Strategic Behavior in Markets and Teams:

Essays in Experimental Economics

Inauguraldissertation
zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaften
der Universität Mannheim

vorgelegt von
Sihong Zhang, M.Sc.

Graduate School of Economic and Social Sciences
Center for Doctoral Studies in Economics

Sommersemester 2018

Abteilungssprecher:	Professor Dr. Jochen Streb
Referent:	Professor Dr. Henrik Orzen
Korreferent:	Professor Dr. Stefan Penczynski

Tag der mündlichen Prüfung: 09 Juli 2018

Dedicated to my beloved parents.
Thank you for your unconditional love and support. I am very grateful for all the things you have done for me. Without you I would not be where I am now.

Acknowledgments

First and foremost, I would like to thank my advisor Henrik Orzen and my co-advisor Stefan Penczynski for their continuous support and guidance during my PhD studies in Mannheim. Without them, the work presented in this thesis would not have been possible. I am very grateful for the productive discussions we had and their incredibly valuable advice on my work. They were always very encouraging and I have benefited tremendously from working with them. I am much obliged to Henrik Orzen for always taking time and for his very thorough comments which shaped two of the three chapters of these thesis. To Stefan Penczynski, I am greatly indebted for introducing me to programming and conducting laboratory experiments enabling a very quick and smooth start into this field of research for me. Thank you both for your time and effort and for being such interested and dedicated advisors!

I would also like to thank Andrej Svorenčík for helping me with the organization of my sessions in the mLab in Mannheim and Jörg Oechssler for offering an elective in experimental economics in Mannheim, which spurred my interest in experimental topics and encouraged my to do my own research in this field.

Furthermore, I would like to thank the experimental economists at ZEW and the members of the HeiKaMax for giving me the opportunity to present my research and for their valuable comments.

I would also like to take this opportunity to thank all my friends and fellow doctoral students not only for making my time in Mannheim so enjoyable but also for their support. I still remember the long hours we spend in the library working together on our problem sets in the first year. And while conducting my research, I could always count on their help whenever there was an urgent situation in the laboratory. I deeply cherish the friendships that formed during my time in Mannheim.

And at last but not least, I would like to say thank you to my friends and family. Thank you for always caring, believing in me, understanding me and supporting me!

Contents

List of Figures

List of Tables

1 General Introduction

In classical economics the baseline assumption when analyzing the strategic behavior and performance of firms and individuals is that decision-makers are completely rational and selfish (focused on maximizing their own payoff). While this assumption is a useful theoretical benchmark, it do not always apply to peoples behavior in reality. Human beings are often only partially rational, which may lead to unsophisticated and inefficient decisions.

In the three self-contained chapters of this dissertation, I analyze strategic decision-making and the dynamics of markets and teams using laboratory experiments. Chapter 2, which is based on joint work with Stefan Penczynski, investigates how much information sellers are willing to disclose under competition and whether consumers make sophisticated buying decisions in the sense that they can draw the right conclusions from the information that is and is not provided to them. Chapter 3, which is based on joint work with Henrik Orzen, considers only the firm side and analyzes the assumption that due to competition and market dynamics only rational firms that decide optimally in terms of profit maximization survive in the long-run. Chapter 3 tackles the question of the validity of laboratory experiments in our market setting. In Chapter 4 I examine dynamic investment behavior in a teamwork setting where investments have to be accumulated over time to complete a certain project. I seek to analyze whether there is cooperative behavior even if collective interests conflict with individual interests as costs are private and benefits are public.

Chapter 2

In B-to-C markets consumers are often overwhelmed by the number of brands and product variations offered and, when trying to make informed decisions, rely on third-party product information. Often they are unable to draw the proper conclusions from the information provided or fail to consider the possibility that information may have been deliberately withheld. Firms can exploit this rather naive behavior by selectively providing information that is advantageous to them. One thing that might help even unsophisticated consumers is competition. Competitive forces and the need to attract more consumers can act as a lever forcing firms to disclose more information.

To analyze the disclosure of verifiable information in settings with and without seller competition we established an experimental framework with varying market and information structures. This enabled us to see how competition on the one hand and sellers ability to purchase further verifiable information on the other affected firms disclosure decisions and consumers conclusions.

We found that increased competition had a significant positive effect on the disclosure of information. Sellers often chose to disclose selected information and buyers on average accounted for this by bidding less than the reported average. However, buyers on average did not fully compensate for sellers' selective disclosures and, surprisingly, did particularly poorly in the competitive setting. When sellers had the option to purchase additional information, unobserved by the buyers, the stronger selection upon purchase was counterbalanced by stronger overall compensation on the buyer side, on average lowering buyers' overbidding compared to the benchmark cases.

Chapter 3

Experiments in industrial organization in which students assume the role of firms are often criticized due to concerns about external validity. There are several facets to this problem. In this paper we discuss one common line of reasoning that uses an evolutionary argument. The idea is that in "real" markets competition favors firms that decide optimally and market forces eliminate firms that do not maximize profits. Therefore, it is safe to assume that firms that are still present in a market conform to the ideals of the rational choice model.

To investigate the relevance of this claim to the validity of laboratory experiments we set up a deliberately complex market setting and compared behavior and market outcomes in two treatments. In the Baseline treatment subjects were allocated at random to markets of different sizes and then interacted strategically by investing in production capacity and by competing on prices over several periods. Losses or poor performance were distinct possibilities but did not lead to any selection. The Market Force treatment was identical except that subjects could self-select into the market as long as they had sufficient liquidity and were expelled from the market if they did not. We examine sorting effects and market performance.

We did not find strong evidence in favor of the market selection argument. Although we observed a clear sorting effect when market entry was endogenous, the self-selection did not affect outcomes significantly. In both treatments, firms oftentimes did not fully exploit their market power in market with lower competitive pressure, while in markets with more competitors capacities and prices tend to be too high affecting profits negatively. Overall, we found that experience had a significantly positive effect on profitability and although the setting was deliberately complex, participants' decisions converged on average to the equilibrium solution.

Chapter 4

The term "procrastination" is often associated with individual behavior. In fact, procrastination is not just an individual issue and can be a major roadblock to teamwork which can lead to the failure of a project. One common remedy to counteract procrastination is setting deadlines.

To investigate dynamic investment behavior in individual and team projects as well as the effects of team size and deadlines, I employ a setting where participants make effort decisions over several periods in order to complete a project. Treatments varied in two different dimensions: team size and type of deadline. Deadlines were set in one of two ways: (i) Team members were allowed to suggest a maximum number of periods for project completion and a deadline was then defined through a mechanism based on the suggestions made (self-imposed deadline). (ii) Teams were told how many periods they had to complete the project (exogenous deadline).

I found that, in the absence of deadlines, there was little procrastination when participants worked alone, and all of them managed to complete their projects. However, teams systematically failed to coordinate on the most efficient outcome. In contrast to working alone, teamwork seemed to foster procrastination, and team size even enhanced that negative effect. Large teams needed almost twice as long as small ones to complete a project. Overall, introducing a deadline helped to mitigate the problem of procrastination. However, it negatively affected project completion rates, particularly when the deadline was exogenous rather than self-imposed. In addition, team size also showed a strong negative effect on the completion rate, independently of the amount of time teams had to finish a project.

2 Disclosure of Verifiable Information under Competition[1]

2.1 Introduction

Many times per day, consumers are confronted in media advertisements and on product packages with verifiable product information that producers choose to convey. This might be a microwave oven's "very good" result in an independent test, the certified absence of a chemical in a plastic container, or the scientifically proven tolerance of humans to cholesterol from egg. How useful the presence and in particular the absence of such information is to the consumer often depends on her skill to judge the selection of disclosed information.

In order to investigate both the seller's selection and the appropriateness of consumer's inference, this study establishes a rich experimental framework with varying market and information structure. In particular, the framework enables us to see how seller competition as well as the seller's ability to purchase further verifiable information change the disclosure behavior of sellers and the inference of buyers.

In contrast to cheap talk games, the verifiability of the information provides a clear link to the underlying state of the world, so that such disclosure games often feature full information revelation. However, for this to work, the seminal theoretical treatments

[1]This chapter is based on joint work with Stefan P. Penczynski.

emphasize the important role of the buyer's sophistication in the form of "skepticism" in the inference process (Grossman, 1981; Milgrom and Roberts, 1986). Recent experimental studies suggest that the level of sophistication is often low enough to cause harmful inferences on the receiver end of the information (Jin, Luca and Martin, 2016; Benndorf, Kübler and Normann, 2015).

Looking at the examples above, a likely aid in favor of the consumer is the competition between sellers. Indeed, for the case of incomplete sophistication, Milgrom and Roberts (1986) show that competition among parties with strongly-opposed interests reestablishes full information revelation as sellers increase disclosure. Although not studied theoretically, competition and choice of seller might also change buyers' sensitivity to a given level of non-disclosure.

On the other hand, a complicating factor for consumers is usually the uncertainty about the size of the set of verifiable information available to the seller. For example, a favorable subset of verifiable information from nutrition science might shape consumers' view of food items such as eggs more than it should if they think that no other evidence exists.[2] Theoretical studies show that uncertainty about the set of evidence makes information transmission more difficult and can inhibit the unraveling of information (Dziuda, 2011; Felgenhauer and Schulte, 2014).

Our experimental framework is designed to investigate in detail seller disclosure and buyer inference in a versatile setting with varying market and information structure. The setup naturally reflects the rich informational setting of product markets as well as the conflict of interest that arises from sellers' incentives to increase buyers' willingness to pay and buyers' incentives to evaluate the product correctly. Sellers have ten pieces of noisy evidence about the product quality, whose true value only they know precisely. For each one, they can decide whether or not to reveal it. The bidding mechanism gives buyers incentives to bid their true valuation of the product.

[2] For example, in a recent influential meta-study of randomized-controlled trials on the effect of dietary cholesterol on blood cholesterol, 11 out of 12 studies chosen to be analyzed were funded by the egg or fish industry (Griffin and Lichtenstein, 2013). The study disregarded a number of studies with different results (McDougall, 2016).

One dimension of our 2×2 experimental design relates to the number of sellers in a market, which is one in the monopolistic setting M and four in the competitive setting C. In the latter setting, buyers chooseone seller and bid for her product. In the other dimension, the benchmark of ten pieces of evidence is extended by the sellers' option to purchase for a small fee an additional set offive evidences about the product quality. Buyers are aware of the option, but cannot observe whether it was exercised.

Across settings, we find that sellers frequently report only a selected set of evidences which predominantly contains the most favorable evidences. Buyers compensate for that selection, but – in accordance with previous findings of limited sophistication – often to an insufficient degree. Although buyers' correct inferences could impose equality of payoffs, sellers gain more than buyers across all four games. Competition increases transparency throughout, albeit not to the level of full disclosure. Despite the increased transparency, buyers compensate surprisingly little for the selection in the competitive settings. Indeed, they almost undo the potential benefit of competition as they earn on average only slightly more when choosing one seller out of four than when not. The influence of the competitive setting on the buyer's inference is thus as important as the sellers' increased disclosure.

The option to purchase additional information on average leads to more evidences being published. However, it does not change the bias in the published evidence because this increase is mostly due to the larger set of evidences to select from. Further, the buyer's skepticism is increased in the monopoly setup, leading to an average bid that almost coincides with the product value. In the competitive setup, the skepticism increases but does not reduce the gap between the bid and the true product value. In both cases, the investment in the additional evidences is worthwhile for sellers.

Overall, we deliver a detailed account of sellers' selection of favorable evidences and buyers' insufficient compensation for it. Full disclosure is rare since the compensatory buyer behavior favors opacity. We see that competition decreases buyers' skepticism surprisingly strongly, while the uncertainty about the available set of evidences increases it moderately.

Thanks to the nature of our games, we not only observe that skepticism is insufficient in the baseline setting, we further observe that it changes in response to the institutional setting. Our results imply that skepticism is a very useful feature of consumers postures, if applied in the given context correctly. Johnson, Meier and Toubia (2015), for example, study a context in which skepticism is too strong and turns out costly for consumers. Conversely, the persistence of objectively misguided consumer myths arising from business-related information provision, such as the perceived consumption necessity of dairy products for bone strength and for coverage of calcium and protein needs, suggests the existence of areas of insufficient skepticism. Overall, the recent investigations in the topic suggest that consumers have potential to improve the calibration of their skepticism and increase their welfare by dealing better with information that is provided by interested parties.

The chapter is structured as follows. Section 2.2 shows how our study relates and contributes to different strands of literature. Section 2.3 presents our experimental games and discusses theoretical considerations and the related literature. Section 2.4 reports the results before section 2.6 discusses them and concludes.

2.2 Related Literature

In economics, the understanding of the negative consequences of information asymmetries motivates the investigation of information disclosure (Akerlof, 1970; Viscusi, 1978). In particular, the potential of voluntary disclosure of information is investigated in order to evaluate the need for mandatory disclosure policies (Grossman and Hart, 1980; Grossman, 1981). The prediction of full voluntary disclosure of verifiable information due to unraveling is important, but also dependent on assumptions such as negligible disclosure costs, sufficient competition, or rationality of the consumer (Milgrom, 1981; Milgrom and Roberts, 1986).[3] The endogenous provision of evidence is a characteristic feature of the large class of persuasion games that have been studied extensively (Glazer and Rubinstein, 2001, 2004, 2006; Kamenica and Gentzkow, 2011)

Apart from reasons for incomplete disclosure such as the large amounts of information available to the consumer and strategic considerations not to disclose information (Dranove and Jin, 2010), recent studies have investigated the sophistication of consumers in more detail. Brown, Camerer and Lovallo (2012, 2013) show that the preventing of movie previews by critics is associated with lower movie quality, a relationship which is possibly underestimated by customers.

Experimental studies such as ours allow to cleanly study simplified situations in order to understand the underlying mechanism more deeply. Forsythe, Isaac and Palfrey (1989) study information disclosure in blind-bid auctions and see that – as theory predicts – all sellers eventually disclose the true value of their product. The auction format and the consequent strategic considerations make it difficult to clearly identify the role and extent of bidders' sophistication.

In a recent study, Jin, Luca and Martin (2016) experimentally implement a simple stylized disclosure game similar to the original setup in Milgrom and Roberts (1986) which is suitable to cleanly investigate buyers' sophistication. They find that unraveling is incomplete due to a lack of skepticism in buyers, which persists over a large number of rounds and under the provision of detailed feedback. This study is the most detailed investigation of buyer behavior in this standard setting. In slightly different settings, two recent studies support the finding of limited sophistication on the buyer's side. Hagenbach and Perez-Richet (2015) experimentally investigate disclosure when types are not monotonically ordered, a case considered theoretically in Hagenbach, Koessler and Perez-Richet (2014). Benndorf, Kübler and Normann (2015) consider disclosure of information in labor markets with the focus on privacy concerns.

[3] The large theoretical and experimental literature on cheap talk considers situations in which statements are not bound to relate to the true seller type (Crawford and Sobel, 1982; Cai and Wang, 2006; Wang, Spezio and Camerer, 2010).

In our study, we extend this literature by proposing a simple and flexible market framework of information disclosure which is directly informative about seller and buyer behavior. Thanks to its flexibility, we can investigate behavior across market and information structures and therefore uncover whether the sophistication of buyers' and its anticipation by sellers changes between them.

Milgrom and Roberts (1986) show that the skepticism of buyers leads to unraveling even in monopolistic markets. Further, competition among sellers with strongly opposed interests is able to compensate for lacking skepticism on the buyer side by the existence of incentives to reveal the truth about products. Empirical studies support the view that competition matters, although it is not clear in which direction. Jin (2005) studies health organizations and shows that competition gives stronger incentives to differentiate in disclosure decisions. Burks, Cuny, Gerakos and Granja (2015) find an increase in competition between banks to raise the level of voluntary disclosure and in particular of negative information.

Our experiment allows us to cleanly observe how competition changes transparency. So far, there is little understanding of whether and how the subjective perception of competition influences beliefs and sophistication of the consumer (Huck and Zhou, 2011).

The possibility of extending the set of evidences relates our study to the literature on strategic experimentation and the trade-off between benefits from exploration and exploitation in bandit problems (Rothschild, 1974; Aghion, Bolton, Harris and Jullien, 1991). Here, the trade-off is simpler since the seller only has to weigh the fixed costs of additional evidence against the transparency and selection benefits of more evidences. In the persuasion context, the buyer's uncertainty with respect to the seller's pool of evidences increases the burden on the evidences and makes partial communication less informative. In a model with an exogenously given set of evidences, Dziuda (2011) shows that persuasion is feasible if the persuader has a high number of favorable evidences. Felgenhauer and Schulte (2014) find the same with private acquisition of new evidences and show that a limit on the number of evidences acquired makes persuasion impossible if the seller's stakes are too high.

Our experiment illuminates sellers' propensity to extend the evidence set and whether this results in increased transparency and/or stronger selection of evidences. Furthermore, we show whether and to which extent the uncertainty increases the skepticism with which evidences are received.

Finally, the determinants and the right level of skepticism is an increasingly covered topic in economics (Johnson et al., 2015). Burdea, Montero and Sefton (2016) show in an experimental study of disclosure of partially verifiable information that receivers' skepticism decreases when they have more control about verification, holding constant the amount of and selection of information revealed. Due to the importance of unbiased information for consumer welfare, we believe that the empirical study of skepticism and its determinants is a topic of high importance.

2.3 Experimental Games

We investigate the effects of competition and of additional disclosable information in a 2×2 experimental design as shown in Table 2.1. In the first dimension, we implement two market structures, monopoly and competition. In the second dimension, on top of the standard number of ten evidences, the option to purchase five additional evidences exists.

		Monopoly M	Competition C
10 Evidences	(X10)	M10	C10
10+5 Evidences	(X10+5)	M10+5	C10+5

Table 2.1: 2×2 experimental design.

Monopoly (M10) The most basic game features one seller and one buyer. The seller offers a good of value v, which is uniform-randomly drawn from $[200, 1000]$. While this distribution is common knowledge, the realization is private information to the seller. The seller cannot communicate the true value of the good to the buyer, but he receives a disclosable set E of ten informative but noisy evidences e_i. The evidences are normally distributed with a standard deviation of $\sigma = 100$ and a mean of $\mu = v$. The number and distribution of evidences is common knowledge. The seller decides which of those evidences, if any, to report to the buyer and thus determines a message $M \subseteq E$. Due to the verifiability of information, the seller cannot change or manipulate the evidences' values. Evidences outside of M remain his private information.

The buyer observes M and bids $b \in [0, 1200]$ to buy one unit of the seller's good. The design of the price mechanism is equivalent to a Becker, DeGroot and Marschak (1964) mechanism: The price for one unit of the good, $p > 0$, is uniform-randomly drawn from the interval $[v-200, v+200]$. p is disclosed only after the buyer places her bid. If $b < p$, the transaction does not take place, leaving both parties with a payoff of 0. If $b \geq p$, the transaction takes place and the buyer gets the value v for the price p. The seller obtains the bid b and incurs costs of $c = v - 50$ upon production of one unit.[4]

In summary, the seller's payoff is

$$\Pi_S(b, p, c) = \begin{cases} (b-c) & \text{if } b \geq p, \\ 0 & \text{otherwise.} \end{cases}$$

and the buyer's payoff is

$$\Pi_B(b, p, v) = \begin{cases} (v-p) & \text{if } b \geq p, \\ 0 & \text{otherwise.} \end{cases}$$

[4]For examples of possible outcomes see Appendix A.1.

It can be shown that – similar to a second price auction – it is a dominant strategy for the buyer to bid what she believes to be the true value of the good. Therefore, this rich setup cleanly implements the natural conflict of interest between the seller that wants the buyer's bid to be as high as possible and the buyer that wants her evaluation of the good to be as accurate as possible.

Competition (C10) The competition game features four sellers and four buyers in each market. Sellers are denoted by $j \in \{A, B, C, D\}$. Each seller offers a good with a value v_j that is independently drawn and private information as before. All aspects of E remain the same. The decisions of each seller to determine $M \subseteq E$ are made simultaneously. Subsequently, the buyers as well as the sellers observe the published evidences of all four sellers.

In contrast to before, buyers choose from which seller to buy before they place their bid b for the chosen seller's product. Prices p_{j^*} and payoffs are determined in the same fashion as before for the transaction between the buyer and her chosen seller j^*. There is no competition among the buyers because a seller can sell up to four units of his good.

Note that the absolute magnitude of v_j is not important to the buyer due to the price mediation. In that sense, the setting is symmetric across sellers. A rational buyer should always decide to purchase from the seller whose good's true value she can estimate most accurately. All sellers have thus the same incentives to reveal evidences.

10+5 Evidences (M10+5 and C10+5) In the X10+5 games, the possibility of purchasing additional evidences is introduced as sellers can – in addition to the initial set of ten evidences – purchase five more evidences for a price of $P_e = 15$ points. The additional evidences are independently distributed in the same way as the initial evidences. Sellers make the purchasing decision after they observe their initial set of ten evidences. Buyers know that sellers have the possibility to purchase these additional five evidences but cannot observe a seller's purchase. Apart from the price P_e which sellers pay upon purchase independently of any transaction, the payoff structure is not different from before.

In the X10+5 games, sellers are given the opportunity to disclose more information to the benefit of the accuracy of the buyers' inference. Depending on the competitive pressure and the price P_e, sellers potentially purchase and disclose additional information. Alternatively, the additional information might lead to stronger selection of disclosed information with buyers not being able to infer whether a purchase was made.

2.3.1 Theoretical considerations

Theoretical background

In order to create a multifaceted and versatile framework of information disclosure, we implement a rich setting with multiple normal signals and the usual conflict of interest between the two market sides. Many of the usual intuitions about voluntary information disclosure as discussed in the literature hold to be true here. Some aspects of the context are more unusual and will deserve particular mentioning.

