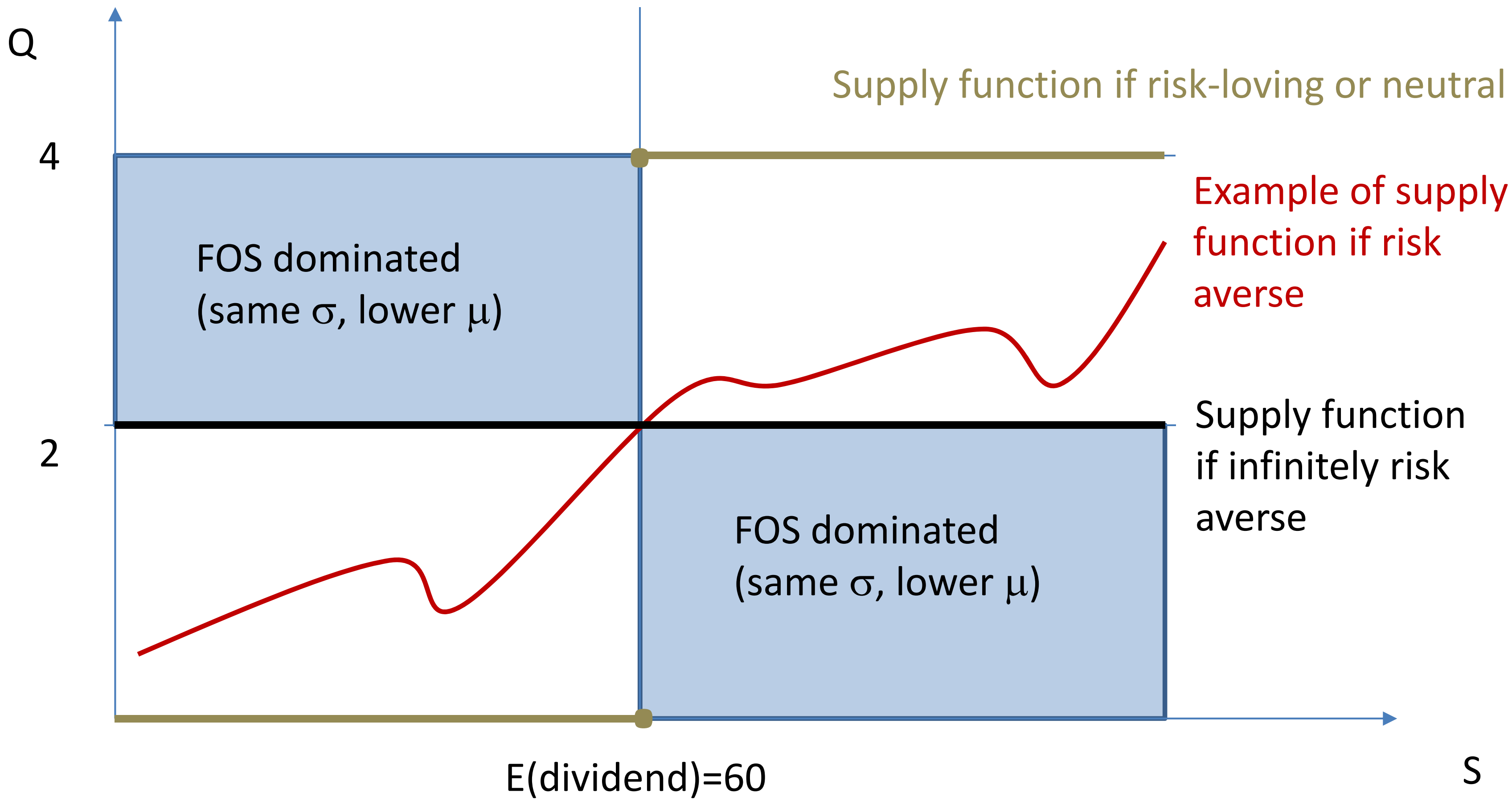


Figure 2, First order stochastically dominated Q

Panel A: Type 1 (seller)



Panel B: Type 2 (buyer)