Using a price mechanism with an exogenous price avoids the introduction of any strategic consideration via an endogenous price determination mechanism. Still, the price levels the field between different seller types, putting emphasis on the amount of information available. For example, a low-v product can be more interesting to the buyer than a high-v product if its value can be more accurately assessed. Therefore, in competition, firms compete on a level playing field despite their different v's.

Traditionally, the Becker-deGroot-Marschak mechanism is an incentive-compatible mechanism mostly used to elicit willingness-to-pay (WTP). If a buyer holds and knows his WTP, the mechanism gives incentives to truthfully reveal it. In our case, the true value v is subject to uncertainty from the point of view of the buyer, so that risk-averse buyers are expected to underbid relative to their expected value (Kaas and Ruprecht, 2006).

Compared to other market settings, the price mechanism in form of the Becker-deGroot-Marschak mechanism provides the buyer with incentives to reveal her willingness to pay and thus her precise inferred expected value of the product. Since the reaction to this mechanism is not subject of our investigation, we explain participants that the bidding of their expected value is optimal.[5]

The buyers' way of inferring the true value of the item is important for our study as it determines in large parts the equilibrium outcome. The early literature of Milgrom (1981) and Milgrom and Roberts (1986) emphasizes the importance of buyer sophistication or, in other words, skepticism for the information revelation in the monopolistic setting. In our setting, skepticism in the face of partial disclosure implies the belief that the disclosed information is the most favorable from the seller's perspective. The nature of the evidences further implies that any disclosed information is always useful and allows to form a posterior belief about the true value. The more information is disclosed, the more accurate the belief will be. The possibility that the seller has preferred to reveal k rather than fewer pieces of information because of k particularly favorable realizations might induce the buyer to further reduce the inferred value. In the face of the skeptical buyer's markdown applied to the nominal value of partially disclosed information, the seller maximizes the buyer's bid by revealing all pieces of information if they are not too extreme.

By the nature of the normal draws, extremely low or high realizations of E occur with very small probability. Very low realizations might reduce the chance of a profitable transaction so much that the seller prefers to not disclose or to only disclose low realizations, practically refraining from the initiation of a transaction. Conversely, very high realizations might lead to partial disclosure if the corresponding markdown is believed to be small compared to the position of the outliers.

[5]The instructions state: "Note that in terms of expected payoff, the Buyer benefits most from bidding what he believes the Value of the item is."

As in other settings, the partially disclosing seller always benefits from a less skeptical buyer who might have a different belief about the ranking of the disclosed information in E or who does not measure well the extent to which the disclosure depends on the realizations. In this case, full revelation of information will not be achieved and the seller is expected to make an extra profit at the expense of the buyer.

We implement a setting which does not exhibit "strongly opposed interests" as modelled in Milgrom and Roberts (1986) via an additional pricing stage. All disclosure incentives are derived from the relative amount of disclosure. Intuitively, this setting is closer to the more commonly observed nature of competition since companies do not frequently reveal information about competitors' products. More importantly, it is interesting to quantify the effect of moderate competition. If it leads to full transparency this is an interesting insight. If it does not, the observation of whether and how buyer sophistication changes in the face of competition with incomplete transparency is equally interesting.

Theoretical predictions

In order to fix ideas about behavior in the experiment, we sketch some theoretical considerations. For simplicity, we first consider a setup in which the seller decides about the message M prior to observing the realization of E. Later, we evaluate what the treatment effects are when the choice is made under private information of E.

Let the belief β specify which evidences M are disclosed from E. The message $M \subseteq E$ induces a posterior belief distribution $f(v|M, \beta)$ and an expected value $\mathbb{E}(v|M, \beta)$.

A *skeptical* buyer holds a belief $\beta = s$ that the seller discloses the highest $\#M$ evidences in E, inducing $f(v|M, s)$ and $\mathbb{E}(v|M, s)$. Analogous to Milgrom and Roberts's (1986) skeptical posture, in the face of a seller that benefits from a higher bid, this belief minimizes the buyer's bid for a given message M. In relation to the average observed evidence in M, $\bar{e}^M = \sum_{i:e_i \in M} e_i$, the bid features a skepticism-induced markdown $S_{\#E}(M) = \bar{e}^M - \mathbb{E}(v|M, s)$. As $\#M$ increases, the inclusion of evidences closer to v leads to a markdown reduction of $\Delta S_{\#E}(\#M) \equiv S_{\#E}(\#M+1) - S_{\#E}(\#M) < 0$. This markdown is zero when all evidences are disclosed and hence no selection is implemented, $S_{\#E}(M = E) = 0$. Given $\#M$, the markdown and its derivative increase in the cardinality of the set of evidences $\#E$ to select from.

A *naïve* buyer holds the belief $\beta = n$ that M is the complete evidence available to the seller and not selected from E. Such a buyer takes M at face-value and forms an expected value $\mathbb{E}(v|M, n) = \bar{e}^M$. The resulting bid therefore does not include a skepticism-induced markdown.

Irrespective of the belief β, due to the remaining uncertainty about the true value v, a risk-averse buyer maximizes expected utility with a bid b^* that is below $\mathbb{E}(v|M, \beta)$ and that generates a risk-induced markdown $R(M) = \mathbb{E}(v|M, \beta) - b^*$ (Kaas and Ruprecht, 2006).

Because more pieces of evidence reduce uncertainty about $\mathbb{E}(v|M, \beta)$, b^* increases with higher $\#M$ and the markdown reduces $\Delta R(\#M) \equiv R(\#M+1) - R(\#M) < 0$. The absolute level of expected utility increases in $\#M$ due to the decrease in risk.

Proposition 2.1 (M10 – Sophisticated buyer) *If the risk-averse buyer is skeptical, the seller will disclose all evidences, $M = E$.*

Proof. Since a skeptical buyer is accounting for the selection of evidences, the seller can only influence his risk-induced markdown and minimize it by giving as much information as possible, $M = E$. □

In this setup, the assumption of risk-aversion leads to incentives of full revelation. A risk-neutral, skeptical buyer would be indifferent between various strengths of selection and many seller strategies could be sustained in equilibrium.

Proposition 2.2 (M10 – Sophisticated and naive buyers) *If the risk-averse buyer is ve with probability ν and skeptical with probability $1 - \nu$, the seller will disclose a number* *of evidences $\#M^*$ such that the incremental reduction of the markdown $\Delta R(\#M^*)$ is just not smaller than the expected gain from the naive's overbidding, $\nu \cdot \Delta S_{10}(\#M)$.*

Proof. We restrict attention to seller strategies that disclose the top $\#M$ evidences because these are the most beneficial strategies in the face of naïve buyers. The seller chooses $\#M$ such that his profit is maximized.

$$\begin{aligned}
\max_{\#M \leq 10} \pi(\#M) = \mathbb{E}(b) - c &= (1-\nu) \cdot (\mathbb{E}(v|M,s) - R(\#M)) && (2.1)\\
&+ \nu \cdot (\mathbb{E}(v|M,s) - R(\#M) + S_{10}(\#M)) - c && (2.2)\\
&= (\mathbb{E}(v|M,s) - R(\#M)) + \nu \cdot S_{10}(\#M) - c && (2.3)
\end{aligned}$$

□

Intuitively, the seller will publish a number of evidences $\#M^*$ up to the point were the marginal or incremental cost from reducing the naive's overbidding is equal to the incremental benefit of reducing uncertainty. An optimal number $\#M^* > 1$ is obtained when the risk-aversion is such that the risk-induced markdown that applies to both types, $\Delta R(1)$ outweighs the seller's benefits from naïve bidding $\nu \Delta S_{10}(1)$. Similarly, since for large messages the reduction in uncertainty $\Delta R(9)$ is relatively small, a substantial share of ν will induce $\#M^* < 10$. Of course, extreme choices of $\#M = 1$ or $\#M = 10$ could still occur for extreme enough risk preferences that determine R.

Proposition 2.3 (M15 – Sophisticated buyer) *When the seller can purchase five further pieces of evidence, he does so only if the resulting change in R is greater than the price of 15 points. After a purchase, all 15 pieces of evidences are disclosed.*

Proof. Due to the sophistication of the buyer, the disclosed information only impacts the risk-induced markdown. If the buyer is sufficiently risk-averse that the increase to 15 evidences more than compensates the cost of the additional five pieces of evidence, the seller benefits from the purchase. □

Proposition 2.4 (M15 – Sophisticated and naive buyers) *If the risk-averse buyer is naive with probability ν and skeptical with probability $1-\nu$ and $\frac{R'}{S'_{15}} < \frac{R''}{S''_{15}}$, the seller will disclose a number of evidences $\#M^*$ such that $\nu = \frac{R'}{S'_{15}}$.*

Proof. In M15, the seller can choose between purchasing and not purchasing additional evidence. Not purchasing additional information would re-establish the case of M10 considered in proposition 2.2.

If the seller purchases additional evidences, then sophisticates know in equilibrium that the disclosed evidence is featuring a stronger publication bias than without the additional evidences. The naive buyer, however, still acts as if $\#M = \#E$.

Conceptually, the seller's choice of $\#M^*$ is isomorph to proposition 2. Generally, $\#E = 15$ opens the possibility to reduce uncertainty further than under $\#E = 10$. For the first ten evidences, the marginal risk-induced markdown is as before and decreases to lower levels as $\#M > 10$.The publication bias for a given $\#M$ is much higher now as is the marginal skepticism-induced markdown.

$$\max_{\#M \leq 15} \pi(\#M) = \nu \cdot S_{15}(\#M) + (\mathbb{E}(v|M) - R(\#M)) - c \tag{2.4}$$

$$\text{FOC:} \quad \nu \cdot S'_{15}(\#M^*) = R'(\#M^*) \tag{2.5}$$

$$\text{SOC:} \quad \nu \cdot S''_{15}(\#M^*) < R''(\#M^*) \tag{2.6}$$

It follows that the resulting optimal message size is therefore higher than in M10. Whether or not the resulting higher profit for the seller justifies the payment for the additional pieces of evidence determines then whether or not the seller chooses to purchase or not. In any case, an equilibrium in which the seller purchases features a higher gross profit (before fee) than in M10.

□

In our experiment, sellers observe their evidences before deciding about M, making an exhaustive analysis very complex. Sellers might select into disclosing $\#M$ on the basis of the realized evidences. For example, $S_{\#E}(M = E)$ could be non-zero and positive when full disclosure only occurs with "good" realizations of E. Indeed, the result of full disclosure from proposition 1 does not necessarily hold here. For example, Shin (1994) shows that full revelation fails in a simpler context when there is uncertainty regarding the informed party's information. For example, E might realize in a fashion that the seller prefers not to put out the whole E.

Importantly for our experimental treatments, as before, the incentives for disclosure decrease when naives enter the population of buyers. Compared to the equilibrium with only sophisticated buyers, we should expect less disclosure in an equilibrium when naives are present. The lack of full disclosure in the monopoly setting can then only be probabilistically attributed to the presence of naive buyers.

Due to the assumed risk-aversion, we can show that under competition, full disclosure is reestablished as equilibrium.

Proposition 2.5 (C10) *Irrespective of the buyer types, under competition,* $\#M = 10$ *and sellers derive the expected profit* $(\mathbb{E}(v|E) - c)$.

Proof. Due to the preference of all risk-averse buyers for products with a lot of information and a resulting low level of risk they choose among the four products in the market the product with the highest $\#M$. Rather than not being chosen by any buyer for a transaction, the sellers discloses $\#M = 10$ and – in the face of four buyers – will be chosen in expectation by 1 consumer for a transaction. □

2.3.2 Experimental Procedures

We conducted the experiments in the Experimental Economics Laboratory at the University of Mannheim (mLab).[6] Across ten sessions, a total of 160 students partici-

[6]The experiment was programmed and conducted with the software z-Tree Fischbacher (2007) and subjects were recruited with ORSEE (Greiner, 2004).

pated. We recruited all students from the general student population of the University of Mannheim.

At the beginning of each session, the instructions are read out aloud and explained to all participants. Subsequently, subjects have three attempts to complete an understanding test of four comprehension questions about the payoff structure. Those who fail get an individual short briefing on the points they did not understand. Finally, subjects have three minutes to generate draws from normal distributions with standard deviation of 100 in order to get familiar with the kind of draws occurring in the experiment.[7]

Participants play two of the four different games sequentially in each session as shown in Table 2.2. To control for order effects, we vary the order of the games between sessions. Apart from the pilot sessions one and two, each game is played twice as game one and game two of the session, respectively, following a latin square design.

Each game is played for ten rounds with two unpaid practice rounds before the start. Participants are randomly matched into markets and randomly assigned the role of seller or buyer. While their market counterparts change randomly in each round, participants keep their role of seller or buyer for five rounds. For the last five rounds, roles are switched and maintained until the end, while counterparts keep changing.

Session	Game 1	Game 2	Session	Game 1	Game 2
1	M10	C10	6	C10	C10+5
2	C10	M10	7	M10	M10+5
3	C10	M10	8	C10+5	C10
4	M10+5	C10+5	9	M10	C10
5	C10+5	M10+5	10	M10+5	M10

Table 2.2: Games in 10 sessions

At the end of each round, the feedback in M games consists of the true item value, the price, the bid, the realization of a transaction and own payoffs. In C games, sellers are

[7] Appendix A.3 provides screenshots of the games and of this tool.

further informed about the published messages of all other sellers, their own number of realized transactions and the total amount bid for their good.[8]

Participants are compensated on the basis of the outcomes in the 20 paid rounds. Individual payoffs are in points and are converted to cash at an exchange rate of 1 EUR for 60 points.[9] The average payoff per subject was 10 EUR. Since payoffs can be zero and even negative in a given round, we established a minimum payoff for the session of 2 EUR which was not binding for any participant.

2.4 Results

2.4.1 The effect of competition

A basic indicator of the effect of competition is the number of evidences that sellers disclose, $\#M$. Figure 2.1a shows that this number is fairly symmetrically distributed around four and five and shifts slightly right in competition, with a pronounced spike at full disclosure of ten evidences. Figure 2.1b shows a slightly stronger difference in later rounds 5+10, suggesting that the competitive pressure leads to even more disclosure over time.[10]

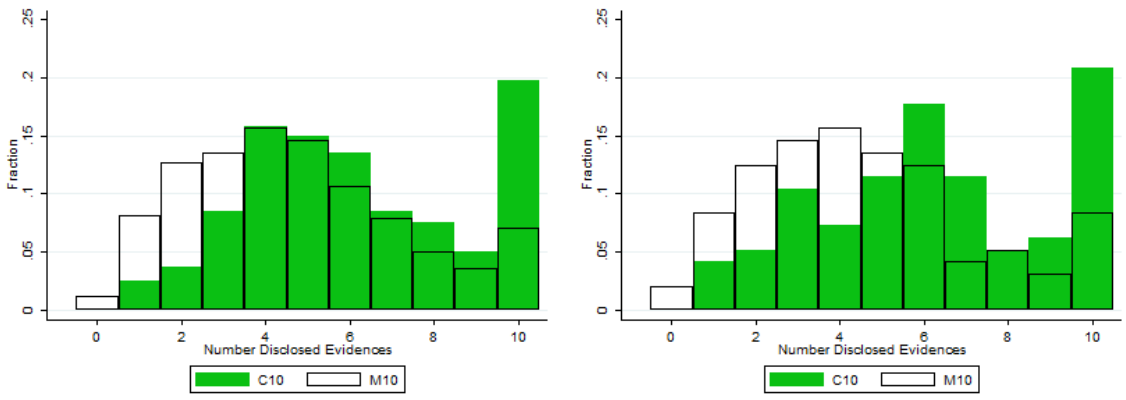
Throughout games and rounds, panel (a) of Table 2.3 shows significant differences in the average number of disclosed evidences between M and C. In M10, over all rounds, 1.47 pieces of evidences less are disclosed compared to C10. The difference is 1.60 in the last rounds of being a seller.

The selection of evidences leads to a positive "publication bias" of $\bar{e}^M - v$ if the mean of the disclosed evidences exceeds the true value of the good. Panel (b) of Table 2.3 shows that the reduced disclosure in M leads to a higher publication bias than in

[8] In the two pilot sessions, sellers were not able to see what the other sellers decided to report.

[9] All values are restricted to integers and given in experimental points.

[10] By considering the 5th and the 10th round we can check for possible learning effects over time and we make sure that each participant is only accounted for once in each role.



(a) All rounds. (b) Rounds 5+10.

Figure 2.1: Number of disclosed evidences, $\#M$.

		All rounds			Rounds 5+10		
		M		C	M		C
(a)	10	4.71	***	6.18	4.65	***	6.25
Nr Published evidences		***		***	[0.24]		**
$\#M$	10+5	5.31	***	6.91	5.16	***	7.38
(b)	10	75.05	***	45.61	82.83	***	42.21
Publication bias		[0.94]		[0.49]	[0.59]		[0.49]
$\bar{e}^M - v$	10+5	75.40	***	43.15	77.35	***	47.24
N	10	480		480	96		96
	10+5	320		320	64		64

Notes: t-tests for equal means, significance level indicated by: *** p <0.01, ** p <0.05, * $p \leq 0.1$, [$p > 0.1$].

Table 2.3: Means of number of disclosed evidences and publication bias

C, as expected when sellers publish the highest evidences. This difference is significant across games and even more pronounced in the last rounds between M10 and C10. Overall, the data show that sellers' disclosure in monopoly is quite low but improves significantly due to competition.

The publication bias and sellers' decisions of which evidences to disclose can be analyzed in further detail when reporting it by the number of published evidences $\#M$ as in Table 2.4. Panel (a) gives benchmark magnitudes from simulations of 10000 draws of ten evidences. The first line indicates the mean difference to the true value for the

kth ranked draw, $e_k - v$, and the second indicates the mean publication bias for the k highest draws, $\frac{1}{k}\sum_{i\in\{1,\ldots,k\}} e_i - v$.[11]

Panel (b) reports the observed publication bias overall (X10) as well as split up between M10 and C10. Across games, for $k \leq 6$, the magnitudes of the publication bias are below the levels theoretically expected under a "top k evidences" disclosure strategy, and above for $k > 6$. This suggests that for $k \leq 6$, sometimes evidences other than the top k are disclosed, possibly to avoid a transaction after a low realization of evidences. The numbers for $k > 6$ witness a selection into disclosing many evidences as a result of favorable evidence realizations. Otherwise, for example, a publication bias of 12.5 for $k = 10$ in X10 is statistically very unlikely. Furthermore, with one exception, the bias is slightly smaller in C10 than in M10, suggesting that evidences other than the top k are disclosed more often in C10. Similar patterns can be observed for X10+5 in Table 2.11 on page 33.

To better understand the sellers' strategies, we can take a closer, exemplary look behind these averages. For $k = 1$, consider the publication bias of 141.3 in M10 and the significantly lower value of only 40.6 in C10. Figure 2.2a shows that almost all sellers in M10 disclose their highest evidence while less than 40% of sellers do so in C10 (2.2b). The latter figure shows that further strategies used may be to disclose the single evidence that is closest to the true product value or to disclose the lowest evidence to avoid a loss-making transaction. The smaller publication bias in C suggests that competition increases the incidence of such strategies.

The buyer now observes the disclosed evidences with a mean of $\bar{e}^M$ and – in order to bid approximately the true value v – should account for the selection by bidding this mean minus a "markdown", $\bar{e}^M - b$. For correct inferences of v, the buyer's markdown and the seller's publication bias should be the same in absolute value. In C, we always refer in the following to the evidences of the seller j^* chosen by the buyer

[11] These calculations give a good indication of the expected publication bias, but do not take into account the interval boundaries of v, which influence the rational inference in the experiment. Simulations in Appendix A.2 show that the boundary only matters in a limited range above the lower interval boundary.

		#M / Rank k										
		10	9	8	7	6	5	4	3	2	1	0
(a)	k	-154.2	-100.4	-65.8	-37.8	-12.5	12.0	37.5	65.5	100.1	153.7	
Mean	Top k	-0.2	16.9	31.6	45.5	59.4	73.8	89.2	106.4	126.9	153.7	
(b) Pub. bias	X10	⁝12.5	⁝25.8	⁝43.7	⁝50.6	46.8⁝	59.9⁝	70.8⁝	81.4⁝	113.0⁝	117.6⁝	
	N	129	41	60	79	116	142	151	106	79	51	6
	M10	⁝10.4	⁝31.6	⁝53.2	⁝57.1	52.4⁝	65.0⁝	72.7⁝	90.2⁝	117.8·	141.3	
	N	34	17	24	38	51	70	75	65	61	39	6
	C10	⁝13.2	21.7	∶37.3	44.6	42.5⁝	54.9⁝	69.0⁝	67.3⁝	96.5⁝	40.6⁝	
	N	95	24	36	41	65	72	76	41	18	12	0
(c) Mark-down	X10	12.3	13.3	21.9	31.6	35.2	30.7	40.7	58.9	96.0	162.2	
	N	208	38	59	85	104	130	133	87	69	41	6
	M10	39.8	16.8	29.5	35.1	45.3	40.2	59.8	75.1	116.2	170.4	
	N	34	17	24	38	51	70	75	65	61	39	6
	C10	7.0	10.5	16.7	28.8	25.4	19.6	16.1	11.0	-57.8	1.8	
	N	174	21	35	47	53	60	58	22	8	2	0
(d)	C10 Chosen	9.8	26.4	40.4	44.2	51.2	56.1	73.5	64.0	140.1	45.3	
Pub. bias	sellers	174	21	35	47	53	60	58	22	8	2	0

Notes: Average publication bias being ⁝above (below⁝) the simulated confidence interval of N Top k draws is denoted with Maya numerals ⁝ (99%), ∶ (95%), and · (90%) to the left (right) of the average.

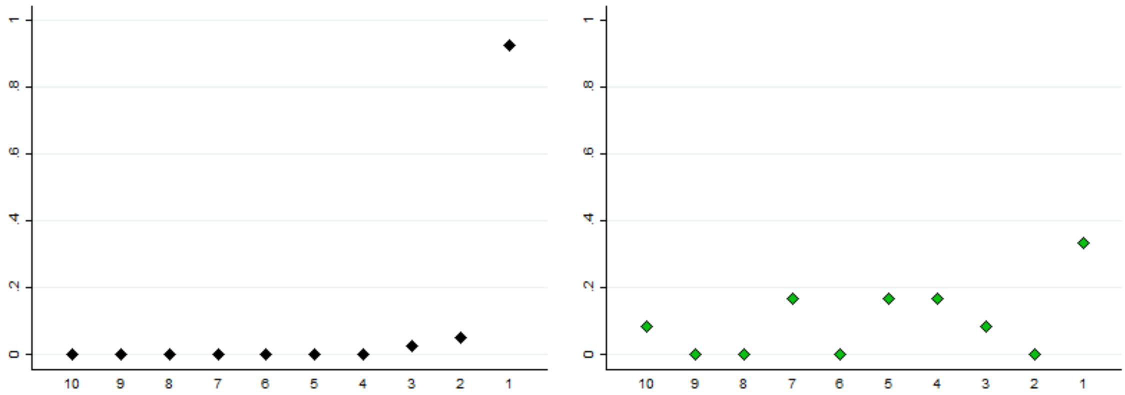
Table 2.4: Summary by number #M and rank k of disclosed evidences (X10 games)

for a transaction. The publication bias of those chosen sellers can be calculated as markdown + precision and differs little from the publication bias across all sellers in Table 2.3.

As expected, panel (a) of Table 2.5 shows that the markdown is lower in competition than in monopoly. Surprisingly, the magnitude of the markdown in competition is very small: in C10 it is only 21% of the publication bias, suggesting that buyers are much less skeptical than in monopoly. We can judge the appropriateness of the markdown by calculating the precision of the bid, $b-v$. Panel (b) of Table 2.5 shows some overbidding in M10 but five times the overbidding in C10, the absolute difference remaining stable over time.[12]

This difference is maintained when conditioning on the observed number of evidences. Panel (c) of Table 2.4 shows the markdown by the number of disclosed evidences and

[12] In the M treaments, the markdown and the precision do not precisely add up to the publication bias due to six (M10) and two (M10+5) sellers that do not disclose any evidence and for which publication bias and markdown cannot – in contract to precision – be calculated.



(a) M10 ($N = 39$). (b) C10 ($N = 12$).

Figure 2.2: Fraction of disclosure for k-ranked evidences.

		All rounds			Rounds 5+10		
		M		C	M		C
(a)	10	67.30	***	13.78	67.68	***	9.60
Markdown		[0.46]		[0.42]	[0.77]		[0.63]
$\bar{e}^M - b$	10+5	73.23	***	19.28	72.67	***	16.44
$\neq 0$	10	***		***	***		[0.33]
	10+5	***		***	***		*
(b)	10	5.63	***	25.22	8.13	[0.22]	26.43
Precision		[0.63]		[0.41]	[0.71]		[0.90]
$b - v$	10+5	1.63	**	19.14	1.30	[0.15]	28.31
$\neq 0$	10	[0.32]		***	[0.45]		***
	10+5	[0.79]		***	[0.93]		**

Notes: t-tests for equal means, significance level indicated by: *** p <0.01, ** p <0.05, * $p \leq$0.1, [$p > 0.1$].

Table 2.5: Buyers' markdown and bidding precision

by M10 and C10. Panel (d) reports the publication bias of sellers j^* chosen by the buyers. The markdown is generally smaller in magnitude than the publication bias. In particular, the minor markdowns even for small $\#M$ are surprising and contribute to the low precision in C10. A very similar picture emerges for the 10+5 settings in panels (e) and (f) of Table 2.11 on page 33.

The buyers' overbidding is reflected in an elevated transaction probability and in relatively low payoffs. Panel (a) of Table 2.7 shows that transaction probabilities are mostly above 0.5, significantly so in M10 and C10. Throughout all games, the sellers'

		Rounds 1-5			Rounds 6-10		
		M		C	M		C
(a)	10	56.59	***	4.23	78.18	***	23.33
Markdown		[0.75]		[0.18]	[0.19]		[0.83]
$\bar{e}^M - b$	10+5	53.11	***	17.34	93.61	***	21.23
$\neq 0$	10	***		[0.54]	***		***
	10+5	***		***	***		***
(b)	10	17.45	[0.08]	35.62	-6.20	[0.06]	14.83
Precision		[0.35]		[0.30]	[0.14]		[0.90]
$b - v$	10+5	28.01	[0.76]	24.73	-24.74	***	13.55
$\neq 0$	10	**		***	[0.47]		***
	10+5	***		***	**		[0.08]

Notes: t-tests for equal means, significance level indicated by: *** p <0.01, ** p <0.05, * $p \leq$0.1, [$p >$ 0.1].

Table 2.6: Buyers' markdown and bidding precision

payoffs are higher than the buyers' payoffs as seen in panel (b). Importantly, buyers do not benefit from competition as their payoff in C10 is almost the same as in M10.

		M10	M10+5	C10	C10+5
(a) Transaction probability	All rounds	0.54	0.53	0.54	0.52
	$\neq 0.5$	*	[0.32]	***	[0.21]
	N	480	320	480	320
	Rounds 5+10	0.60	0.55	0.59	0.53
	$\neq 0.5$	**	[0.46]	***	[0.38]
	N	96	64	96	64
(b) Payoff All rounds	Seller	58.73	46.53	59.26	50.08
	Δ	18.24	4.45	16.56	8.33
	$\Delta \neq 0$	***	[0.47]	**	[0.23]
	Buyer	40.49	42.08	42.70	41.75
	Seller, gross		51.31		55.23
	Δ		9.23		13.48
	$\Delta \neq 0$		[0.14]		*

Notes: t-tests for equal means, significance level indicated by: *** p <0.01, ** p <0.05, * $p \leq$0.1, [$p >$ 0.1]. "Gross" refers to the payoffs before cost of purchasing additional evidences is subtracted.

Table 2.7: Transaction probability and payoffs

In addition to the bidding behavior, in competition, the buyers' choice of seller can discipline the seller. Table 2.8 reports the difference in disclosure characteristics x_{j^*} between the chosen seller and the other three sellers $\frac{1}{3}\sum_{j\neq j^*} x_j$. The chosen seller on average reports 1.51 pieces of evidence more than the average competing seller.

A higher number of disclosed evidence thus increases the seller's probability of being selected. The mean and median of the disclosed evidence of the chosen seller are relatively higher in C10, reflecting that their nominal magnitude has a positive and significant effect on the buyer's choice of the seller. That this effect is not significant in C10+5 might be due to the stronger reliance on differences in $\#M_j$. The standard deviation of disclosed evidence is slightly higher for the chosen sellers, an effect that is fully attributable to their higher number of evidences.

Characteristic x	C	C10	C10+5
$\#M_j$	1.51***	1.40***	1.67***
$\bar{e}^M$	32.24***	49.17***	6.83
SD M_j	7.31***	7.15***	7.55***
N	800	480	320

Table 2.8: Numbers for chosen seller relative to other three sellers

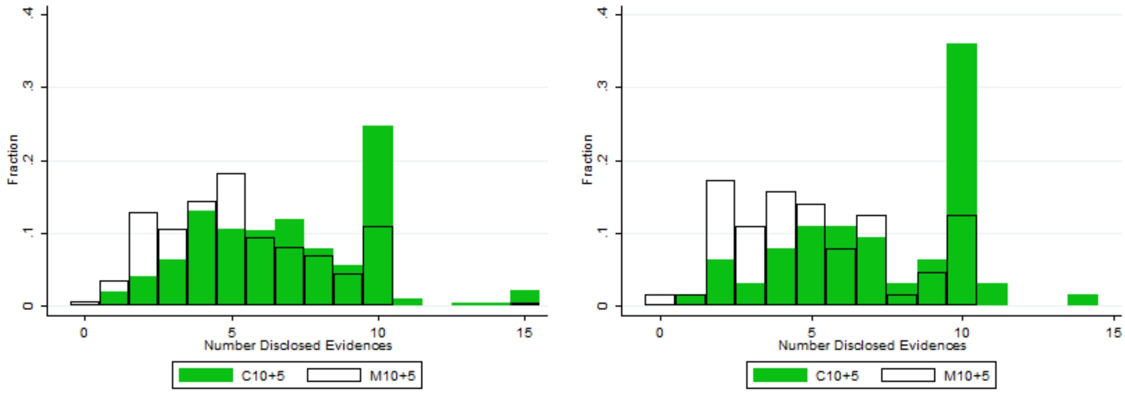
2.4.2 The effect of purchasing additional information

When sellers have the option to purchase additional evidences, panel (a) of Table 2.9 indicates that roughly one-third of the time they do so. Competitive pressure in C10+5 increases this fraction slightly, but not significantly. In the last rounds, this fraction increases throughout, an effect that can be attributed to the payoff benefits that purchasing brings. Panel (b) shows that – across settings – purchasing (+5) leads to a higher payoff than not purchasing further evidences (+0), a difference that even increases over time.

			10+5	M10+5	C10+5
(a) Purchase +5 (%)	All rounds		33.13	31.88	34.38
				[0.56]	
	N		640	320	320
	Rounds 5+10		39.84	35.94	43.75
				[0.47]	
	N		128	64	64
(b) Payoff	All rounds	+5	62.16	53.99	69.74
		$\Delta \neq 0$	***	[0.28]	**
		+0	41.44	43.04	39.78
	Rounds 5+10	+5	72.04	47.00	90.79
		$\Delta \neq 0$	*	[0.96]	**
		+0	38.19	45.76	29.58

Table 2.9: Sellers' purchasing decisions and average payoffs

Distributions of disclosed evidences are very similar to the ones before with the difference of a more pronounced fraction of disclosure of ten evidences across games and rounds, see Figure 2.3. Throughout, very few sellers reveal that they made a purchase by disclosing more than ten evidences. Panel (a) of Table 2.3 shows average numbers of disclosed evidences in X10+5 games that are throughout roughly 0.5 higher than in the ten benchmark. Interestingly, this happens at a maintained level of publication bias. The previously observed differences between monopoly and competition in X10 are approximately replicated.

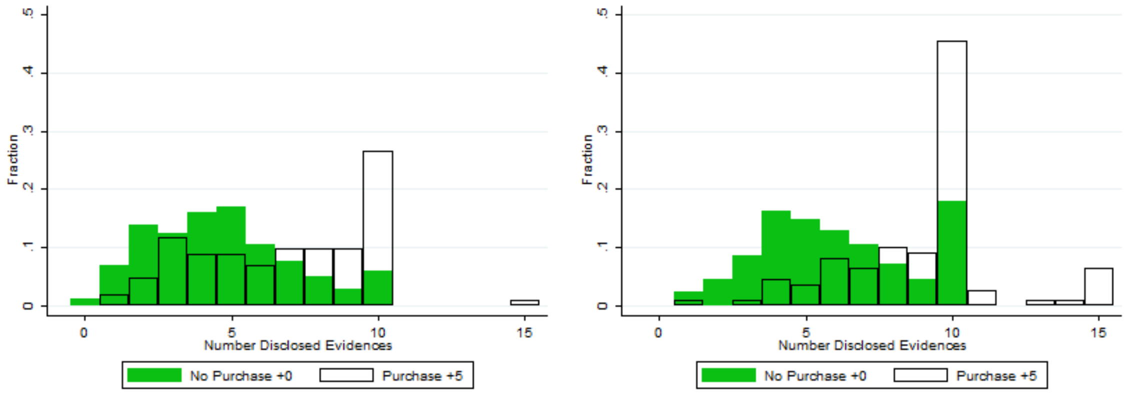


(a) All rounds.

(b) Rounds 5+10.

Figure 2.3: Number of disclosed evidences, $\#M$, in C10+5 and M10+5.

The results so far suggest that purchases were mostly made to enable even more selection. Figure 2.4 supports this view graphically and illustrates disclosure as a function of the purchasing decision. Among purchasers, disclosing ten evidences is most common, by far so in competition. This way they leave the buyer in the unknown about their purchasing decision while still selecting and appearing transparent.



(a) M10+5.

(b) C10+5.

Figure 2.4: Number of disclosed evidence $\#M$ by purchasing decision

The uncertainty about the seller's pool of evidences makes buyers more skeptical, they increase the size of their markdown compared to the benchmark X10 settings, as Table 2.5 shows. While this leads to an almost perfect precision in the M10+5 game, the markdown change in C10+5 is not sufficient to improve the precision by much.

In the last rounds, the precision deteriorates as the markdown – in contrast to the publication bias – is even smaller in absolute terms.

Table 2.7 shows that the overall payoffs of sellers is reduced in 10+5, in terms of both net payoff including the fee and gross payoff. Since the payoffs of purchasing sellers are much higher, this loss comes from non-purchasing sellers. The payoffs of buyers are not changed much and differences between buyers' and sellers' net payoff are not significant anymore.

Fixed-effects panel regressions in Table 2.10 summarize how the behavior of sellers ($\#M$, Pub. bias) and of buyers depend on game characteristics. Importantly, the effect of competition on the number of disclosed evidences and the publication bias is clearly and significantly emerging. Further, the increase in markdown due to competition is larger in absolute terms than the decrease of the publication bias.

Regressions (1) to (4) show that the effect of X10+5 is strongly influenced by the possibility to purchase additional evidences. While X10+5 has a positive, non-significant effect on the number of disclosed evidences, controlling for the purchasing decision in regression (2) reverses this sign. The effect of subjects' role-switching is minor as the buyers' markdown improves only slightly and marginally significantly in rounds 6-10 (p-value=0.12, regression 5). Regressions (1)-(4) further show that a higher value v increases the number of disclosed evidences significantly and reduces the publication bias. At the same time, a higher absolute value of observed evidences $\bar{e}^M$ decreases the markdown (regression 5).[13] Further, the characteristics of the set of ten evidences E are relevant. A higher mean of the evidences relative to the value increases significantly the number of reported evidences. Still, this higher mean and a higher standard deviation lead to an increased publication bias. This reflects the selection into more disclosure when E consists of high evidences in the high ranks.[14]

[13]This relationship exists close to the lower bound of v as well as over the whole interval. See Appendix A.2 for a discussion of the optimal increase in the markdown close to v's lower bound.

[14]Neither gender nor an economics-related field of study have a significant effect on any of these dependent variables.

	$\#M$ (1)	$\#M$ (2)	Pub. bias (3)	Pub. bias (4)	Markdown (5)
Competition	0.87***	0.87***	-22.58***	-22.53***	-40.32***
	(0.21)	(0.21)	(3.82)	(3.82)	(7.80)
X10+5	0.31	-0.30	7.23	2.05	-3.43
	(0.27)	(0.26)	(5.81)	(6.66)	(11.68)
Competition	-0.16	-0.20	6.69	6.39	9.82
$\times$ X10+5	(0.29)	(0.28)	(5.56)	(5.55)	(10.25)
Rounds 6-10	-0.13	-0.03	-0.46	0.37	9.52
	(0.20)	(0.19)	(3.37)	(3.37)	(6.13)
Value v	0.0007***	0.0007***	-0.01*	-0.01**	
	(0.00)	(0.00)	(0.00)	(0.00)	
Mean $E - v$	0.01***	0.01***	0.67***	0.67***	
	(0.00)	(0.00)	(0.04)	(0.04)	
SD E	-0.01***	-0.01***	0.60***	0.60***	
	(0.00)	(0.00)	(0.05)	(0.05)	
Purchase +5		1.79***		15.21***	
		(0.27)		(4.93)	
$\bar{e}^M$					0.09***
					(0.01)
Constant	5.43***	5.42***	12.20*	12.04*	-2.08
	(0.31)	(0.30)	(6.93)	(6.91)	(9.60)
N	1600	1600	1592	1592	1592
Subjects	160	160	160	160	160
R^2 overall	0.10	0.19	0.26	0.26	0.11

Notes: Panel fixed-effects regressions. Cluster-robust standard errors (subject level) are provided in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level. In 7 instances, sellers choose $\#M = 0$ and do not feature in (3)-(5).

Table 2.10: Determinants of $\#M$

		#M / Rank k														
		15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
(a)	k	-173.5	-124.7	-94.7	-71.4	-51.4	-33.4	-16.3	0.2	16.7	33.7	51.7	71.7	95.0	125.1	173.8	
Mean	Top k	0.2	12.6	23.1	33.0	42.4	51.8	61.3	71.0	81.1	91.8	103.5	116.4	131.3	149.5	173.8	
	X10+5	∶7.4	-2.5	-10.6⁝		52.5	53.7	46.9⁝	51.5⁝	70.2⁝	74.0⁝	89.0⁝	77.2⁝	134.7	125.6	166.6	
(b)	N	8	1	1	0	3	77	20	21	17	16	13	14	13	5	3	
Pub. bias	M10+5	∶20.2					⁝59.8	49.2⁝	59.6⁝	79.2	93.4	95.6	108.4	138.7	125.6	143.8	
Purchase	N	1	0	0	0	0	27	10	10	10	7	9	9	12	5	2	
	C10+5	5.5	-2.5	-10.6⁝		52.5	50.5	44.7⁝	44.1⁝	57.4⁝	58.9⁝	74.0⁝	21.1⁝	86.3·		212.1	
	N	7	1	1	0	3	50	10	11	7	9	4	5	1	0	1	
(c)	k						-154.2	-100.4	-65.8	-37.8	-12.5	12.0	37.5	65.5	100.1	153.7	
Mean	Top k						-0.2	16.9	31.6	45.5	59.4	73.8	89.2	106.4	126.9	153.7	
	X10+5						⁝6.3	∶27.8	32.3	41.5·	43.5⁝	53.2⁝	62.2⁝	75.4⁝	102.8⁝	113.7⁝	
(d)	N						37	12	26	47	47	79	74	41	49	14	2
Pub. bias	M10+5						-9.5·	24.4	21.3⁝	40.4	61.6	61.1⁝	79.2⁝	78.8⁝	111.5∶	153.4	
No	N						8	4	12	16	23	49	37	22	36	9	2
Purchase	C10+5						⁝10.6	∶29.6	⁝41.8	42.0	26.1⁝	40.3⁝	45.2⁝	71.5⁝	78.8⁝	42.2⁝	
	N						29	8	14	31	24	30	37	19	13	5	
	X10+5	-4.1	-43.8	69.9		5.5	27.1	23.1	25.3	39.4	47.3	60.3	50.6	78.5	99.1	119.7	
	N	17	3	2	0	6	159	34	52	54	47	79	76	47	48	14	2
(e)	M10+5	-0.9					26.1	34.2	19.8	73.6	49.7	74.7	80.0	95.7	120.9	166.7	
Markdown	N	1	0	0	0	0	35	14	22	26	30	58	46	34	41	11	2
	C10+5	-4.3	-43.8	69.9		5.5	27.4	15.4	29.3	7.6	43.1	20.5	5.5	33.4	-28.8	-52.8	
	N	16	3	2	0	6	124	20	30	28	17	21	30	13	7	3	
(f)	C10+5 Chosen	2.3	-2.5	-10.6		52.5	37.2	42.3	46.7	30.7	9.3	50.6	43.8	75.9	81.9	52.2	
Pub. bias	sellers	16	3	2	0	6	124	20	30	28	17	21	30	13	7	3	

Notes: Average publication bias being ⁝above (below⁝) the simulated confidence interval of N Top k draws is denoted with Maya numerals ⁝ (99%), ∶ (95%), and · (90%) to the left (right) of the average.

Table 2.11: Summary by number $\#M$ and rank k of disclosed evidences (X10+5 games)

2.5 What explains lack of skepticism in competition?

The results so far raise the question why we observe the lack of skepticism in competition. In the following we investigate three possible explanations. First, fairness concerns might motivate buyers to reciprocate the generous information provision of the chosen seller by not reducing their bid too strongly. Second, the very choice of a seller might give buyers an illusion of control and make them optimistic about the selection of evidence that they face. Third, the knowledge of the competitive pressure might lead to the belief that the selection of published evidence is less biased than under monopoly.

We investigate these explanations with the help of three further games that only feature buyers. All three games change details of the C10 game and confront new buyers with data that had been generated by sellers in the previous sessions. First, in "C10B", buyers face markets and published evidences as generated in C10 but their bids only have financial consequences for themselves, not for any seller. Avoiding consequences on other participants removes any relevance of social preferences and reciprocity and should thus re-establish levels of skepticism as observed in monopoly. Otherwise identical to C10B, "C10B-NC" (No Choice) imposes on buyers the choices of buyers in the original C10 markets, thus removing any own choice and any resulting feeling of control that could lead to different behavior between competition and monopoly. Lastly, "C10B-4M" (four Monopolists) confronts buyers with markets that are made up of four monopolists' products and published evidences as previously observed in M10. Since the behavior of those monopolists is not shaped by competitive pressure the observed level of skepticism towards the chosen seller should be as in monopoly.

2.5.1 Experimental procedures

The additional six sessions were conducted in Mannheim (MA, three sessions) and in the AWI-Laboratory of the University of Heidelberg (HD, three sessions). Each session features a sequence of the three games with each game occurring once in each sequence position (Latin Square design), see Table 2.12. Each session features five rounds of each treatment, using data from the first five rounds of the respective games in sessions one to ten.