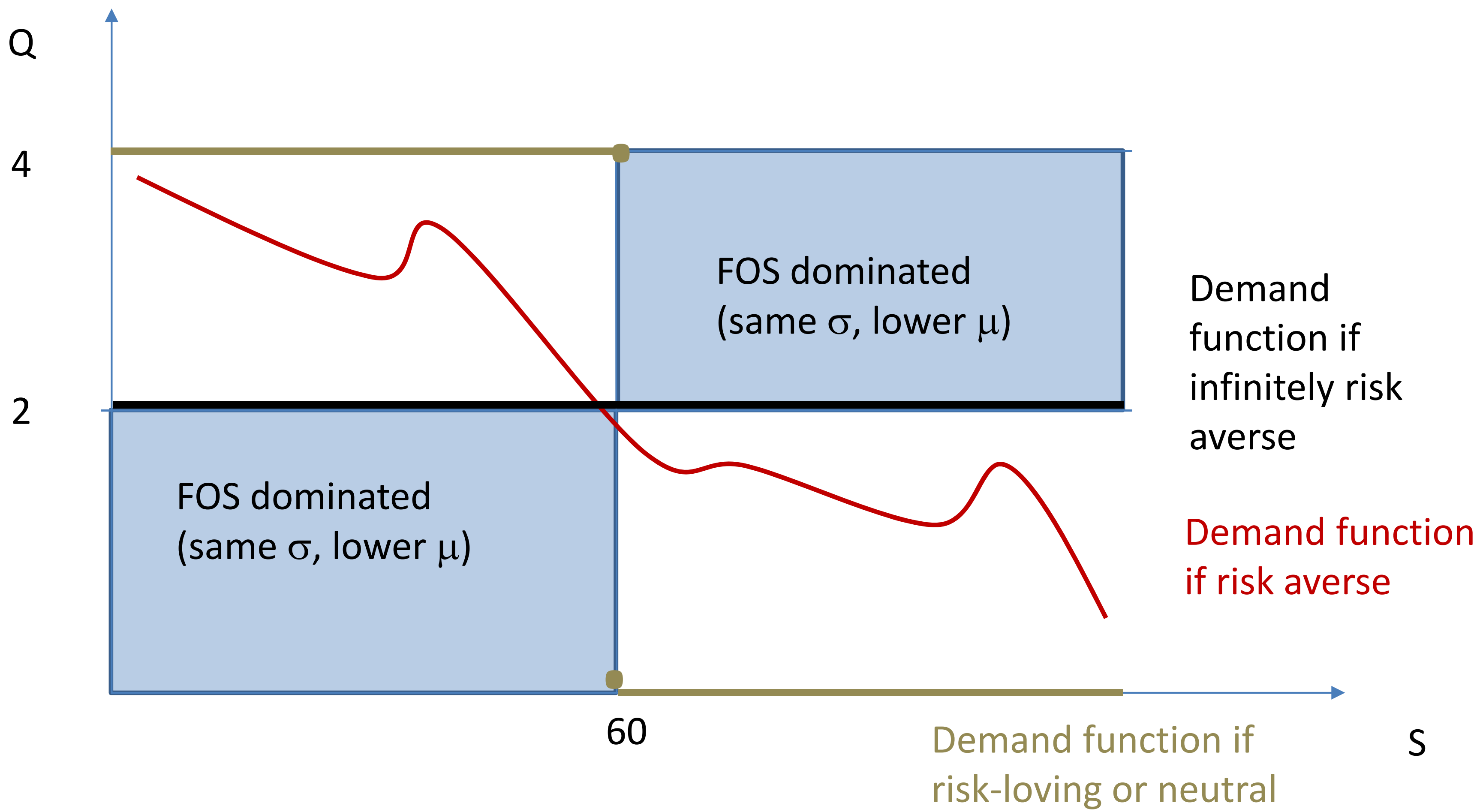


Figure 2, Panel C: First order stochastically dominated Q (Type 3, buyer)