Session	Sequence position		
MA, HD	1	2	3
11, 14	C10B	C10B-4M	C10B-NC
12, 15	C10B-NC	C10B	C10B-4M
13, 16	C10B-4M	C10B-NC	C10B

Table 2.12: Games in extension sessions 11-16.

As opposed to the eight seller and eight buyers in previous sessions, each session now includes 16 buyer. Thus, each set of three sessions (MA, HD) uses the data from the six C10 games in sessions one to ten. In the C10B-4M games, data from the six M10 games are fed in. Due to non-shows of three participants, a total of 93 subjects provided 279 observations per game.

2.5.2 Results

Table 2.14 presents the results by games. By design, panels a and b coincide with averages of periods one to five in C10 and M10, respectively. The non-shows introduce small differences, also between C10B and C10B-NC. With respect to our hypotheses regarding the buyers' behavior, the markdowns are the most interesting evidence here. Recall from Table 2.6 that the round one to five markdown in C10 was -4.23 and in M10 -56.59. Strikingly, we can see that the markdown in all games is less pronounced than in C10.

		C10B		C10B-NC		C10B-4M	
(a) Nr Published evidences		5.77		5.80		4.9	
(b) Publication bias		47.13		46.55		74.45	
(c) Markdown		0.91	[0.86]	1.83	[0.79]	0.57	[0.94]
	$\neq 0$	[0.79]		[0.63]		[0.86]	
(d) Precision		38.61	[0.86]	37.56	***	55.76	***
	$\neq 0$	***		***		***	
N		465		465		465	

Notes: t-tests for equal means and non-zero means, significance level indicated by: *** p <0.01, ** p <0.05, * $p \leq$0.1, [$p > 0.1$].

Table 2.13: Buyers' markdown and bidding precision

With respect to the explanations for the low skepticism in competition, this allows us to conclude as follows. The similar behavior in C10 and C10B suggests that generosity is not behind the non-skeptical attitudes under competition. Further, the similar behavior in C10B and C10B-NC suggests that the own choice of a seller has relatively little influence on skepticism either and does not explain the large skepticism in M10. Finally, the similar behavior in C10B and C10B-4M suggests that the competition does not influence much the beliefs about the selected evidence either.

So overall, the results of these additional games allow us to rule out some simple explanations for the difference in skepticism, but we cannot single out what would be required to establish skepticism. So it is possible that a different reason is behind the observed effect. Further, we cannot rule out an interaction of NC or 4M with the removal of social preferences B, although we are lacking any intuition why that should exist.

2.5.3 Some checks

None of the markdowns is significantly different from zero.

		C10B	C10B-NC	C10B-4M
(a) Markdown		2.33	1.25	-0.75
Seq. position 1	N	155	160	150
(b) Markdown		2.91	2.05	0.24
Seq. position 2	N	160	150	155
(c) Markdown		3.53	2.22	2.12
Seq. position 3	N	150	155	160

Table 2.14: Buyers' markdown and bidding precision over rounds

One hypothesis that we have not checked is about the complexity of competition versus monopoly. So just the fact that you are confronted with four sellers and only after your choice are back in the situation as in monopoly.

2.6 Conclusion

In this study, we establish a rich experimental market framework to study disclosure of verifiable information by interested parties as well as the inference behavior on the consumer side. The setup cleanly implements the natural conflict of interest inherent in any seller-buyer interaction without adding further confounding strategic elements. In particular, changes in the market and in the informational structure allow us to investigate behavioral responses to those institutional variations. In the monopolistic baseline setting, we replicate previous findings of insufficient skepticism on the buyer's side and resulting incomplete disclosure on the seller side.

Competition raises the amount of information disclosed and should help consumers' inference. However, the skepticism of buyers turns out so low that it nearly compensates the potential payoff improvements and leaves buyers as well off as under monopoly. This suggests that buyers feel protected by competition to an extent that they act less cautiously.

Introducing the possibility for sellers to purchase additional evidences creates uncertainty about the amount of information available to the sellers. Buyers account for this uncertainty and bid more skeptically, particularly in monopoly.

The suggested framework can be extended in order to understand further influences on behavior that have been discussed theoretically. For example, it would be interesting to investigate disclosure about outside options that sellers know more about than consumers. Milgrom and Roberts (1986) predict that cigarette companies would not reveal information about the benefits of not consuming any cigarettes. Further interesting aspects could be partial verifiability and costly influence on the realization of evidence. For example, Lesser et al. (2007) show in the context of nutrition-related research that funding sources are predictable of the study outcome.

The recent investigations of skepticism in the literature show that it is an important feature of consumer behavior which can be very useful when it is calibrated well to the disclosure behavior and driving interests. We believe that it will be a very relevant and important field of study which can increase consumer welfare significantly. This is especially true for products whose full effects on welfare are understood only via complex investigations, such as cigarettes and food items.

3 Selection in Markets: An Experimental Approach[1]

3.1 Introduction

Over the last several decades laboratory experiments have become an extremely popular tool for studying economic behavior in depth. Perhaps unsurprisingly, one of the earliest traditions in this literature was the investigation of competition in markets, starting with Chamberlin (1948), Smith's early contributions (Smith, 1962, 1964) and the works of Sauermann and Selten (1959), E. Fouraker and Siegel (1960, 1963), and Friedman (1963, 1967). Utilizing the controlled conditions of the laboratory as a basic testing ground for theories of competition has the key advantage that the experimenter can recreate the specific conditions (in terms of costs, demand, information, timing, number of competitors etc.) that theories assume in order to derive predictions. On the other hand, there are worries about the external validity of experiments that recruit from the typical subject pool - mostly undergraduates, sometimes graduate students - for tasks relating to the decision-making of oligopolistic firms, and such worries appear to be more pronounced than in experiments in which students essentially act as consumers, as workers, or as nondescript individual decision makers who are asked to make basic choices under risk or uncertainty. Thus, a fundamental question is: Can

[1] This chapter is based on joint work with Henrik Orzen.

we draw any useful conclusions from laboratory research that examines the behavior of firms without letting actual firms participate in the experiments?

While this is a reasonable question to ask, it cannot be answered by armchair economics. Specifically, a general a priori “sense of uneasiness” about employing student volunteers to represent firms in the laboratory is too nebulous to be convincing. Instead, it must be spelled out what exactly may prevent the parallelism between what is observed in experimental markets and what is going on in naturally-occurring ones. On such a basis empirical or experimental tests can be carried out to examine specific points of criticism (see Siakantaris, 2000; Guala, 2002).

Indeed, the issue of whether firm behavior can be legitimately studied in the laboratory has multiple facets. Clearly, decisions in firms are made by people, and experimental subjects are people. However, some may suspect that managers and entrepreneurs are a particular breed of people who are perhaps more risk loving or more creative than the average student and who will also be more experienced. Secondly, the process that leads to important decisions in large companies typically involves groups, for example a management board or an advisory committee of experts and such decisions may have to be justified to other stakeholders. Yet, in most market experiments firms are represented by individuals who act completely on their own. Furthermore, the decisions of real firms often involve high stakes, whereas student volunteers are paid on average little more than the minimum wage for their time spent in the laboratory. Some skeptics also voice concerns that the small-group settings in the lab may cause participants to be more inclined towards cooperative behavior amongst each other than are real competitors at the market place.

There are counterarguments to some of these claims. For instance, regulators recognize tacit collusion as a relevant phenomenon in many naturally-occurring markets, and entrepreneurs, CEOs or other top-level managers often take individual responsibility for vital business decisions. However, while clarifying whether or not there are close field analogs of situations or outcomes we typically find in the laboratory would be an intriguing undertaking, we are more interested here in learning whether specific concerns about external validity can be shown to matter for appropriate experimental designs or procedures in order to establish a higher degree of parallelism between the lab and the field. One example is the suspicion that replacing students by professionals in the relevant field leads to systematically different results. If this can be established we would know that professionals differ from students to such an extent - in terms of their knowledge, experience or preferences - that we ought to pay close attention to these factors. This question has already received some attention in the literature (see Frchette, 2015, for an overview). For better or worse, the current evidence does not send strong signals that the student subject pool leads to very biased results. However, to date the volume of this research is limited and thus it is difficult to draw clear conclusions.

In this paper, we contribute to the debate by examining whether and in how far a popular line of reasoning to support the assumption of rational firms in models of competition is relevant to how experimental oligopoly research ought to be conducted. The argument dates back to at least Friedman (1953) and goes something like this: Even though there is no guarantee that any individual business enterprise will always comply with the principles of rational firm behavior, market forces favor businesses that maximize profits and reject businesses that do not. Thus, since those who make good decisions will thrive and those who make bad decisions will sooner or later be forced to exit, we should expect most firms - at least in established markets - to conform to the model of the sophisticated rational firm. This is a powerful argument and while its main role is to defend the use of the rationality assumption in theoretical work it also provides a line of attack against experiments: Confronted with an unfamiliar situation student subjects may resort to simple, improvised heuristics when deciding what to do and may be unable to discover optimal strategies in complex market settings. Although bad performance is punished to some extent by low payoffs, conventional experimental protocols do not stipulate the removal of unsuccessful entrepreneurs or provide nonentrepreneurial activities as alternatives. Hence, experimental results obtained in any other than the very simplest settings will be biased towards naive behavior.

We employ the following strategy to investigate the importance of such market force effects to the validity of experimental results. Our subjects interact as competing “firms” and we confront them with a multi-layered strategic market setting, which is intentionally more complex than most experiments on oligopolistic competition without venturing into the absurd and without deviating too much from the kinds of settings that are typically implemented.[2] The Baseline treatment is procedurally conventional in that student subjects are allocated at random to markets of different sizes even though in most experiments such allocations are done between-subjects whereas we employ a within-subject protocol. This treatment also resembles many other experimental studies in that poor performance is a distinct possibility but does not lead to any selection. Thus, there is not much evolutionary pressure for bad decisions to be weeded out over time, which may have repercussions also on the behavior of other players who have a better understanding of how to maximize profits as they will take their competitors’ suboptimal choices into account. The Market Force treatment addresses this potential shortcoming by introducing endogenous market entry. That is, subjects now self-select into the market but they can do so only as long as they have sufficient liquidity and are expelled from the market if they do not. We expect there to be sorting effects and we ask whether this leads to adjustments over time such that market outcomes change in meaningful ways. We also investigate whether and how individual decisions in our markets vary systematically with other measurements we take from our participants. For this we run a battery of post-tests that examine risk aversion, loss aversion, cognitive skills and social value orientation at the level of individuals.

In our market experiments we consider duopolies, triopolies and quadropolies. Once the market size is known in a given round, subjects first choose capacity levels and then compete by setting prices in three consecutive periods. Thus, our setting resembles the

[2]For comparisons see the review of the more recent literature by Potters and Suetens (2013) and the review of the earlier literature by Holt (1995).

Kreps and Scheinkman (1983). However, we use a demand function that avoids a situation in which only mixed-strategy solutions exist on certain game paths. Instead, unique pure-strategy predictions are obtained in all subgames. This has the advantage that there are always clear theoretical benchmarks against which we can compare actual behavior. We do not claim that our setting provides an accurate description of any specific real world market. Our goal is a different one, namely to investigate whether outcomes will be driven by naive play and whether market selection forces can discipline our student subjects in this respect. There are several details in the predicted equilibrium path that naive players could "get wrong". For example, choosing a suitable capacity level is an important aspect of individually optimal behavior: Capacities that are too high lead to losses; the equilibrium solution is characterized by firms artificially withholding capacity; and too small capacities impede sizable profits in the pricing stage. We have a somewhat counterintuitive comparative prediction here: More firms lead to lower total outputs. This is because in equilibrium firms contract their capacities extremely much when more entry occurs. Profits are also predicted to decrease with the number of competitors, and so are prices despite the reduction in total output.

We observed strong self-selection into markets when entry was endogenous. However, self-selection did not seem to affect capacity and pricing decisions substantially. The outcomes in both treatments – Baseline and Market Force – did not differ significantly.

In both treatments participants performed best in three-firm markets in the sense that, on average, their choices were very close to the equilibrium predictions. In duopolies firms did not fully exploit their market power, especially in the first rounds of the experiment. There, mean capacities and prices tend to be lower than ex-ante predicted. By contrast, mean capacities and prices were higher than predicted in four-firm markets.

Although we constructed a deliberately complex market setting, participants very quickly adjusted to the market situation in both treatments and average capacities and prices converged towards the equilibrium values.

We found that the market entry decision did not vary significantly with any of the characteristics we measured in the post-tests. Experience and cognitive skills affected decisions and profits significantly in the sense that with increasing experience and cognitive skills capacity and price choices were closer to the equilibrium predictions and profits increased. This applied to both treatments.

This chapter is organized as follows. Section 3.2 gives a more detailed description of the experiments and the post-tests. Section 3.3 discusses theoretical considerations and predictions before the results are presented in section 3.4. The last section 3.5 concludes.

3.2 Experimental Design

3.2.1 Baseline treatment

Main market

Participants are matched into groups of five. Each participant represents a firm. For each group there is a single main market. Each round consists of three stages:

Stage one: Market allocation

Stage two: Capacity choice

Stage three: Price choice in three consecutive periods

In the Baseline treatment participants are randomly allocated to either the main market or the side market at the beginning of each round (stage one). Those who are on the main market will compete with those in their group who have also been selected to the main market. Participants on the side market do not compete with other firms since there is a separate side market for each firm.

Everyone is informed about the number of firms on the main market. Participants who are on the side market receives this information as well. The demand on the main market for firm i's good in period t is given by

$$q_{it}(p_{it}, \bar{p}_{-it}, n) = \max\left\{0, \frac{120}{n} - 2p_{it} + \bar{p}_{-it}\right\} \tag{3.1}$$

.

where n is the number of firms present in the main market, p_{it} is firm is price in period t and $\bar{p}_{-it}$ is the competitors' average price ($\bar{p}_{-it} = 0$ if $n = 1$) in that period. In stage two each firm i decides on its production capacity $\bar{q}_i \in [0, 50]$ with $\bar{q}_i \in \mathbb{Z}$ for the current round. A firm's production capacity determines how many units it can maximally produce in each period of stage three. Once the decision about the production capacity is made, it cannot be changed in stage three. Each capacity unit costs 50 ECUs and has to be paid upfront, prior to the start of the pricing stage.

At the beginning of stage three the production capacities of all firms on the main market are announced. Then firms in the main market compete in three consecutive periods by choosing prices $p_{it} \in [0, 50]$ with $p_{it} \in \mathbb{Z}$ for their goods. In period two and three firms are informed about all firms' price choices of the previous period. After the third period, the round ends.

At the end of each round, participants are informed about their round profit. In addition to that, they also receive a market summary including all firms' capacity levels and the average prices in each period.

Side market

All firms that are not in the main market operate on a firm-specific side market. Similar to the main market firms first determine their production capacity $\bar{q} \in [0, 20]$ in stage ne and then choose prices $p \in [0, 50]$ in three consecutive periods in stage two. The maximum possible capacity is only 20 units because demand levels are generally lower than on the main market. The demand level on the side market fluctuates randomly from round to round and is characterized by

$$q_t(p_t) = A - 2p_t \tag{3.2}$$

where A is drawn at the beginning of each round from the set $\{53, 54, , 59\}$ with any number being equally likely. Subjects are informed about the value of A only after having made the decision to enter the side market but before making any choices. Stage two and stage three correspond to the stages on the main market.

At the end of each round, participants receive feedback on their expenditures for capacity and their revenue from the three periods. Moreover, they also get a summary of what has happened on the main market in the particular round (number of firms, chosen capacity levels, average prices).

Using an outside option task in which participants had to make the same decisions as on the main market with the only difference that there is no competition, has several advantages for our experiment. Firstly, when both tasks – main and outside option – are the same, we avoid that participants select the main market or side market based on how much they like or dislike a particular task. This would lead to a selection bias that we do not desire. Secondly, the side market is a version of the main market without strategic risk as participants do not complete against others there. Hence, it might attract those players who are seeking to avoid this type of strategic interactions with others. And lastly, being on the side market gives participants the opportunity to further familiarize themselves with the setting and the decisions they are supposed to make.

3.2.2 Market Force treatment

In the Market Force treatment the way firms operate and make their decisions is the same as in the Baseline treatment. The only difference is that in the Market Force treatment firms are allowed to choose whether they want to be on the main market or rather operate on the side market for each round. Participants make this decision simultaneously and individually at the beginning of each round.

At the end of each round, firms that are present on the main market individually decide for themselves whether they wish to remain on the main market for the subsequent round or wish to exit the main market. Firms that decide to exit will operate on the side market during the subsequent round. All firms whose account balance falls below 300 ECUs are not allowed to operate on the main market. This minimum capital requirement forces firms that were previously on the main market and made unprofitable decisions to exit the main market. It is worth noting that this is not a pure bankruptcy rule but stricter. Market forces will eliminate those firms who did not make profitable decisions or do not have enough capital to operate on the market.

In the side market the only difference to the Baseline treatment is that in the Market Force treatment, firms decide at the end of each round whether they want to enter the main market or remain on the side market. Due to the minimum capital requirement, entering the main market requires an account balance of at least 300 ECUs.

For firms who were previously forced to exit the main market as their account balance dropped below 300 ECUs, being on the side market can also be seen as a "recovery phase" to rebuild their resources. As the average expected payoff on the side market is between 160 ECUs and 330 ECUs, firms can reach the necessary balance threshold for entering the main market within one to two rounds.

We allowed the main market to only grow by one firm from round to round. Hence, not every firm who wants to enter the main market is admitted. If more firms want to enter than allowed, the selection of the firms who may enter is done at random. This way, we could avoid high uncontrolled variations of the market size. Once market entry decisions are made simultaneously there is a high probability of coordination failure, which can lead to high fluctuations in terms of market size. Although this volatility due to coordination failure an interesting phenomenon on its own, it is not the focus of our work here.

3.2.3 Post-test

In order to see whether and how certain individual traits affect entry decisions and market behavior, we constructed an additional test consisting of four independent smaller tasks. We picked four different measurements to test for, as we believe that these might affect individual choices in the following ways:

(i) Loss aversion: Subjects who are more loss avers may not want to enter the main market and rather operate in the side market where their outcome does not depend on the interaction with other players. Furthermore, highly loss avers subjects may choose to produce less or charge lower prices to ensure that they do not have unsold overcapacity.

(ii) Cognitive skills: The hypotheses here is that higher cognitive skills lead to "better" capacity and pricing decisions and hence higher profits.

(iii) Risk aversion: Participants who are more risk avers may prefer to stay on the side market, as there is no strategic risk there. As for capacity and price choices, risk avers subjects might be more careful to not overproduce or overprice.

(iv) Social preferences: Participants with a higher pro-social preference might be less competitive on the market, for example in term of price setting.

Task 1: Loss aversion

We use a loss aversion design as in Rau (2014) for this part with ten lotteries each with a 50-50 chance of a good outcome and a bad outcome (see Table 3.1). Each subject receives an initial endowment of 25 points and has to decide for every of these ten lotteries whether or not s/he would be willing to play it. Once all choices have been made, the computer randomly selects one of the ten lotteries. If the subject's previous choice was to play this lottery, his/her payoff is the initial endowment plus or minus the amount indicated depending on whether the good or bad outcome was drawn. If the individual chose not to play this lottery s/he receives the initial endowment of 25 points.

Lottery	Bad outcome	Good outcome
1	-4 points	20 points
2	-6 points	20 points
3	-7 points	20 points
4	-8 points	20 points
5	-10 points	20 points
6	-12 points	20 points
7	-14 points	20 points
8	-17 points	20 points
9	-20 points	20 points
10	-25 points	20 points

Table 3.1: Lotteries of the loss aversion test

Task 2: Raven's standard progressive matrices

In order to test participants' cognitive skills we used the Raven's standard progressive matrices (referred to as Raven's SPM). These matrices were originally developed by John C. Raven (Raven, 1936, 1981), designed to test the reasoning and problem solving abilities of the test-takers. In each question of this test participants are shown a panel containing a pattern of geometric shapes with one element missing. Each pattern is in the form of a 3x3 matrix. In each case there are eight proposed solutions. An example of such a task is given in Figure 3.1. Participants are asked to identify the solution that is best suited to complete the pattern. In our experiments participants were shown 12 panels in total with increasing difficulty. They had 12 minutes to find the 12 missing elements.

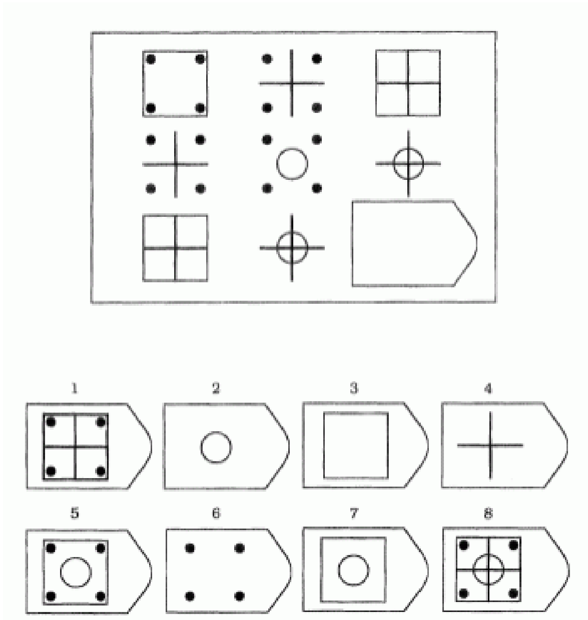


Figure 3.1: Example of one panel from the Ravens test

Task 3: Risk aversion

We used the Eckel-Grossmann test (see Eckel and Grossman, 2002; Eckel et al., 2007) to analyze the risk aversion of participants. Participants were confronted with six lotteries as depicted in Table 3.2, each with a 50% probability of receiving the higher outcome. From these six lotteries subjects were asked to choose one. The lotteries are increasing in variance and expected value allowing us to categorize the participants into five risk categories. A completely risk avers subject would chose lottery one, which has the lowest expected payoff but guarantees a safe outcome of 28 points. A risk neutral player who only maximizes expected payoffs would be indifferent between choosing lottery five and six. However, considering the variance of the lotteries, lottery six is a mean-preserving spread of lottery five and hence could be seen as even more risky than lottery five by the participants.

In addition, participants are ask to state whether they would consider themselves as a person who is willing to take risks in general. They are asked to indicate their

Lottery	**Payoff: Heads**	**Payoff: Tails**
1	28 points	28 points
2	36 points	24 points
3	44 points	20 points
4	52 points	16 points
5	60 points	12 points
6	70 points	2 points

Table 3.2: Lotteries of the Eckel-Grossmann test

own assessment on a scale from 1 to 10 with 1 = not at all willing to take risks; 10 = completely willing to take risks.

Task 4: Liebrand's ring measure of social values

As the name already says this ring measure is used to assess a subject's social value orientation (SVO) and was first devised by W. B. G. Liebrand in 1984 (Liebrand, 1984). The SVO assesses how much weight a person attaches to others' outcome in relation to their own.

For this task participants are randomly matched into groups of two. The identity of the partner remains anonymous. Participants face a set of 24 scenarios each containing a pair of options (option A and option B), which consists of a combination of the subject's own payoff and the payoff of the other participant.

Example scenario:

	Option A	Option B
Your payoff	13.00	14.50
Other's payoff	7.50	3.90

Each scenario corresponds to a point on the SVO ring which is a circle centered at the origin of the Cartesian plane. The vertical axis represents the payoff of the other subject and the horizontal axis represents a subject's own payoff.

Adding up all 24 allocations choices of a participant yields an own-other-payoff allocation vector. The angle of this vector indicates the subject's social value orientation. There are four common orientations: altruistic, individualistic, cooperative and competitive. The altruistic subject maximizes the other's payoffs, while the individualistic subject maximizes his/her own payoffs. The cooperative subject prefers win-win situations and maximizes joint payoffs and the competitive subject maximizes his/her relative advantage (difference between own payoff and the other's payoff). The length of this own-other-payoff allocation vector indicates how consistent and efficient the subject's choices were. If the length of the own-other-payoff allocation vector is equal to the radius of the SVO ring all choices were completely consistent and pareto efficient.

3.2.4 Experimental Procedure

We conducted the experimental sessions in the Experimental Laboratory of the University of Mannheim (mLab).[3] Across eight sessions, a total of 120 undergraduate and postgraduate students participated. All students were recruited from the general student population of the University of Mannheim. Compensation of all participants was calculated based on their ECU balance at the end of the experiment. The exchange rate was 175 ECUs = 1 Euro.

Throughout the entire experiment participants are provided with a calculator. The purpose of this calculator was firstly to help participants familiarize themselves with the possible scenarios and how they affect their ECU earnings for different market sizes. And secondly, to assist participants with finding their optimal capacity choice and prices based on their believes about the other firms' choices.

At the beginning of each round, the calculator for the main market has four sliders that allow participants to set (a) the number of firms on the main market, (b) their production capacity, (c) their price and (d) what they believe the average price of the other firms on the main market is (see Figure 3.2). Given the chosen values for (a), (b), (c) and (d), the calculator showed how many units would be sold, the capacity costs and the expected revenues.

[3] The experiment was programmed in zTree (Fischbacher, 2007) and subjects were recruited through ORSEE (Greiner, 2004).

Once the actual number of firms on the main market is determined and announced to everyone, the first slider disappears and the calculator uses only the actual number of firms.[4] After stage two in which all participants chose their production capacities the second slider also disappears and the calculator uses only the actual chosen capacity. Hence, in stage three (pricing stage) the calculator only has two sliders ((c) and (d)).

For those who were on the side market a calculator with two sliders, one for the production capacity and one for own prices, was provided.

Figure 3.2: Calculator

At the beginning of each session, the instructions were read out and explained to all participants. Participants were randomly allocated to groups of five. Throughout the entire experiment participants stayed in these groups and only interacted with the other four members in their own group but not with the members of other groups.

The entire experiment consisted of ten rounds. Each round constituted of one two stage game in which firms chose their capacities and prices as described before.

At the beginning of the first round participants had up to four minutes to work with the calculator. After these four minutes participants were randomly allocated to either the main market or the side market in the Baseline treatment. In the Market Force treatment, participants had to decide whether they want to enter the main market or operate in the side market within these four minutes. By clicking either the "yes"

[4] In the Baseline treatment this happens after the random draw has allocated subjects to the main market and in the Market Force treatment it is after subjects have made their entry decisions.

or the "no" button on the screen they were able to submit their decision. If participants did not press a button within the four minutes, they will automatically enter the side market. Once participants were allocated to the corresponding markets, they were asked to chose capacities and prices.[5] Participants had certain time limits to submit their capacity choice and price choice decisions. The time limits for price choices were shorter in the later five rounds of the experiment. If a participant did not submit a decision within the time limits, the value that they had selected for the relevant variable in their calculator (production capacity or price) was automatically submitted.

In order to achieve a better comparability, we programmed the sequence of market sizes in the Baseline treatment based on our observations in the Market Force treatment. We constructed the Baseline treatment in a way that each session exactly mirrors one Market Force treatment session in terms of the size and the order of the main markets. As we wanted to focus our analysis on duopolies, tripolies and quadropolies, we replaced $n = 1$ by $n = 2$ and $n = 5$ by $n = 4$ in the Baseline treatment.[6]

The post-tests were conducted separately from the main experiment. Additional invitations for the post-tests were send out to all participants of the main experiment after all sessions of the main experiment were completed. 79% of subjects who took part in the main experiment participated in the post-tests. In those post-test sessions all participants worked on their own and completed the four tasks as described above in section 3.2.3 individually. At the end of a session they received feedback on their choices, the lottery outcomes and their payoffs. They were also informed about how much they were given by their counterparts in the Liebrand's ring measure of social values task. All earnings were given in points and the exchange rate was 12 points =

[5] Appendix B.1 provides screenshots of the games.

[6] A monopoly situation only occurred in two markets for one round each and $n = 5$ only occurred in one market for one round in the Market Force treatment.

1 Euro. In addition to what they earned each participant also received a show-up fee of 4 Euros.

3.3 Theoretical Considerations

The market setting we investigate is characterized by differentiated goods with capacity-constrained firms and with or without endogenous entry. As described before, the sequence of events in a given round in the endogenous-entry variant is as follows:
Stage one: Firms independently and simultaneously decide whether or not to enter the main market.
Stage two: Entrants observe the number of entrants and then independently and simultaneously choose their production capacities. Each unit of capacity comes at a cost of c.
Stage three: Firms observe the vector of chosen capacities and then independently and simultaneously choose prices in T consecutive periods.
Recall that the demand for firm i's good in period t is

$$q_{it}(p_{it}, \bar{p}_{-it}, n) = \max\left\{0, \frac{120}{n} - 2p_{it} + \bar{p}_{-it}\right\}.$$

where n is the number of firms present in the main market, p_{it} is firm i's price in period t and $\bar{p}_{-it}$ is the competitors' average price ($\bar{p}_{-it} = 0$ if $n = 1$) in that period. Thus, the demand for firm i's product decreases hyperbolically in the number of firms, decreases linearly in firm i's own price and increases linearly in the other firms' mean price.[7] Marginal costs of production in stage three are constant and normalized to zero.

[7] This demand system has the intuitive property that a firm's price induces a positive externality on its rivals. It has the unintuitive property that this externality holds even when the demand for the firm's own product drops down to zero. While there are various ways in which this could be fixed, we chose to disregard this issue because (i) doing so avoids unnecessary additional complications in the demand function, (ii) it is irrelevant for the equilibrium solution, and (iii) the price range in our experiment is limited and potential effects would be small.

In the following we first derive the subgame-perfect equilibrium (SPE) of the game for n market entrants and subsequently determine n by considering the alternative to entering the main market. To find equilibrium prices in the stage three pricing subgame we must take into account the capacity choices made in the previous stage two. If firm i's capacity, $\bar{q}_i$, is sufficiently large at given prices, such that $\bar{q}_i \geq q_{it}(p_{it}, \bar{p}_{-it}, n)$, it is not capacity-constrained and will sell $q_{it}(p_{it}, \bar{p}_{-it}, n)$ units.[8] Conversely, if the demand exceeds firm i's capacity, it will sell $\bar{q}_i$ units.

As long as firm i's rivals charge sufficiently high prices, its best response is to operate at maximum capacity and to select a price p_{it} such that the demand $q_{it}(p_{it}, \bar{p}_{-it}, n)$ is precisely equal to $\bar{q}_i$:

$$p_{it}^{BR1}(\bar{p}_{-it}, n) = \frac{60}{n} - \frac{1}{2}\bar{q}_i + \frac{1}{2}\bar{p}_{-it}. \tag{3.3}$$

Should the competitors choose very low prices, demand for firm i's product will be limited and it is not optimal to lower p_{it} so much that it sells out. In this case firm i's best response is

$$p_{it}^{BR2}(\bar{p}_{-it}, n) = \frac{30}{n} + \frac{1}{4}\bar{p}_{-it}. \tag{3.4}$$

Firm i finds it optimal to withhold excess capacity in this way if $\bar{p}_{-it} < 2\bar{q}_i - \frac{120}{n}$. Consider now the possibility of an equilibrium in which m firms are at their capacity limit and best respond according to (3.3) and $n - m$ firms have excess capacity and best respond according to (3.4). In the following we will use the letter i to index full-capacity firms and the letter j to index excess-capacity firms. Both types of best response functions can be reformulated to express own price as a function of the sum of all prices, P_t:

$$p_{it}^{BR1} = \frac{n-1}{2n-1}\frac{120}{n} - \frac{n-1}{2n-1}\bar{q}_i + \frac{P_t}{2n-1} \tag{3.5}$$

$$p_{jt}^{BR2} = \frac{n-1}{4n-3}\frac{120}{n} + \frac{P_t}{4n-3} \tag{3.6}$$

[8] In our setting a firm's capacity level indicates how much it is able to produce in each of the T periods.

Summing up equation (3.5) over the m full-capacity firms and equation (3.6) over the $n-m$ excess-capacity firms, and denoting $\bar{Q}_m = \sum_{i=1}^{m} \bar{q}_i$ as the sum of capacities of all full-capacity firms, we obtain

$$P_t = \frac{n-1}{2n-1}\frac{120m}{n} - \frac{n-1}{2n-1}\bar{Q}_m + \frac{mP_t}{2n-1} + (n-m)\left(\frac{n-1}{4n-3}\frac{120}{n} + \frac{P_t}{4n-3}\right) \quad (3.7)$$

and solving for P_t yields the equilibrium sum of prices:

$$P_t^* = \frac{2m(n-1)+n(2n-1)}{(6n-3-2m)}\frac{120}{n} - \frac{4n-3}{6n-3-2m}\bar{Q}_m.$$

Using P_t^* in equations 3.5 and 3.6 we obtain

$$p_{it}^* = \frac{4n-3}{6n-3-2m}\frac{120}{n} - \frac{n-1}{2n-1}\bar{q}_i - \frac{1}{2n-1}\frac{4n-3}{6n-3-2m}\bar{Q}_m \quad (3.8)$$

and

$$p_{jt}^* = \frac{2n-1}{6n-3-2m}\frac{120}{n} - \frac{1}{6n-3-2m}\bar{Q}_m \quad (3.9)$$

.

In equilibrium the demand for a full-capacity firm is, by construction,

$$q_{it}^*(p_{it}^*, \bar{p}_{-it}^*, n) = \frac{120}{n} - 2p_{it}^* + \bar{p}_{-it}^* = \bar{q}_i$$

and the demand for an excess-capacity firm is

$$q_{jt}^*(p_{jt}^*, \bar{p}_{-jt}^*, n) = \frac{120}{n} - 2p_{jt}^* + \bar{p}_{-jt}^* = \frac{2(2n-1)}{6n-3-2m}\frac{120}{n} - \frac{2}{6n-3-2m}\bar{Q}_m$$

.

This solution can be used to find equilibrium prices and revenues for any stage three subgame, i.e. for any vector of capacity levels that was chosen in stage two. A simple numerical algorithm is required to establish m in each case. We use this approach to determine theoretical benchmarks at given capacities for the price-setting stage in our experiment.

In the subgame-perfect Nash equilibrium, however, no firm will hold unnecessary excess capacity since capacity is costly. The consequence of this is that on the equilibrium path equation (3.5) is a representation of *any* firms' best response. This allows us to rewrite equation (3.8) as $P_t^* = 120 - \bar{Q}$ where Q is the *total* capacity on the supply side. From this we obtain

$$p_{kt}^* = \frac{120}{n}\frac{n-1}{2n-1}\bar{q}_k - \frac{1}{2n-1}\bar{Q}$$

. Thus, in effect firm k chooses q_k in stage two to maximize

$$\pi_k^*(\bar{q}_k) = T\left(\frac{120}{n} - \frac{n-1}{2n-1}\bar{q}_k - \frac{1}{2n-1}\bar{Q}\right)\bar{q}_k - c\bar{q}_k$$

where T is the number of stage three periods ($T = 3$ in our experiment) and c is the marginal cost of capacity ($c = 50$ in our experiment).

This yields

$$\bar{q}_k^*(n) = \frac{2n-1}{3n-1}\left(\frac{120}{n} - \frac{c}{T}\right)$$

as the equilibrium capacity level and therefore

$$p_{kt}^*(n) = \frac{120}{3n-1} + \frac{2n-1}{3n-1}\frac{c}{T}$$

as the SPE price conditional on n. Per-round profits are given by

$$\pi_k^*(n) = \frac{(120T - nc)^2}{nT}\frac{2n-1}{(3n-1)^2}$$

Table 3.3 lists the equilibrium predictions for different values of n.[9] As the table shows, profit is severely affected by market entry as there is increased competitive

[9] For a monopolist the optimal stage three price is $p(\bar{q}) = 60 - \frac{\bar{q}}{2}$ for any $\bar{q}$ in the feasible range.

pressure on prices. Notice the curious prediction that total market capacity falls with the number of firms. We shall see whether this is borne out by the data.

Market size	Capacity per firm	Market capacity	Price	Per-period revenue	Per-round profit
$n = 1$	43.3	43.4	38.3	1661.1	2816.7
$n = 2$	26	52	34	884	1352
$n = 3$	14.6	43.8	25.4	370.7	382.8
$n = 4$	8.5	33.9	21.5	182.6	123.4
$n = 5$	4.7	23.6	19.3	90.9	37

Table 3.3: Equilibrium predictions

Subjects that do not enter the main market operate on firm-specific side markets in which they assume the role of monopolists. The demand in period t is given by

$$q_t(p_t) = A - 2p_t \tag{3.10}$$

where A is drawn at the beginning of each round from the set $\{53, 54, ..., 59\}$ with any number being equally likely. Subjects are informed about the value of A only after having made the decision to enter the side market but before making any choices. Stage two and stage three correspond to the stages on the main market. The firm's optimal price in stage three is given by

$$p^*(A) = \max\left\{\frac{A - \bar{q}}{2}, \frac{A}{4}\right\}.$$

That is, the firm will choose a price that ensures that all units are sold unless the capacity is so high that it is more profitable to withhold capacity. Obviously, it would not be reasonable to purchase such excess capacity in the first place. In stage two the firm's maximization problem boils down to maximizing

$$\pi(\bar{q}) = T\left(\frac{A - \bar{q}}{2}\right)\bar{q} - c\bar{q}$$

which yields

$$\bar{q}^*(A) = \frac{A}{2} - \frac{c}{T}$$

and hence

$$\pi^*(A) = \frac{(TA - 2c)^2}{8T}$$

as the optimal profit. Given $T = 3$ and $c = 50$ and given the uniform distribution of A, the expected profit on the side market is $E(\pi^*) = 194.2$.

Thus, if three firms operate on the main market they fare better in equilibrium than they would have done on the side market. However, four firms on the main market will individually conclude that the side market would have been the better option. Therefore, standard theory predicts $n = 3$.[10]

3.4 Results

3.4.1 Summary

Table 3.4 gives a brief overview of the average capacities and prices chosen as well as the average realized profits for both treatments – Baseline and Market Force – depending on the market size. The equilibrium values (ex-ante predictions) were also included as a reference point.

At first glance, there seems to be no substantial difference regarding the average capacity and price choices as well as the profits between the Baseline treatment and the Market Force treatment given market size. Testing for the differences using the Mann-Whitney U test verifies this first impression as all p-values were between 0.157 and 0.908.

In the following, we use the Mann-Whitney U test for between-subject comparisons and the Wilcoxon signed-rank test for within-subject comparisons if not stated otherwise.

[10]Note that while side market profits are risky (they range from $\pi^* = 145.0$ for A =53 to $\pi^* = 247.0$ for $A = 59$) they are guaranteed to exceed equilibrium profits on the main market for $n = 4$ and guaranteed to fall short of equilibrium profits for $n = 3$. Thus, risk aversion does not provide a motive for excess entry into the main market.

We observed a striking difference between the actual mean profits and the ex-ante equilibrium profits. In three-firm markets the mean profits were 44% lower than the equilibrium profits in the Baseline treatment and 34% lower in the Market Force treatment. In duopolies firms were on average only about half as profitable as predicted by the equilibrium. The gap between actual profits and ex-ante predicted profits is even larger in four-firm markets. There, average profits were only 10% of the equilibrium profits in the Market Force treatment and even negative in the Baseline treatment. All these differences were highly significant with p-values between 0.003 and 0.004.

The considerable difference between actual and equilibrium profits suggests that either capacities or prices were not chosen optimally or a combination of both.

In duopolies the mean capacities and prices were both lower than the ex-ante equilibrium values for both treatments. In the Baseline treatment the difference is not significant for capacities ($p = 0.167$) but highly significant for prices ($p = 0.005$). In the Market Force treatment both differences were significant. For capacities the p-values is 0.023 and for average prices it is 0.082.[11]

In both treatments the mean capacities were very close to the ex-ante prediction in three-firm markets ($p = 0.367$ and $p = 0.556$). However, the average posted prices were significantly lower than the equilibrium price ($p = 0.003$ for both treatments).

In quadropolies average capacities were significantly higher than the equilibrium capacities in both treatments ($p = 0.003$). In line with the theoretical predictions, average prices were significantly lower than the ex-ante equilibrium prices ($p = 0.021$ in the Market Force treatment and $p = 0.005$ in the Baseline treatment). Low average profits in these markets may result from prices that were either too low to cover capacity costs or still too high to sell all units produced. We find evidence for the latter explanation.

[11] The difference between average posted prices and equilibrium prices is significant for the first period ($p = 0.004$) but for periods two and three it is not significant anymore ($p = 0.142$ and $p = 0.625$).

Some firms set high capacities in conjunction with high prices. Especially first-period prices were on average far too high. This led to excess capacity and high losses.

N	Share	Treatment	Capacity	Price				Profit
				Period 1	Period 2	Period 3		
2	23%	Baseline	24.4	26.8	29.0	30.4	28.7	734.2
		Market Force	19.5	29.4	32.4	34.0	31.9	763.4
		Equilibrium	26.0	34.0	34.0	34.0	34.0	1352.0
3	55%	Baseline	15.0	22.6	22.9	23.3	23.0	221.7
		Market Force	15.0	23.4	23.8	24.2	23.8	251.8
		Equilibrium	14.6	25.4	25.4	25.4	25.4	382.8
4	22%	Baseline	10.6	19.8	19.7	19.4	19.6	-13.9
		Market Force	10.0	20.4	20.0	19.7	20.1	12.5
		Equilibrium	8.5	21.5	21.5	21.5	21.5	123.4

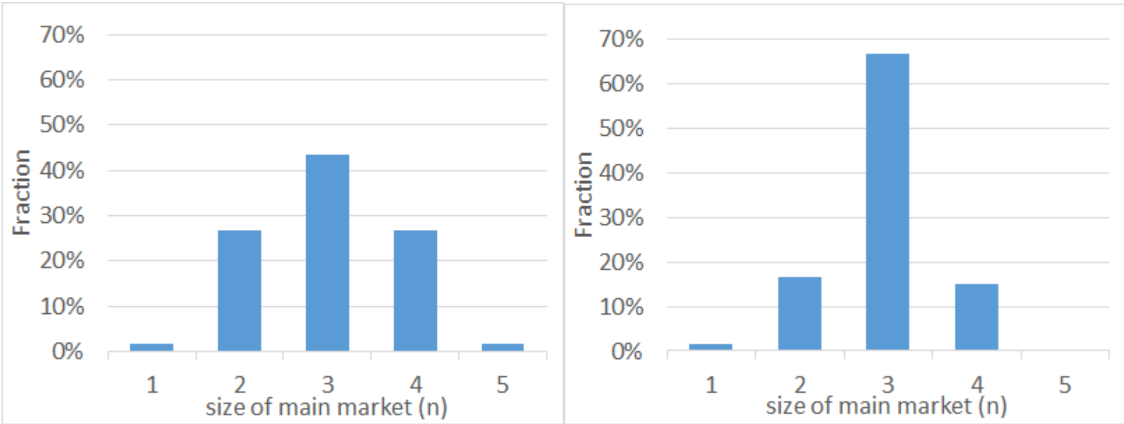
Table 3.4: Average capacity, price and profit in the main market over all ten rounds

Column two of Table 3.4 shows the distribution of market size. As we can see, markets with three firms occurred most frequently. Figure 3.3 gives a slightly more detailed picture of the market size distribution showing that there is a clear convergence towards $n = 3$ over time. In rounds one to five markets with three firms occurred 42% of the time. This share increased to 68% for rounds six to ten. In equilibrium, our parameterization implies that ex-ante expected profits in the side market were lower than expected profits with three firms in the main market. This is reversed for expected profits with four firms in the main market. Therefore, this convergence is in line with the theoretical predictions.