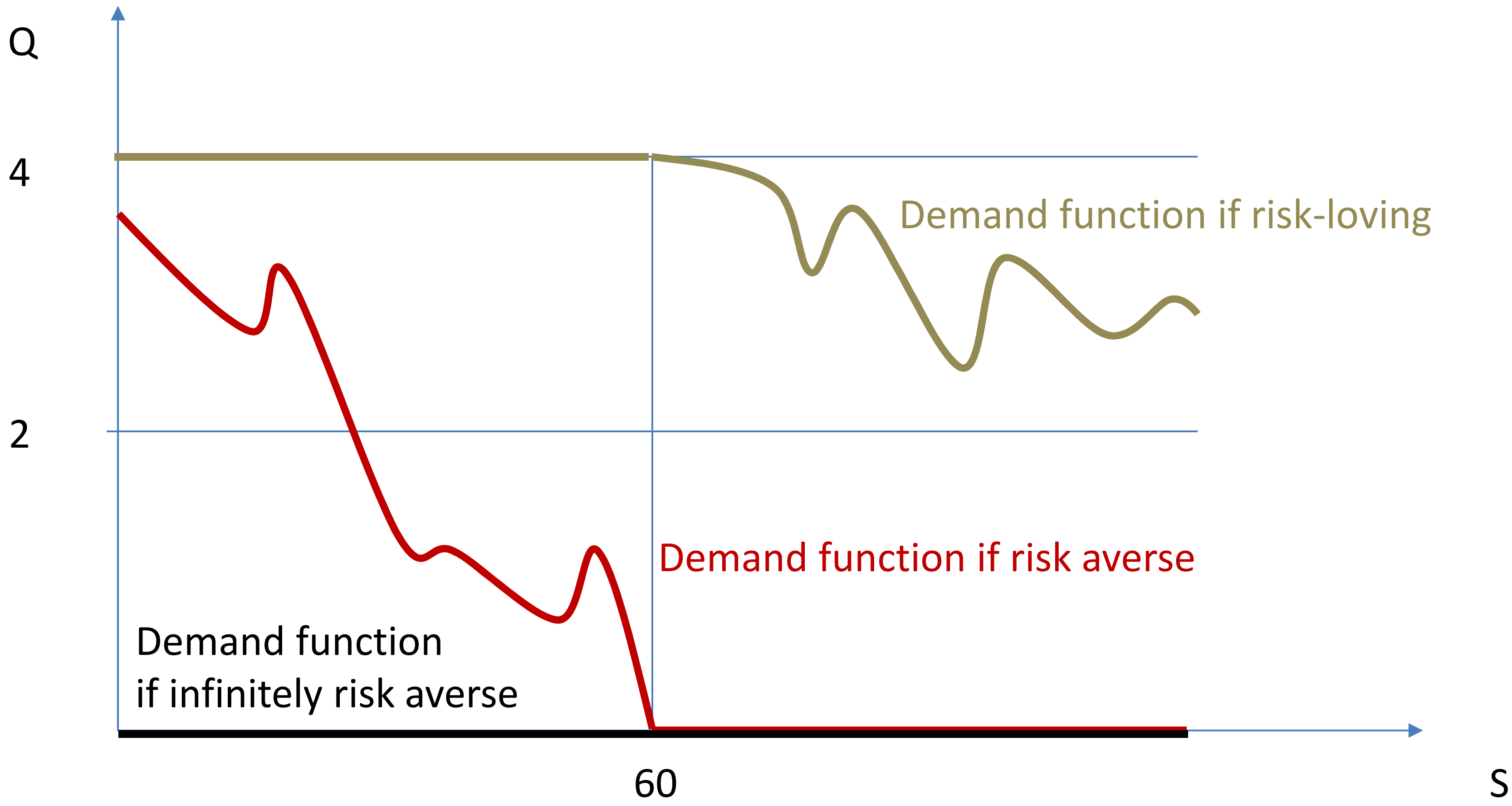


Figure 3: p-values for Wilcoxon signed-rank test of H0 that demand not significantly different in call & random price

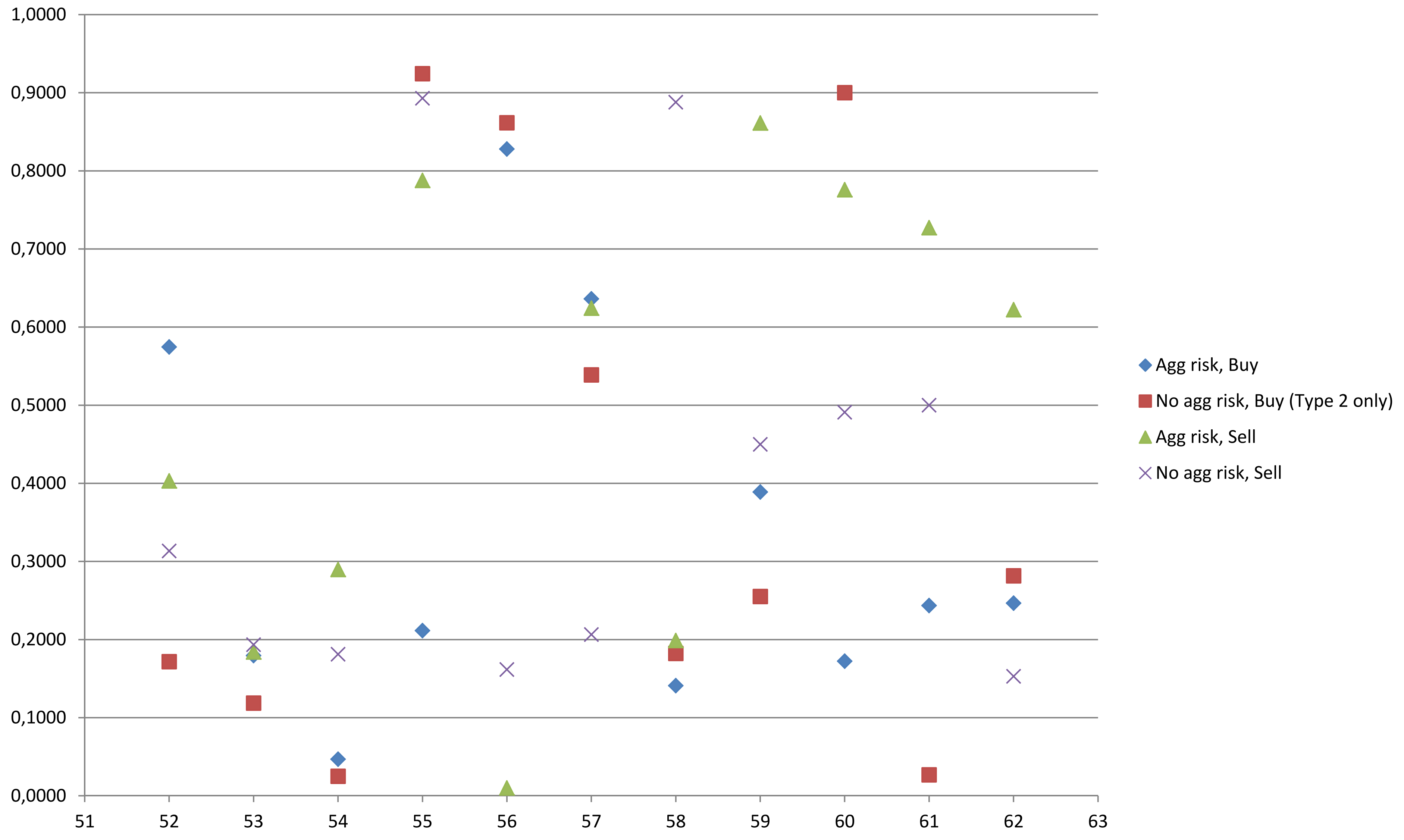


Figure 4: Market clearing price in the 2 treatments (average across replications)