3.4.2 Baseline treatment

Capacity choice

By construction more firms lead to lower total market outputs in our setting. According to the theoretical predictions the total market output in duopolies is 52 units, in tripolies 43.8 units and in quadropolies 34 units. The data shows that in two- and three-firm markets total capacities per market averaged over all ten rounds were not significantly different from the equilibrium predictions. In duopolies the average to-



(a) Rounds 1 to 5 (b) Rounds 6 to 10

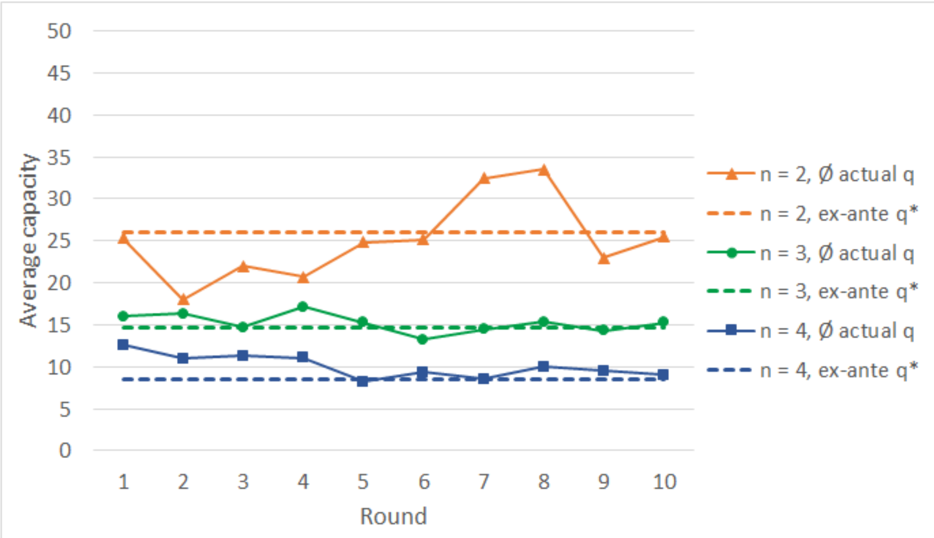
Figure 3.3: Distribution of market size in the Market Force treatment

tal output was 47.3 units. This was slightly but statistically not significantly lower than predicted ($p = 0.167$). In tripolies the actual total output was 43.4 units and not significantly different from the equilibrium prediction ($p = 0.367$). However, for quadropolies the average total market capacity (44.9 units) was significantly higher than the equilibrium value ($p = 0.012$). Using a Wilcoxon signed rank test shows that there was no systematic difference between the average market outputs of these three market sizes ($p > 0.23$).

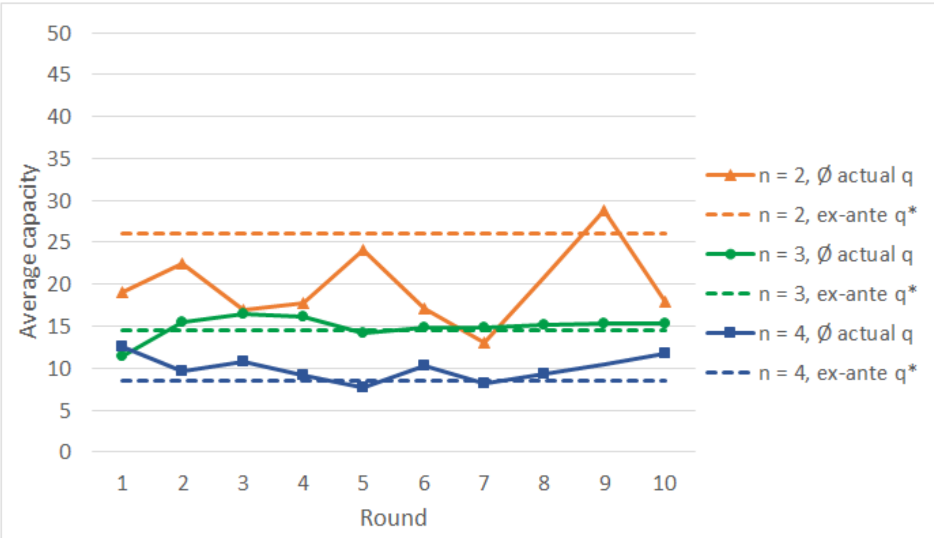
Considering rounds one to five separately from rounds six to ten, the average total market outputs were slightly closer to the equilibrium predictions in the second half of the experiment. The market outputs in the first five rounds were lower than the equilibrium value in duopolies (45.9 units) and higher in three- and four-firm markets (47.9 units and 44.6 units, respectively). In rounds six to ten total market output increased in duopolies to 48.8 units and decreased in tripolies as well as quadropolies to 43.3 units and 37.3 units. These differences, however, were not significant for duopolies and quadropolies ($p > 0.45$). For tripolies the difference was only marginally significant ($p = 0.099$).[12]

[12]The data shows that for quadropolies those independent observations with a total market output above 50 units in the first five rounds did not encounter a main market of size $n = 4$ in the later five rounds of the experiment.

Although the mean market outputs were closer to the ex-ante equilibrium outputs in rounds six to ten, they were still statistically not significantly different from each other for the different market sizes ($p > 0.2$).



(a) Baseline treatment



(b) Market Force treatment

Figure 3.4: Mean capacities vs. ex-ante equilibrium capacities over all ten rounds

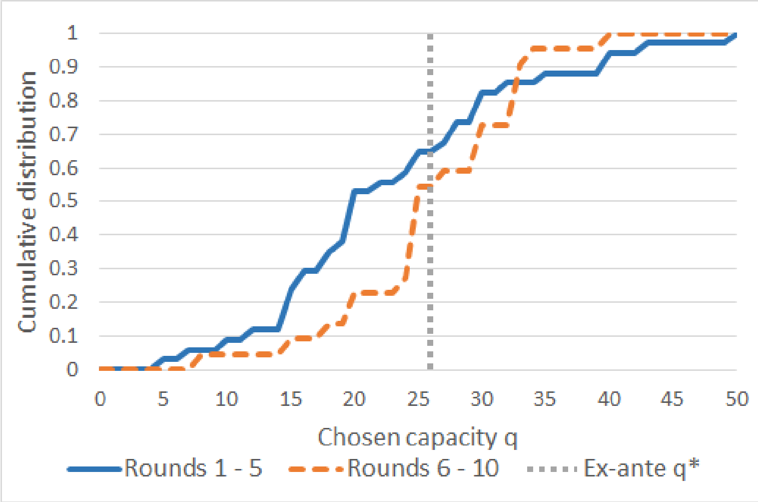
In contrast to the total market outputs, the firms' average capacity choices were significantly different between duopolies, tripolies and quadropolies with p-values between 0.003 and 0.004.

Figure 3.4a suggests that the capacity choices converged towards the ex-ante predicted equilibrium (q^*) given market size, especially in three- and four-firm markets. Visually, mean capacities in tripolies were close to the equilibrium value already in the first rounds. This impression is supported by testing for the difference between the ex-ante predicted capacities and the average chosen capacities, averaged over rounds one to five and over rounds six to ten, respectively. For tripolies, the difference is statistically not significant ($p = 0.224$ for rounds one to five and $p = 0.845$ for rounds six to ten). For duopolies and quadropolies the difference is significant in the first five rounds ($p = 0.041$ and $p = 0.009$, respectively) but not significant in the later five rounds ($p = 0.271$ and $p = 0.115$, respectively).

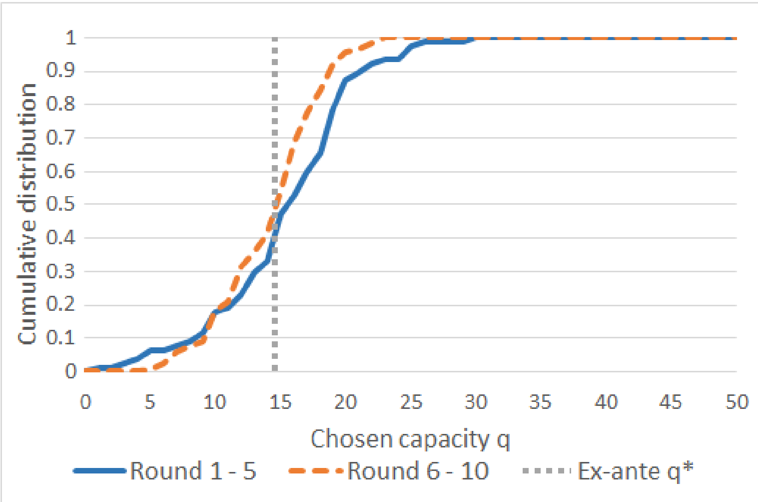
The cumulative distribution of the chosen capacities were slightly steeper around the equilibrium values for the second half of the experiment (see Figure 3.5). The share of firms choosing values close to the equilibrium capacity more than tripled in duopolies and almost doubled in markets with four firms.[13]

From Figure 3.4a capacity choices seemed more volatile in duopolies than in three-player and four-player markets. The average coefficient of variance within each independently observed market was indeed higher in the first five rounds for duopolies (0.44) than for tripolies (0.37) and quadropolies (0.29). However in the later five rounds, the average coefficient of variance reduced to 0.23 in duopolies, which is lower than for markets with three and four firms (both 0.30). When we compare the mean capacities of the different independent observations, the coefficient of variance was approximately 0.2 for all three market sizes in the first five rounds. This coefficient reduced to 0.13 for markets with three and four firms in the later five rounds. For duopolies, however, the coefficient did not change. This suggests that the variance between different independently observed duopolies did not change over time, but within a market participants' capacity choices assimilated.

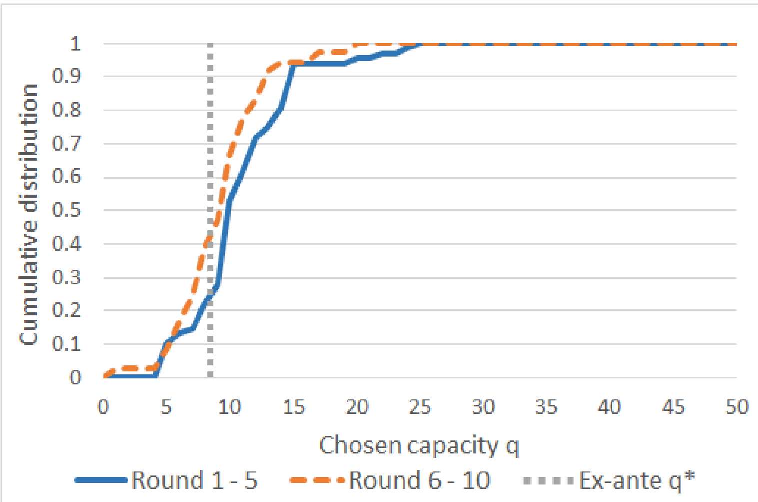
[13] In duopolies the share increased from 9% to 32% for choosing 25-27 units. For tripolies 18% of all firms chose a capacity of either 14 or 15 units in the first five rounds. This share remained almost constant over time. In quadropolies participants choosing eight or nine units increased from 13% to 22%.



(a) n = 2



(b) n = 3



(c) n = 4

Figure 3.5: Cumulative distribution of chosen capacities – Baseline treatment

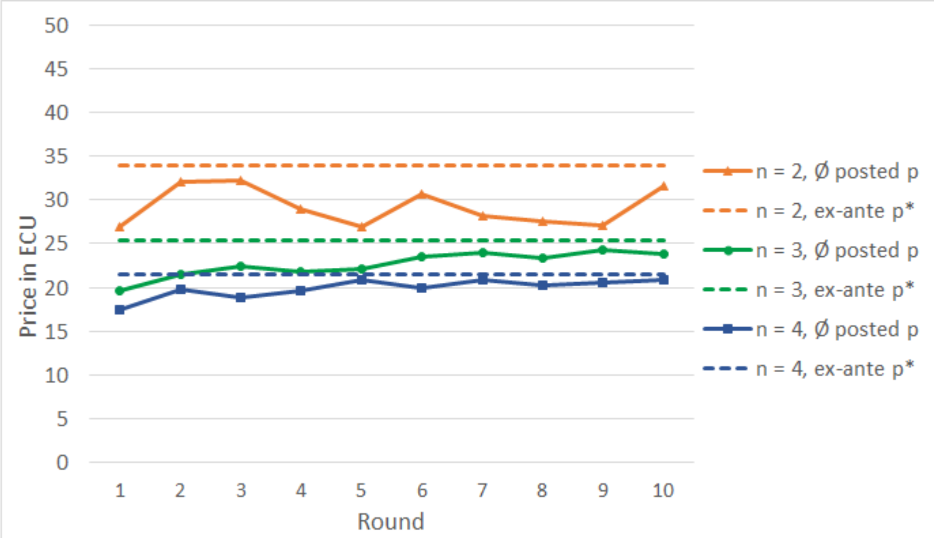
Price choice

After choosing capacities, firms on the main market chose their prices in three consecutive periods.

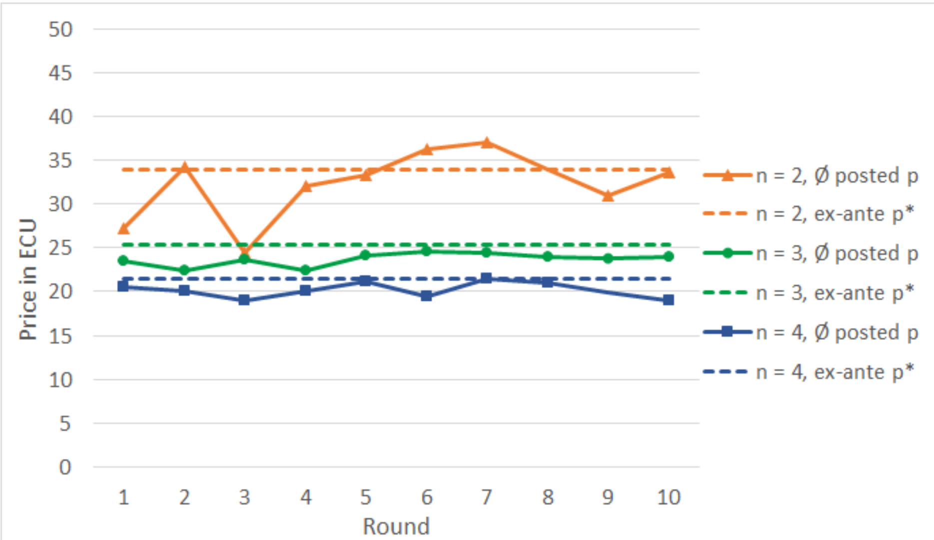
To get an overall understanding of the price choices, first look at the development of the average posted prices in Figure 3.6. The ex-ante predicted prices of a symmetric equilibrium (p^*) were included for comparison.

The average posted prices differed significantly between the two-, three and four-firm markets with p-values below 0.004. As Table 3.4 indicates, subjects chose prices that were on average lower than the equilibrium price regardless of the market size. Figure 3.6a offers support for this finding: Throughout all ten rounds mean posted prices were always below the ex-ante equilibrium prices for the corresponding market size. But for markets with three and four firms average prices seemed to converge towards the equilibrium. Considering the first five rounds separately from rounds six to ten, the average chosen prices were significantly lower than ex-ante predicted in the first half of the experiment regardless of market size. Testing for the difference between average posted prices and ex-ante equilibrium prices gives p-values of 0.005 in duopolies, 0.003 in tripolies and 0.008 in quadropolies. For the second half of the experiment, the differences remained significant for three- and four-firm markets ($p = 0.003$ and $p = 0.036$, respectively). For duopolies the difference was statistically not significant anymore ($p = 0.109$).

According to the theoretical predictions, lower average prices should correspond to higher average capacity choices. However, recall that in duopolies the average chosen capacities were significantly lower than ex-ante predicted in the first five rounds. Hence, average posted prices were too low in the first half of the experiment. In the second half of the experiment neither capacities nor prices differed significantly from the ex-ante predictions. In tripolies capacity choices were not significantly different to the ex-ante equilibrium values throughout the entire experiment. Therefore, we would expect average posted prices to be not significantly different from the ex-ante equilib-



(a) Baseline treatment



(b) Market Force treatment

Figure 3.6: Mean posted prices vs. ex-ante equilibrium prices over all ten rounds

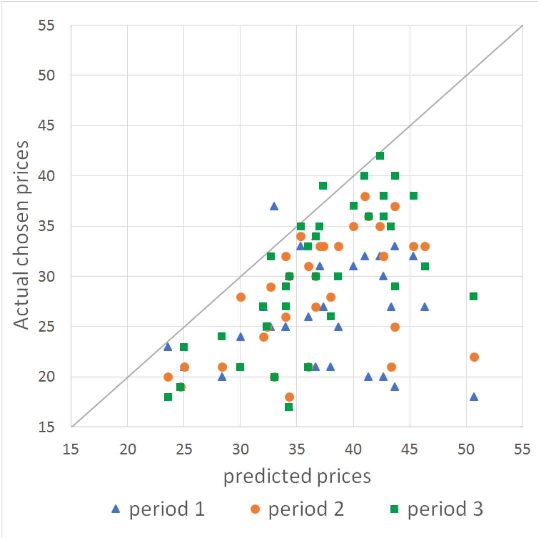
rium value as well. But as we have seen above, average posted prices were significantly lower than the ex-ante equilibrium prices. The same applies to quadropolies in the second half of the experiment. Since capacities in quadropolies were significantly higher than q^* in the first five rounds, average posted prices were in line with the theoretical predictions.

As ex-ante predicted prices were calculated based on ex-ante equilibrium capacities q^*, it is not the most accurate benchmark for analyzing whether participants set prices optimally in stage three after capacities were chosen. Therefore, we calculated an individual profit maximizing price for each firm based on their capacity choices and the total market outputs. Actual posted prices were then compared to these ex-post predicted equilibrium prices (see Figure 3.7). Each point in Figure 3.7 represents the price set by one firm in one of three periods of a round. The 45 degree line constitutes of all points for which the actual chosen price is equal to the ex-post predicted price. For all points above the 45 degree line the actual prices were larger than the predicted ones and for all points below the 45 degree line actual prices were below the predicted ones. Here again we differentiate between the first five rounds and the later five rounds of the experiment in order to capture price adjustment effects.

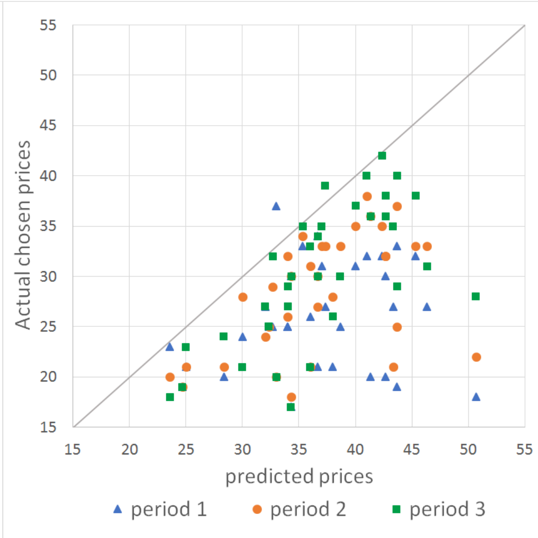
Figure 3.7 also supports the assertion that posted prices tended to be lower then ex-post predicted in two- and three-firm markets as most of the points fall below the 45 degree line. For quadropolies prices were more evenly distributed above and below the 45 degree line. A non-parametric test shows that for duopolies as well as tripolies actual prices were on average significantly lower than ex-post predicted prices throughout all ten rounds ($p = 0.004$ and $p = 0.003$, respectively). For quadropolies the difference between the average chosen prices and the predicted prices is not significant ($p = 0.262$).

In Figure 3.7 the points seemed to be less scattered in the second half of the experiment implying that actual prices were closer to the ex-post predictions in later rounds. This is affirmed when considering the standard deviation of the posted prices within markets. In duopolies the standard deviation reduced by about 50% between rounds one to five and rounds six to ten. For tripolies and quadropolies it reduced by 35% and 40%.[14]

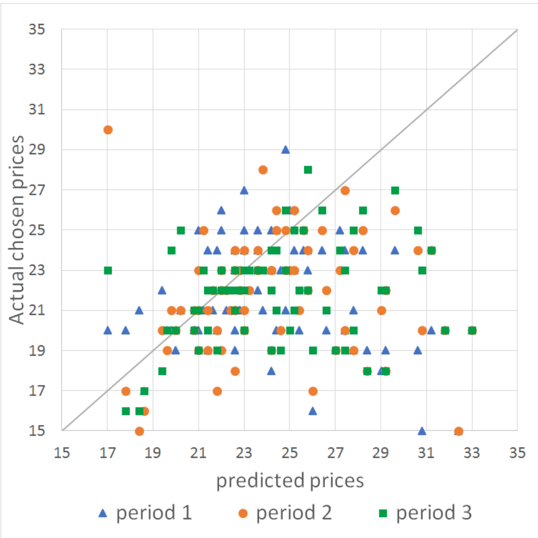
[14]For $n = 2$ the standard deviation reduced from 4.4 to 2.2, for $n = 3$ from 2.4 to 1.6 and for $n = 4$ from 2.2 to 1.3.