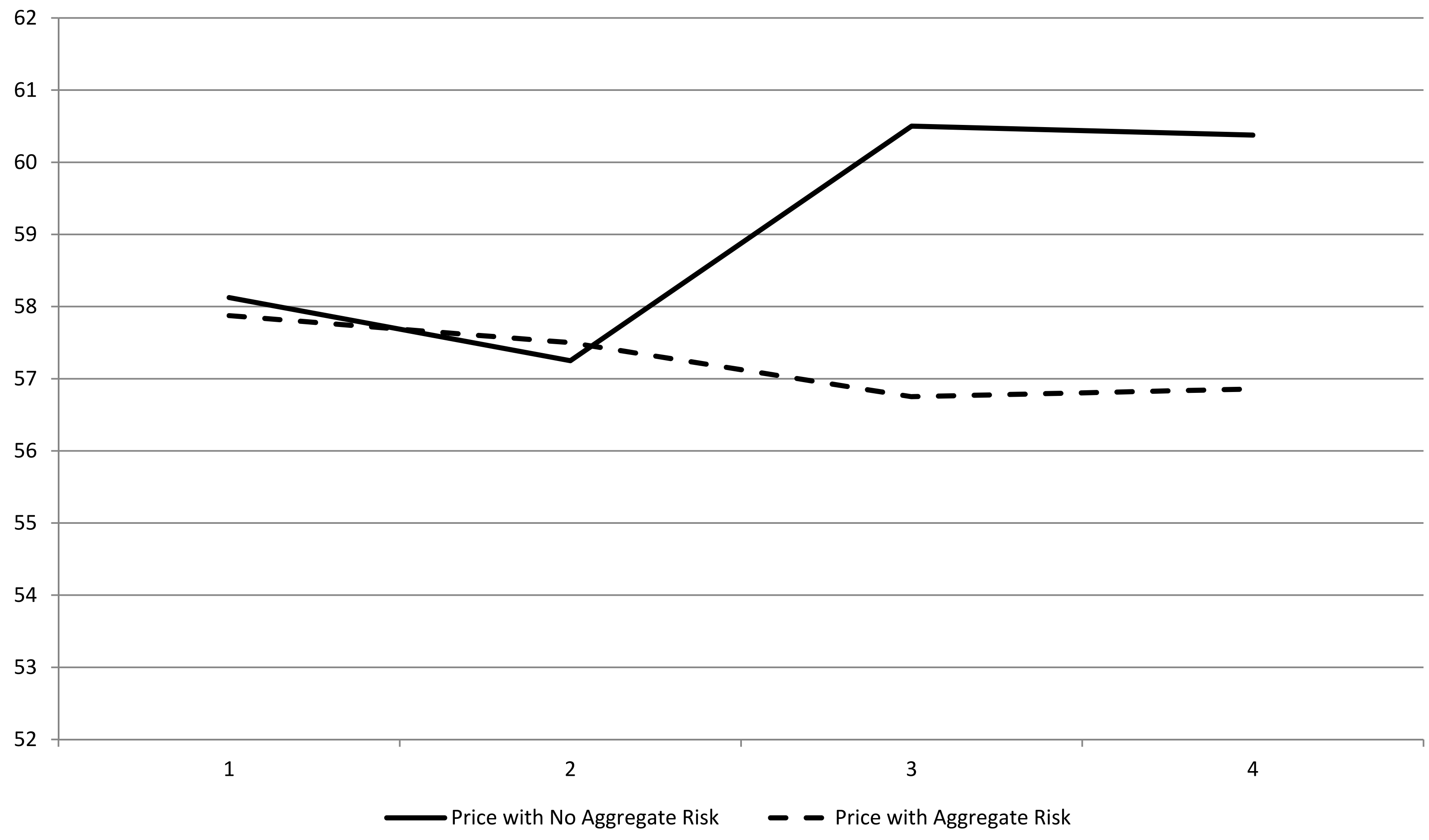
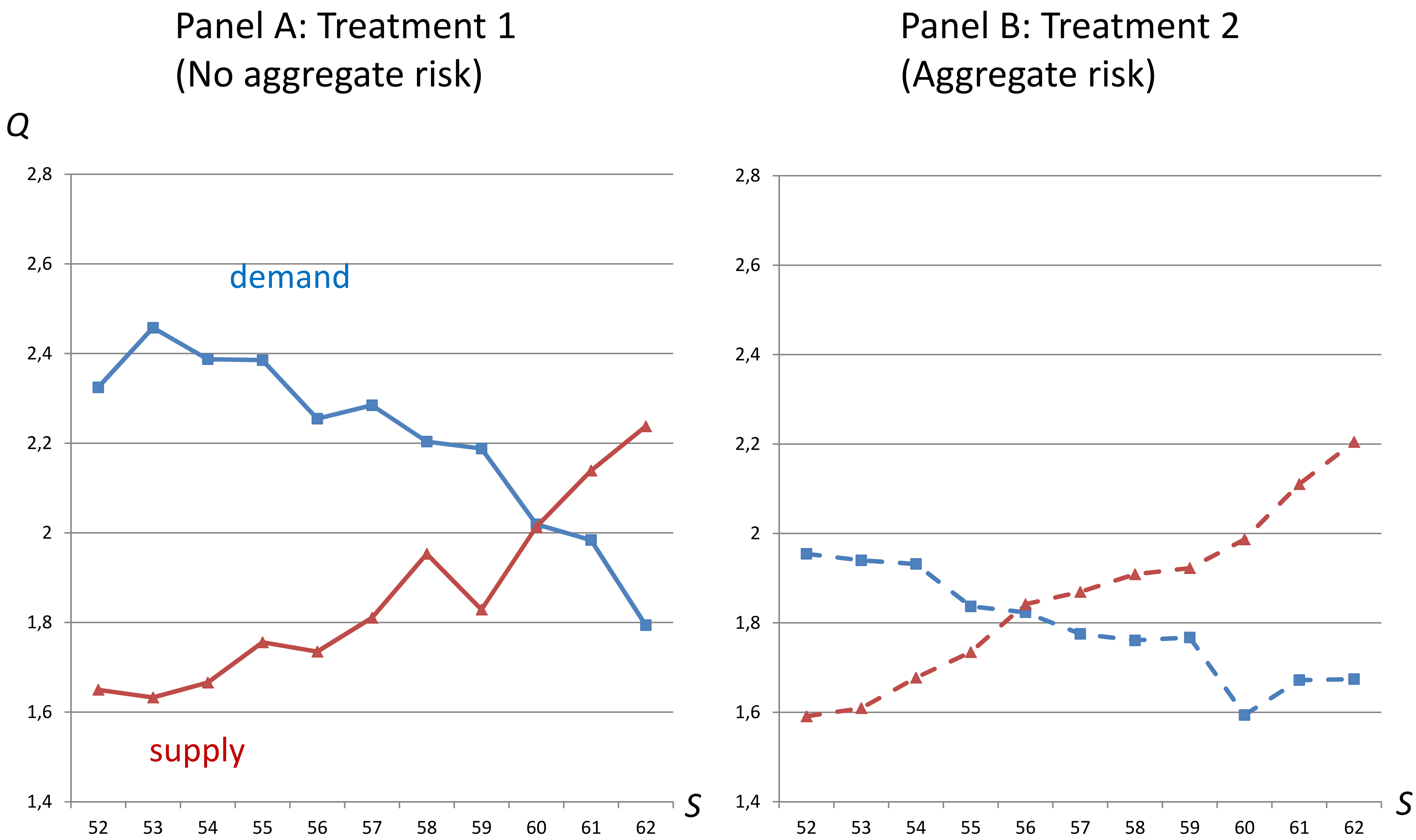
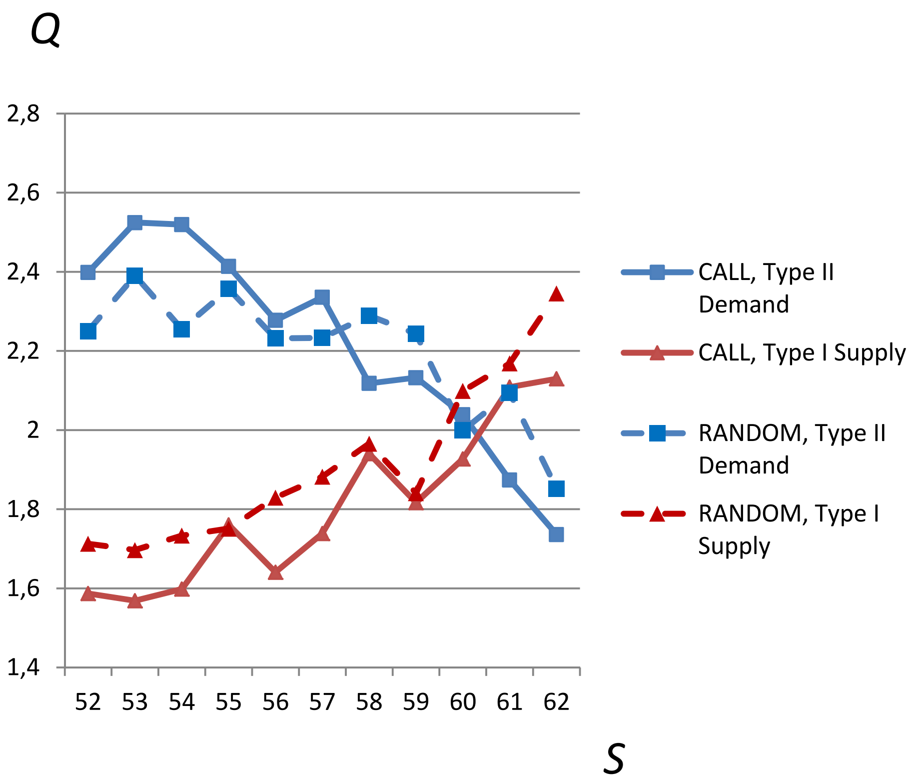