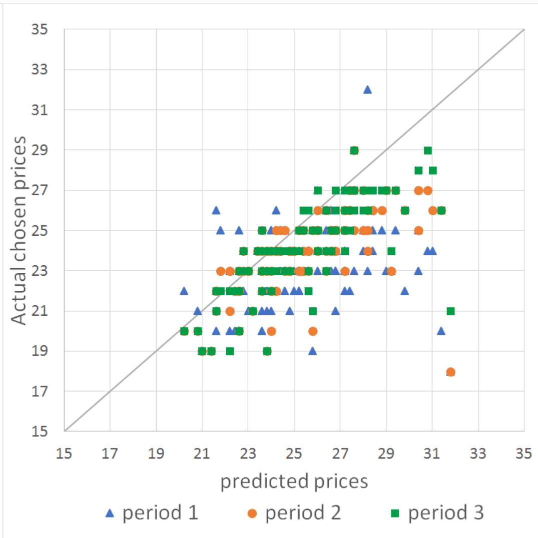
(a) n = 2, rounds 1 to 5



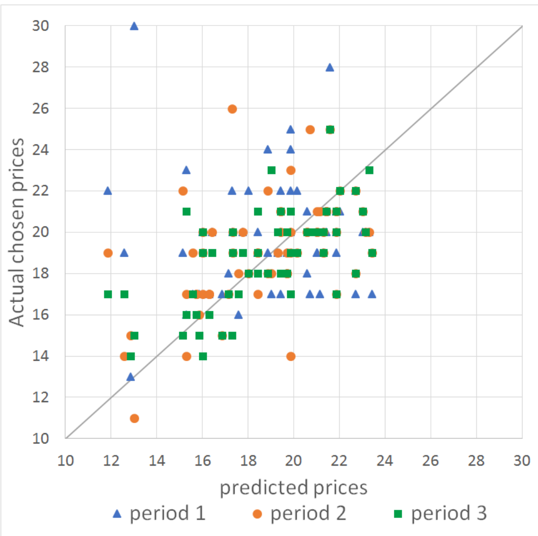
(b) n = 2, rounds 6 to 10



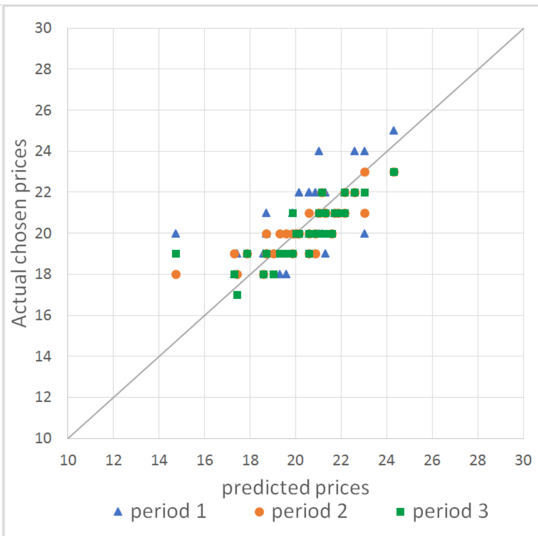
(c) n = 3, rounds 1 to 5



(d) n = 3, rounds 6 to 10



(e) n = 4, rounds 1 to 5



(f) n = 4, rounds 6 to 10

Figure 3.7: Actual prices vs. ex-post predicted prices – Baseline treatment

To test whether actual prices were closer to the predicted prices in the second half of the experiment, we first calculated the absolute differences between the actual average prices of each round and the predicted prices for each firm.[15] Then for each independent observation the average is taken across all these price differences for the first five rounds ($|\Delta P^1|$) as well as for the later five rounds ($|\Delta P^2|$). If the average prices were closer to the predicted prices in the second half of the experiment $|\Delta P^2|$ should be systematically smaller than $|\Delta P^1|$. For duopolies average posted prices were indeed significantly closer to the predicted prices in the later five rounds ($p = 0.035$). For tripolies and quadropolies the difference, however, was statistically not significant ($p = 0.126$ and $p = 0.201$, respectively).[16]

Using a similar method as described above, we tested whether posted prices indeed converged towards the ex-post predicted prices within a round. We first calculated the absolute differences between the posted prices and the ex-post predicted prices for each firm per period per round. Again, we used the absolute values since differences could be positive or negative and we did not want them to negate once we aggregate over all ten rounds. Then, for each independent observation we obtained a mean (absolute) price difference per period. Based on these three means (one for each period) we constructed a trendline with the average absolute difference as the dependent variable and periods as the independent variable. If posted prices were to converge to the predicted prices, the slopes should be systematically negative, which can be tested using a Wilcoxon signed-rank test.

With this method we obtained that for duopolies the slopes of the trendlines were significantly negative with a p-value of 0.009 and hence posted prices converged to-

[15]We used absolute values to measure the difference between actual prices and predicted prices as differences can be positive or negative.

[16]For quadropolies a within-subject comparison between rounds one to five and rounds six to ten is possible only to a limited extent as only four independent observations played a four-firm market in both halves of the experiment. There were ten independent observations in the first five rounds and six in the later five rounds in total.

wards the predicted prices within a round. For tripolies this effect was only marginally significant ($p = 0.092$) while for quadropolies it was not ($p = 0.790$).

In Figure 3.7a to Figure 3.7d, lower predicted prices appeared to materialize better while for high predicted prices the outcome fell short. Calculating the correlation between the average posted prices and ex-post predicted prices over all ten periods, we see that for duopolies the correlation coefficient is only 0.22. This indicates that participants with high predicted prices did not exploit their market power. They priced far below their predicted values. When only considering third-period prices the correlation coefficient increases to 0.45 indicating that within a round prices increased. For markets with three firms the correlation between average posted prices and predicted ones over all ten rounds was slightly negative (-0.19). This suggests that actual prices even decreased when predicted prices increase. As depicted in Figures 3.7c and 3.7d there were some outliers setting prices that were only half their predicted prices. If only considering third-period prices the correlation becomes positive (0.35 for all ten rounds and 0.69 for rounds six to ten only). This is in line with our finding that prices converged towards the equilibrium prices. As expected by looking at Figures 3.7e and 3.7f the correlation between actual and predicted prices is very high for quadropolies with a correlation coefficient of 0.9.

Profits

To evaluate a firm's market performance, we shall have a closer look at the profits. As expected average profits were significantly different between the three market sizes ($p < 0.004$).

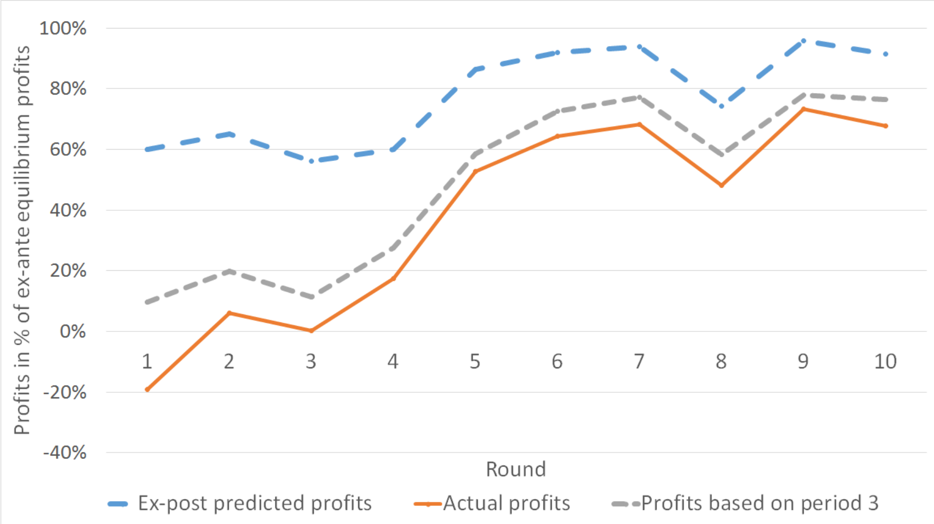
Figure 3.8a gives an overview of how average profits developed over time. For better comparability, we used the ex-ante predicted profits of each market size as a benchmark (100%). We calculate the ex-post predicted profits, the actual profits and profits extrapolated based on third period prices and sales as a share of the ex-ante profits.[17] Normalizing the profits like this also allows us to aggregate over different market sizes.

Average profits – actual, third period and ex-post predicted – were always lower than ex-ante equilibrium profits (see Figure 3.8a). The ex-post predicted profits were lower than the ex-ante equilibrium profits which indicates that capacities were not chosen optimally. This is especially evident in the first four rounds of the experiment, which corresponds to our finding that average total market outputs were too low for duopolies and too high for quadropolies in the first five rounds. Figure 3.8a also suggests that average prices were set too low as there is a clear gap between the ex-post predicted profits and the realized profits. Although price adjustments within a round had an positive effect on profits – "Profits based on period 3" were always above the "actual" profits –, average prices in the third period were still too low.

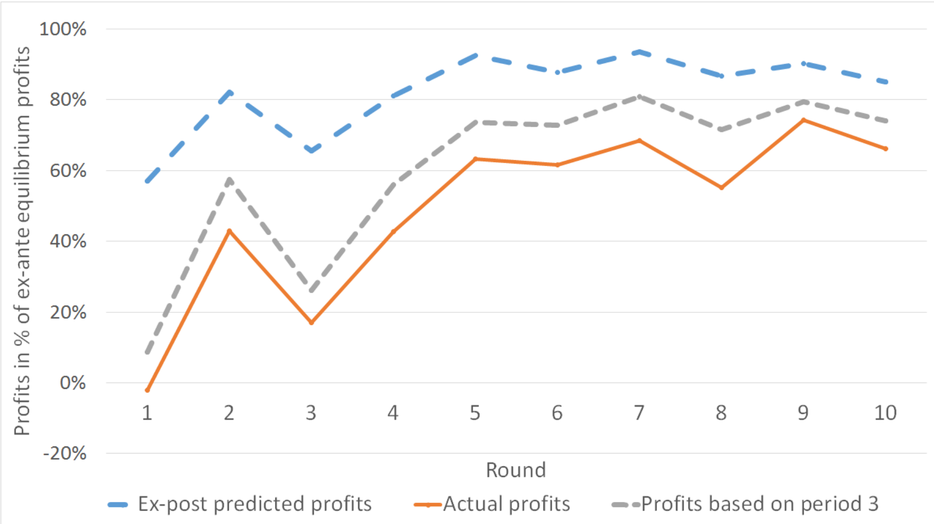
As Figure 3.8a shows average profits increased rather quickly. While firms were unprofitable in the first three rounds, the realized profits were over 60% of the ex-post equilibrium profits from round five onwards. The gap between predicted profits and the actual profits closed over time implying that participants prices choices approached the ex-post equilibrium prices.

As profits were determined by the combination of capacities and prices, the question for which capacities profits were the highest might occur. In Figure 3.9 we illustrate the average profits given chosen capacities. For each market size we calculated the median over all rounds across all sessions of the Baseline treatment. We then divided the chosen capacities into three percentiles: the lowest 33%, the highest 33% and the remaining 34% around the median. For each of the three capacity intervals Figure 3.9 depicts (i) the average ex-post predicted profits, which were calculated based on the chosen capacity and the ex-post predicted prices mentioned above, (ii) the actual

[17] We calculated the "Profits based on period 3" by simply assuming that the capacities and prices in periods one and two were the same as what was chosen by the subjects in the third period of each round. We do this to measure the effects of the price adjustments within a round.



(a) Baseline treatment



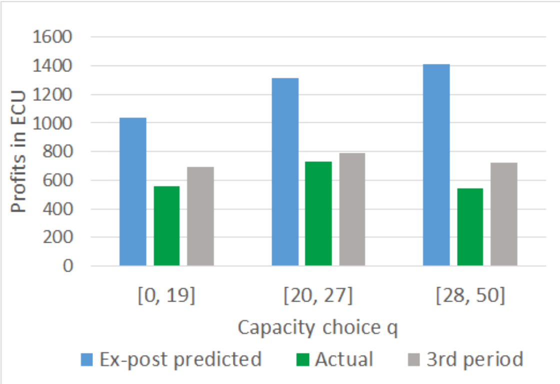
(b) Market Force treatment

Figure 3.8: Development of profits across all ten rounds

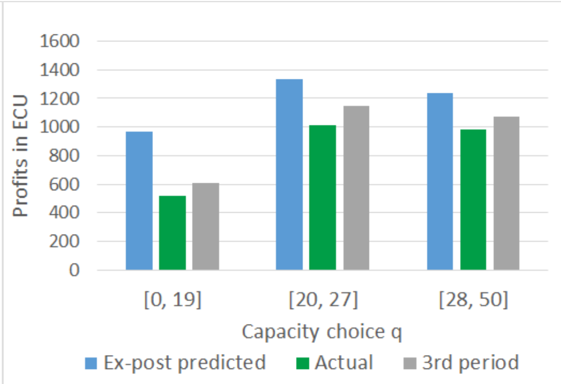
profits that were observed during the experiment, (iii) the profits based on the third period outcomes.[18]

Predicted profits were highest for capacities around the median and thus around the equilibrium capacities in almost all graphs of Figure 3.9. Only in Figure 3.9a predicted profits were increasing with capacity. Having a closer look at the data, we can see that firms with a very high capacity had market counterparts with very low capacities. Hence, these market counterparts were capacity constrained. However, those firms with

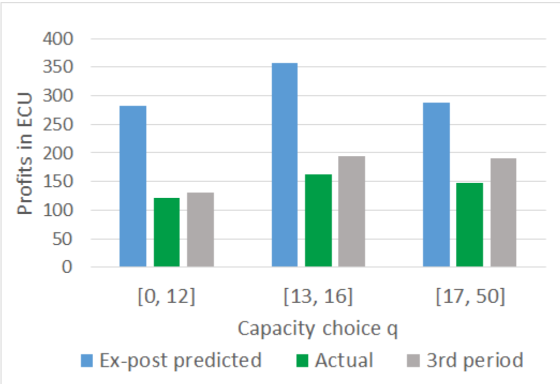
[18]For the Baseline treatment the median is very close to the equilibrium capacities in markets with two and three firms. There, the equilibrium capacities were in the interval around the median. For $n = 4$ the equilibrium value was at the border between the lower 33% and the median interval.



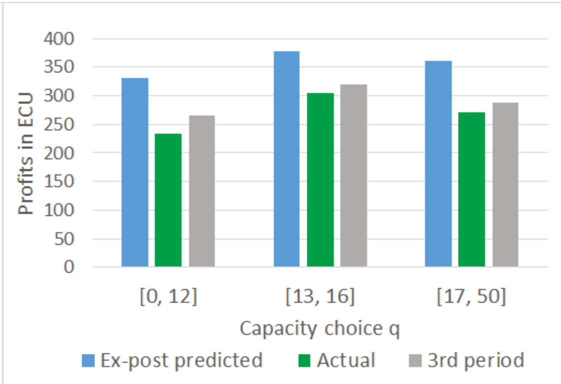
(a) n = 2, rounds 1 to 5



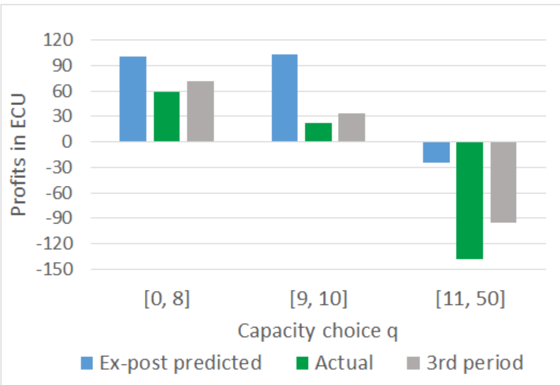
(b) n = 2, rounds 6 to 10



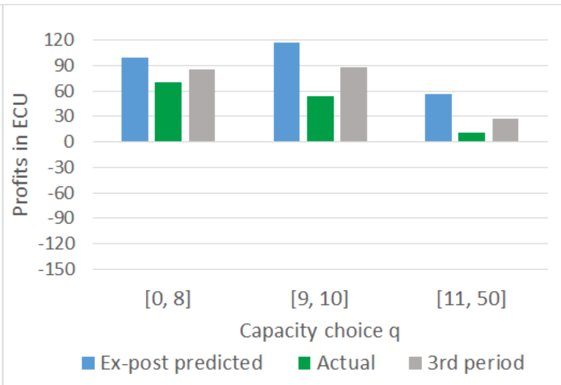
(c) n = 3, rounds 1 to 5



(d) n = 3, rounds 6 to 10



(e) n = 4, rounds 1 to 5



(f) n = 4, rounds 6 to 10

Figure 3.9: Average actual and predicted profits given capacities – Baseline treatment

high capacities did not exploit their market power as they set prices that were much lower than what they could have charged. This lead to a very notable gap between the predicted profits and actual profits.

Not only were the predicted profits highest for capacities close to the equilibrium capacities, but the actual profits were also highest in those intervals. The difference between the average profits for capacities around the median and the remaining capacity levels was significant for duopolies ($p = 0.041$) and marginally significant for tripolies ($p = 0.092$). For quadropolies average realized profits were significantly higher in the low capacity interval than for higher capacity levels ($p = 0.03$).

We have addressed the very low profits in four-firm markets before. Figure 3.9f shows that for high capacity choices average profits were highly negative. Posted prices were too high, especially in the first period. Hence, firms were left with a lot of excess capacity.

Independent of the market size, an increase in profits given capacities could be observed over time when comparing the actual profits in the first five rounds to the later five rounds of the experiment. For duopolies and tripolies this increase was significant ($p = 0.022$ and 0.003), while for quadropolies it was at the boarder of marginal significance ($p = 0.1$). Profits in the second half qwew much closer to the predicted profits which corresponds to the findings regarding capacity choice and prices stated above.

For all three market sizes profits based on third period outcomes were significantly higher than actual profits showing that price adjustments within a round had a positive effect on profits ($p-$values < 0.004).

To sum up, total market outputs did not differ significantly between the different market sizes. However, mean capacities and prices of the three market sizes under consideration were significantly different. In duopolies average capacities were lower than predicted by the ex-ante equilibrium. Given capacity choices, prices were too low on average, especially for firms with high ex-post predicted prices. Most of them did not make use of their market power. In tripolies mean capacities were already very close to the ex-ante equilibrium values in the first few rounds. Average posted prices were slightly but significantly lower than ex-post predicted prices. In four-firm markets mean capacities were significantly higher than in the ex-ante equilibrium and prices were on average higher than ex-post predicted. This led to very high losses for some firms.[19] Overall capacities as well as prices converged towards the equilibrium values over time. Price adjustments were also observed within each round in the sense that third-period prices were closer to the ex-post predicted prices. The profitability of firms increased substantially over time. It seems that although participants did not actively choose to be on the main market in the Baseline treatment, they very quickly adjusted to the market situation.

3.4.3 Market Force treatment

As mentioned above the main and only difference between the Baseline treatment and the Market Force treatment was that subjects in the Market Force treatment were asked to choose whether they want to enter the main market or be on the side market. In case they were already on the main market, they had to decide whether they wish to stay on it. We will refer to this choice as their entry decision hence on.[20]

By comparing the decisions made in this treatment to the Baseline treatment, we wish to find out whether (i) subjects self-selected into either main or side market and (ii) these selection effects led to systematically "better" choices.

Entry decision

Before turning to the analysis of subjects capacity and price decisions, let us first investigate whether there is a self-selection effect. Figure 3.10 depicts the distribution of how many rounds subjects operated on the main market. The figure suggests that there is self-selection once participants were given the chance to choose: The mode is very clearly at ten rounds. 27% of all subjects were on the main market in ten out of ten rounds of the experiment while 7% of all participants never entered the main market but remained on the side market during the entire session.

[19]Losses of up to 700 ECUs were observed.
[20]Recall that the entry balance requirement was 300 ECUs.

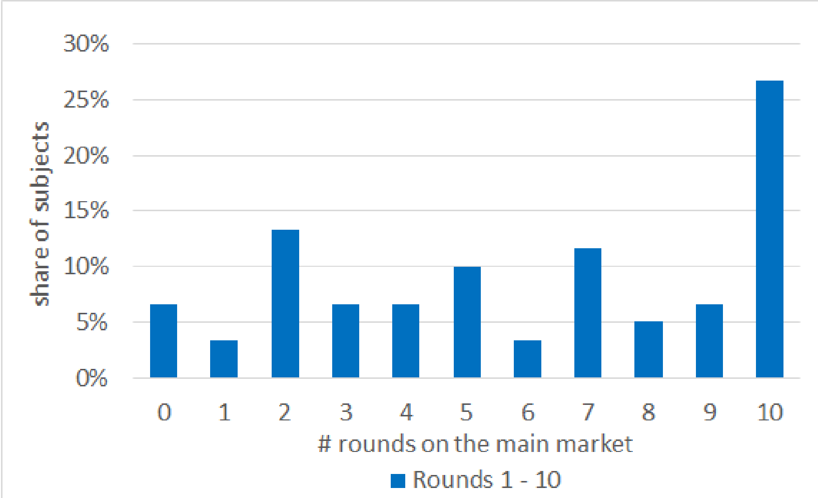


Figure 3.10: Share of total market entries for all ten rounds

The self-selection becomes more evident when considering the entry decisions in rounds one to five and rounds six to ten separately (see Figure 3.11). The shape of the distribution in rounds six to ten has a u-shape with the most subjects either being on the main market for all five rounds or not at all. While in the first five rounds the distribution forms a small u-shape between two rounds and five rounds. This suggest that while some experimentation took place in the earlier rounds, participants "specialized" themselves over time and either operated on the main market or stayed on the side market.