Figure 5: Aggregate supply and demand



Panel C: Treatment 1
(No aggregate risk)



Panel D: Treatment 2
(Aggregate risk)

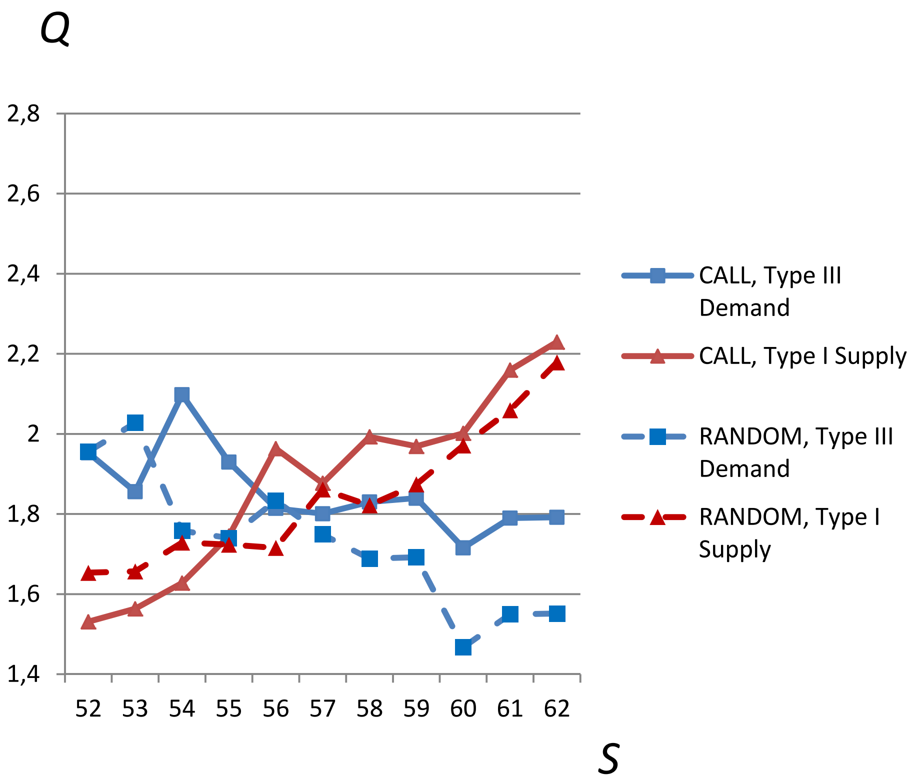


Figure 6: Supply functions of 4 participants

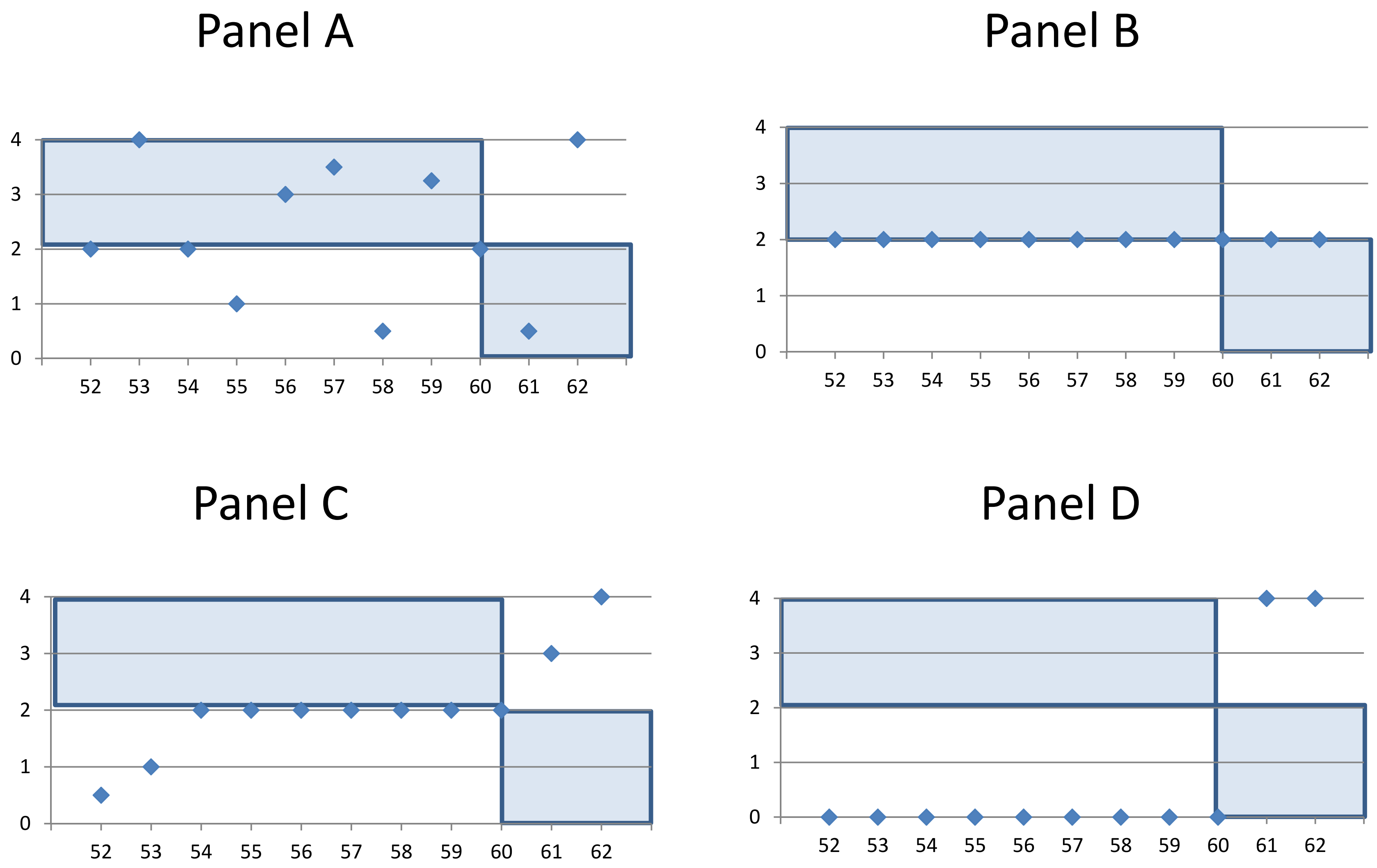


Figure 7, Frequency of first order stochastically dominated actions

Average frequency cross participants (solid line)

90% confidence interval (dashed lines)

Panel A: Type 1 (sellers) throughout
8 replications

Panel B: Type 2 (buyers) throughout
4 replications

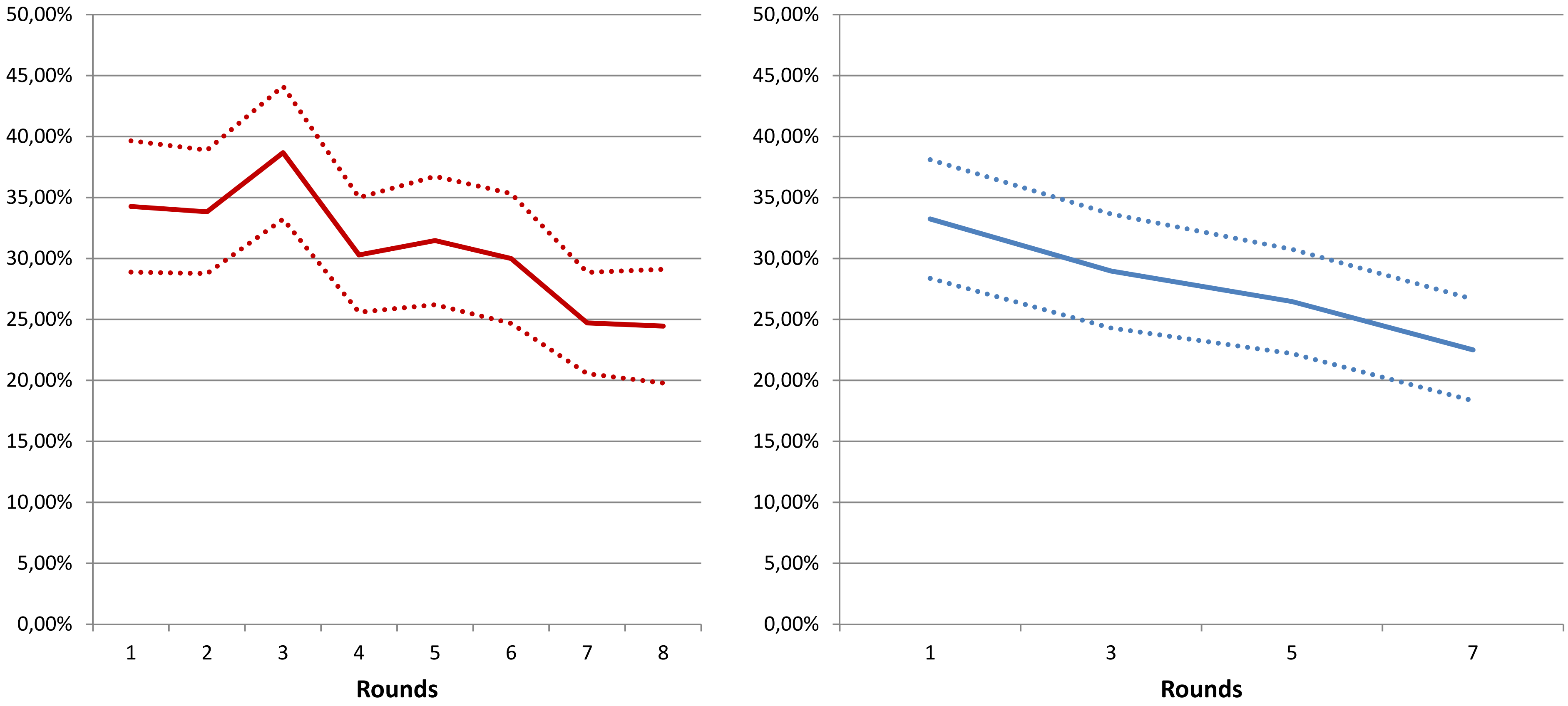


Figure 8: Distribution of transaction size across participants

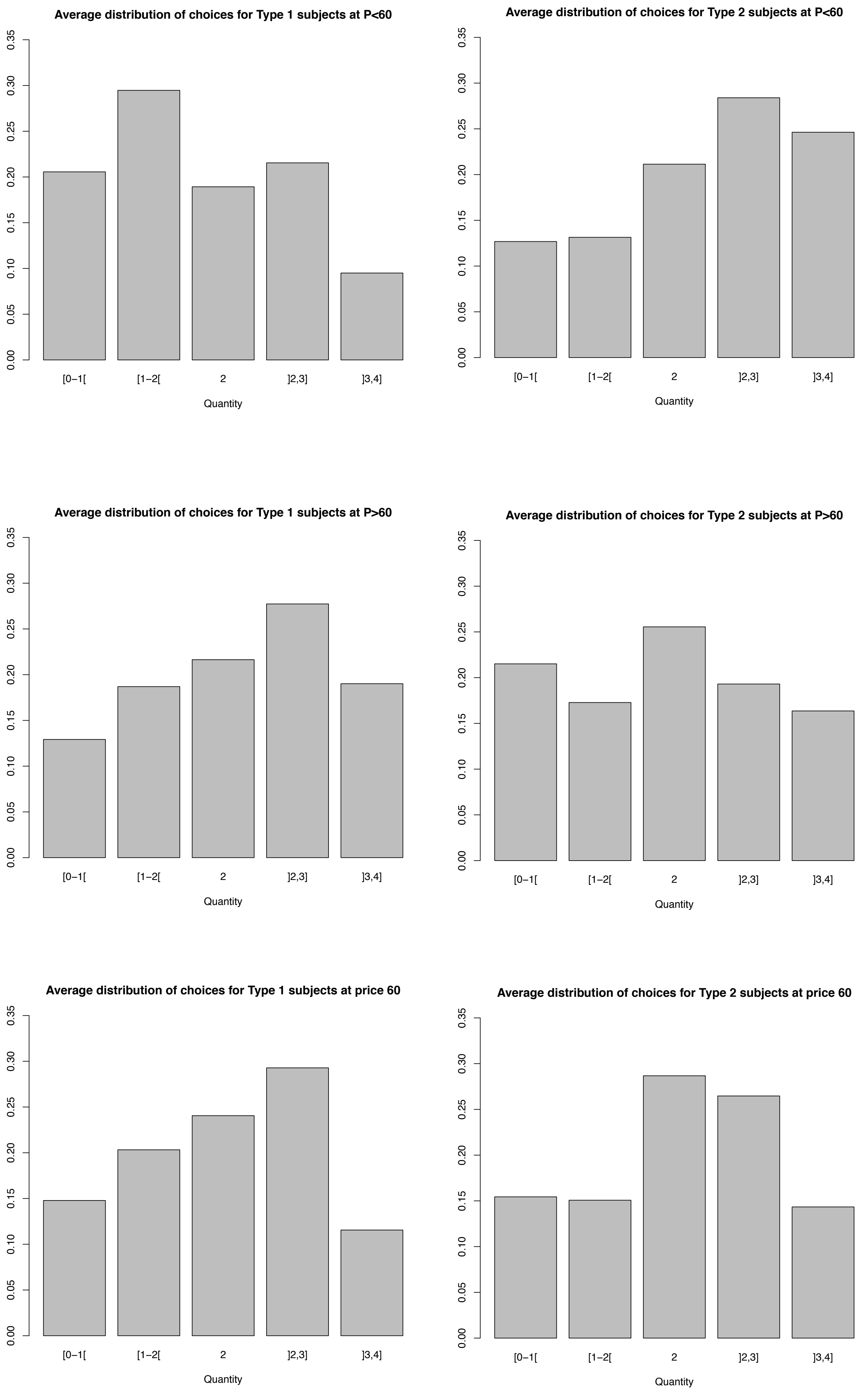


Figure 9: Frequency of FOS-dominated actions leading to large and small risk exposure, throughout 4 Treatment 1 replications

Average frequency cross participants (solid line)

90% confidence interval (dashed lines)