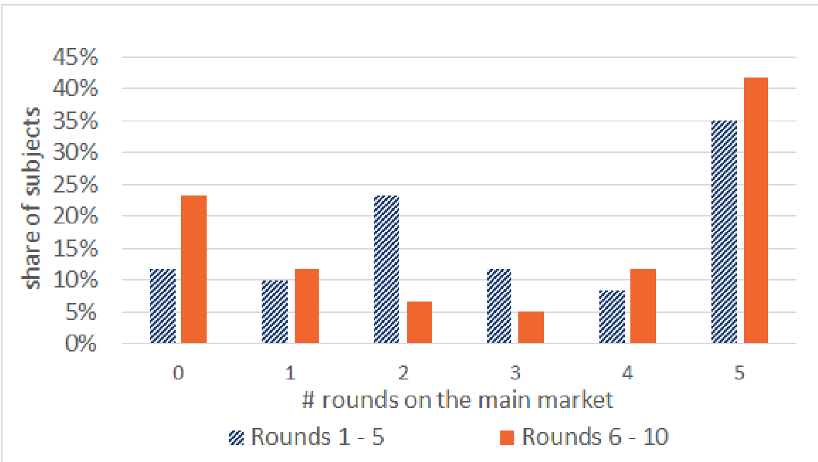


Figure 3.11: Share of total market entries for rounds 1–5 and 6–10

We defined and calculated an index to quantify the magnitude of this specialization for each group of five participants. The specialization index (SI) takes on values between 0 and 1 and is defined such that 1 is equal to total specialization. As standard theory predicts $n = 3$ in equilibrium for our setting, $SI = 1$ implies that three out of five firms were on the main market and the remaining two firms operate on the side market for all ten rounds of the entire experiment. For $SI = 0$ each of the five firms would be on the main market in six out of ten rounds.

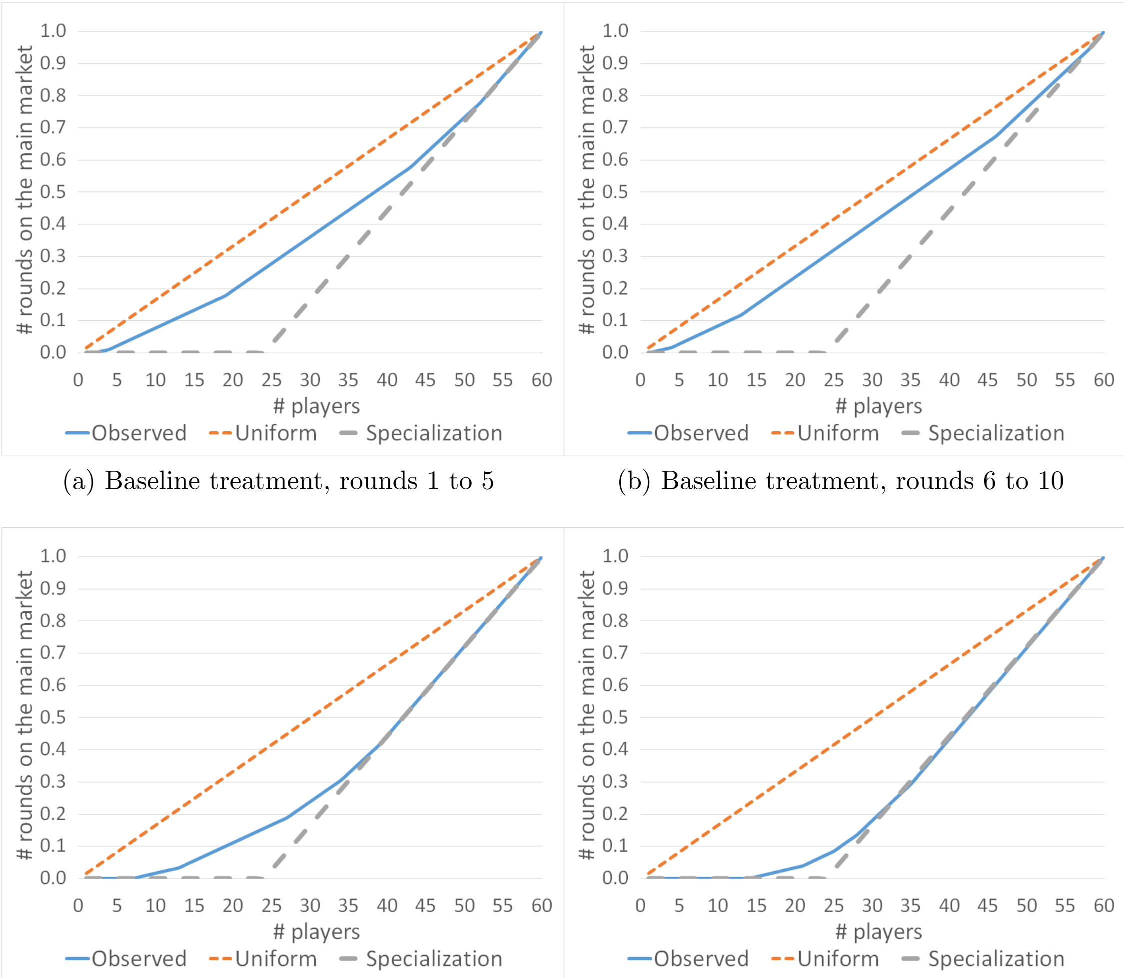
Calculating the index for all independent observations, we find clear evidence for specialization in the Market Force treatment (see Table 3.5). Table 3.5 also shows that specialization was increasing over time. In the later five rounds of the experiment 67% of all markets had an index value of 1. This offers support for the visual impression that there was self-selection into the main market from Figure 3.11.

This selection effect is also illustrated in Figure 3.12. There the cumulative distribution of rounds participants were on the main market is very close to the cumulative distribution of total specialization. In rounds six to ten, it almost completely overlaps with the distribution for $SI = 1$ (see Figure 3.12d).

As expected, we did not find systematic specialization in the Baseline treatment since participants were randomly allocated to either main or side market (see Table 3.5). Therefore the cumulative distribution is closer to the uniform distribution (see Figure 3.12a and 3.12b).

	Market Force		Baseline	
	Rounds 1-5	Rounds 6-10	Rounds 1-5	Rounds 6-10
$\varnothing SI$	0.81	0.95	0.38	0.26
SI_{min}	0.42	0.77	0	0
SI_{max}	1	1	0.78	0.58
$SI = 1$	25%	67%	0%	0%

Table 3.5: Summary of the specialization index



(a) Baseline treatment, rounds 1 to 5

(b) Baseline treatment, rounds 6 to 10

(c) Market Force treatment, rounds 1 to 5

(d) Market Force treatment, rounds 6 to 10

Figure 3.12: Cumulative distribution of rounds subjects operated on the main market

Does this high degree of specialization have a systematic effect on capacity and price choices and how does it affect the profitability of firms?

Capacity choice

First, consider the total market output. The average total market outputs were 40.2 units in duopolies, 45 units in tripolies and 40.4 units in quadropolies. Recall that the ex-ante equilibrium market outputs are 52 units, 43.8 units and 34 units, respectively. Hence, as before in the Baseline treatment, average total outputs were significantly lower than predicted by the ex-ante equilibrium in duopolies ($p = 0.023$), and significantly higher than predicted in quadropolies ($p = 0.009$). For tripolies the total outputs were almost equivalent to the ex-ante predictions with a p-value of 0.556. Comparing average total outputs in both treatments, we did not find a significant difference regardless of market size ($p > 0.19$).

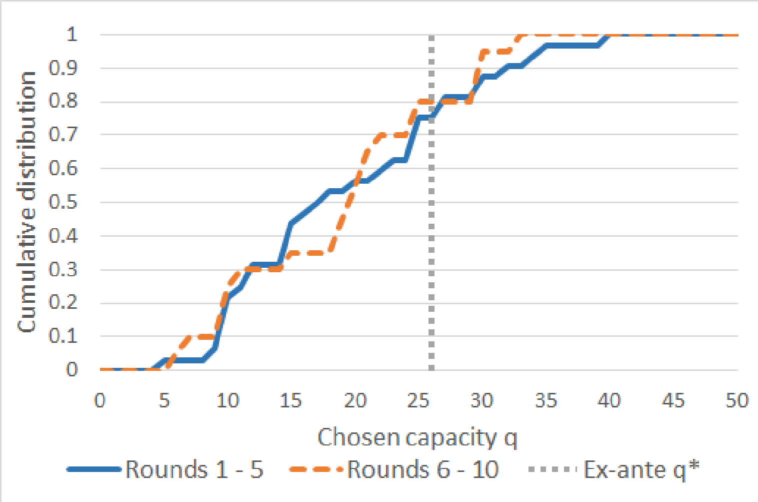
Separately looking at the total market capacities in rounds one to five and rounds six to ten, the total market output neither increased nor decreased significantly for either market size ($p > 0.4$).

Considering comparative statics, the mean market outputs did not differ systematically between two- and three-firm markets ($p = 0.221$) as well as between two- and four-firm markets ($p = 0.965$). This is equivalent to our findings in the Baseline treatment. However, by contrast, the difference between mean market outputs of three- and four-firm markets was marginally significant ($p = 0.0917$).

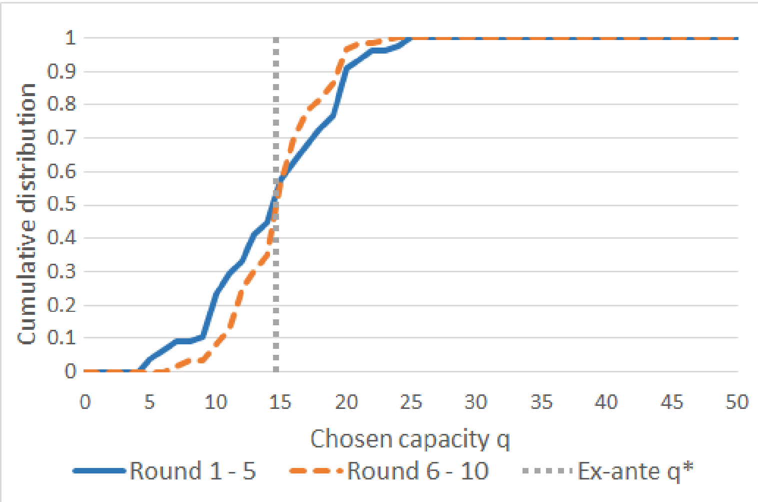
As with the findings in the Baseline treatment, the mean capacities differ significantly between two-, three- and four-firm markets ($p < 0.037$).

In Figure 3.4b the development of the mean capacities over time appears to be similar to what was observed in the Baseline treatment (see Figure 3.4b) in the first half of the experiment: Mean capacities were significantly lower than ex-ante predicted in duopolies ($p = 0.016$), not significantly different in tripolies ($p = 0.969$) and significantly higher in quadropolies ($p = 0.011$). In the second half of the experiment, the difference remains significant for two- and four-firm markets ($p = 0.036$ and $p = 0.090$, respectively) and insignificant in tripolies ($p = 0.367$). Whereas, in the Baseline treatment mean capacities were not significantly different from the equilibrium predictions in the later half of the experiment for all three market sizes. However, there is no significant difference in mean capacities between both treatments ($p > 0.2$).

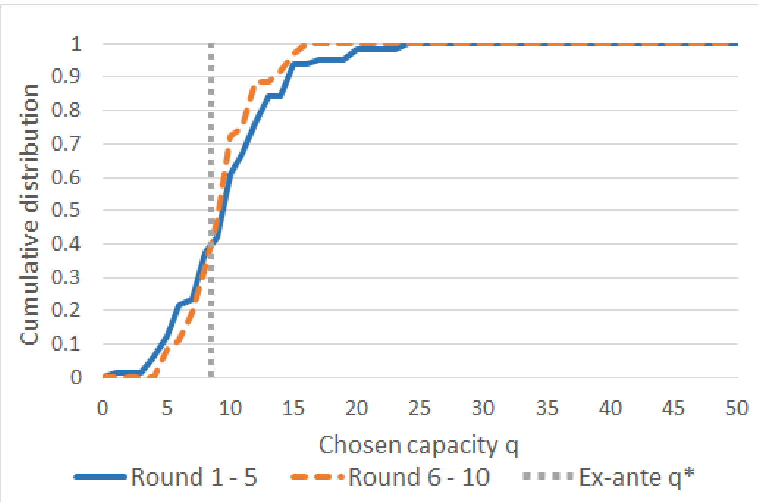
In Figure 3.13 we can also see that the cumulative distribution between rounds one to five and rounds six to ten were quite similar. For markets with three and four firms the distribution curve for the later five rounds appears to be slightly steeper around the ex-ante equilibrium value. In fact, the share of firms choosing either 14 or 15 units



(a) n = 2, Market Force treatment



(b) n = 3, Market Force treatment



(c) n = 4, Market Force treatment

Figure 3.13: Cumulative distribution of chosen capacities – Market Force treatment

in tripolies increased from 17% in the first five rounds to 27% in the later five rounds. For quadropolies we observed a similar increase. 19% of firms chose either 8 or 9 units in rounds one to five and 28% in rounds six to ten.

In total, capacity choices in the Market Force treatment were not significantly different from the Baseline treatment. For duopolies the mean capacities were even lower when entry was endogenous. However, this difference was not significant.

Price choice

As in the Baseline treatment price choices differed significantly between the different market sizes (p-values between 0.003 and 0.004).

In Figure 3.6, which shows the development of the average posted prices over time, prices for all three market sizes seemed to converge towards the ex-ante equilibrium values.

Recall that average capacities were significantly lower than ex-ante predicted in duopolies. In tripolies there was no significant difference between predictions and average chosen capacities and in quadropolies actual capacities were significantly higher than equilibrium values. Hence, based on the theoretical predictions, we would expect average prices to be higher than ex-ante predicted in duopolies, not significantly different from predictions in tripolies and significantly lower than ex-ante equilibrium prices in quadropolies.

However, for duopolies average posted prices were lower than ex-ante predicted in rounds one to five ($p = 0.083$). In rounds six to ten the difference was not significant anymore ($p = 0.8$). In tripolies average actual prices were significantly lower than ex-ante predicted throughout all ten rounds. Although average prices seemed to converge to the equilibrium prices the differences remained significant ($p = 0.013$ for rounds one to five and $p = 0.021$ for rounds six to ten). Thus, average prices were lower than expected in duopolies and tripolies. For quadropolies mean posted prices were significantly lower in the first half of the experiment ($p = 0.006$) which is in line with the theoretical predictions, as capacities there surpassed equilibrium predictions significantly. For the second half, prices converged towards the equilibrium prices and the difference was not significant anymore ($p = 0.205$). As average capacities remained significantly lower than predicted in the second half, average chosen prices were higher than we would expect according to theory.

When comparing the actual posted prices to the ex-post instead of the ex-ante predicted prices, we obtain the same results. Prices were significantly lower in duopolies and tripolies and not significantly different in quadropolies.[21] Given average capacities chosen this implies that prices were set too low in markets with two and three firms while they were too high in markets with four firms. This result is equivalent to what we have found in the Baseline treatment.

The visual impression from Figure 3.14 is similar to the Baseline treatment before: (i) Actual posted prices seemed closer to the ex-post predicted prices in the second half of the experiment, and (ii) posted prices converged towards the equilibrium prices within a round.

Using the same method as described before, we test for the significance of the observed convergence. And indeed for duopolies and tripolies average posted prices were significantly closer to the equilibrium prices in the second half of the experiment. But for duopolies the difference between first and second half was only marginally significant ($p = 0.076$). In contrast to the Baseline treatment where the posted prices were not significantly closer to the predictions in tripolies, the difference between the first five rounds and the later five rounds was significant in the Market Force treatment ($p = 0.025$). For quadropolies the result was similar to the Baseline treatment: The difference was not significant ($p = 0.281$).

[21]Considering the first five rounds separately form the later five rounds $p = 0.004$ in the first half and $p = 0.022$ in the second half of the experiment in duopolies. In tripolies the $p-$values were 0.007 and 0.013 respectively and in quadropolies they were 0.541 and 1.

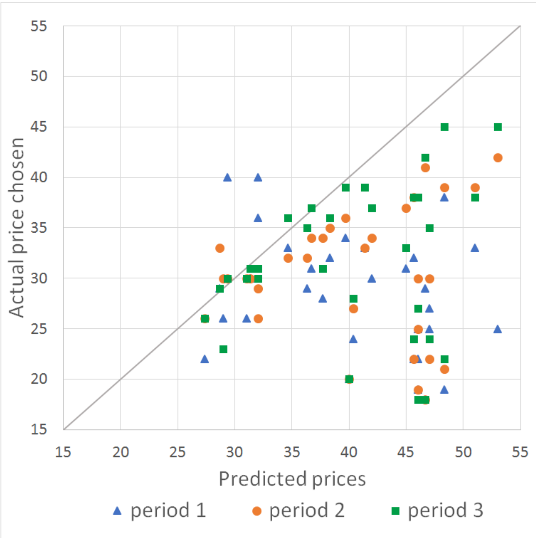
Equivalent to what we observed in the Baseline treatment prices converged towards the equilibrium prices in two- and three-firm markets ($p = 0.007$ and $p = 0.092$, respectively) while they did not in quadropolies ($p = 0.530$).[22]

Comparing the average prices between the Baseline treatment and Market Force treatment, no significant difference could be found for all three market sizes ($p > 0.19$).

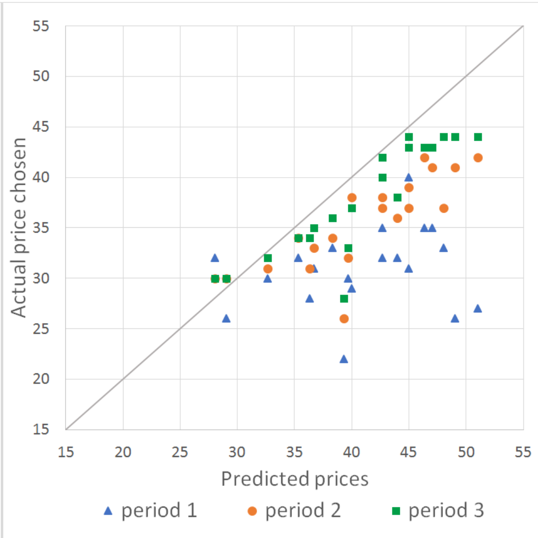
Similar to the Baseline treatment, participants in the duopoly market did not exploit their market power. The correlation coefficient between posted prices and ex-post predicted prices were even slightly negative (-0.05) which indicates that for higher predicted prices posted prices tended to be even lower. When only considering actual prices in the third period the correlation coefficient was positive but still low (0.18). Recall that in the Baseline treatment the coefficient was 0.45. This implies that higher predicted prices materialized better in the Baseline treatment. As opposed to the Baseline treatment where the correlation was negative (-0.19) in three-firm markets, posted prices and ex-post predicted prices were strongly correlated with each other (0.8) in the Market Force treatment. When comparing Figures 3.14c and 3.14d to Figures 3.7c and 3.7d we can see that the points were less scattered in the Market Force treatment. Especially for higher predicted prices some firms set very low actual prices in the Baseline treatment. For quadropolies however, the correlation (0.69) was slightly lower in the Market Force treatment than in the Baseline treatment (0.9). As Figure 3.14e show there were some outliers charging only half their predicted prices.

As with the capacity choices, we did not observe a significant difference between the price choices in the Market Force treatment and the Baseline treatment. Only for tripolies, convergence to the ex-post equilibrium prices was slightly stronger and posted prices were more strongly correlated to the predicted prices in the Market Force treatment .

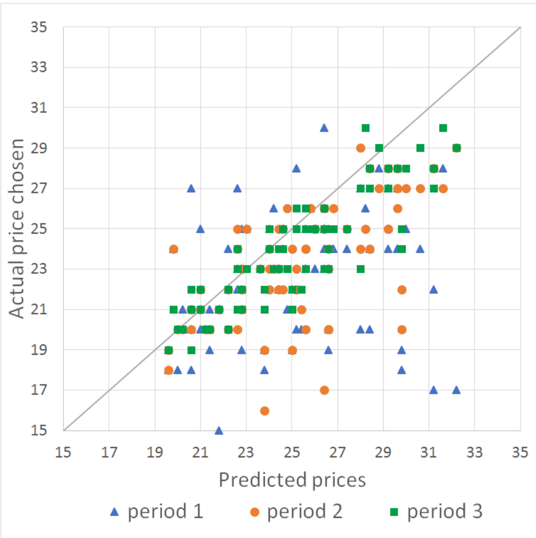
[22]Here we used the same method as described before in section 3.4.2.



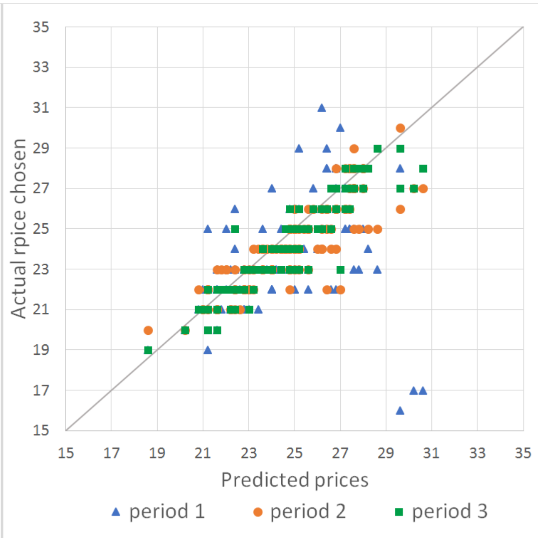
(a) n = 2, rounds 1 to 5



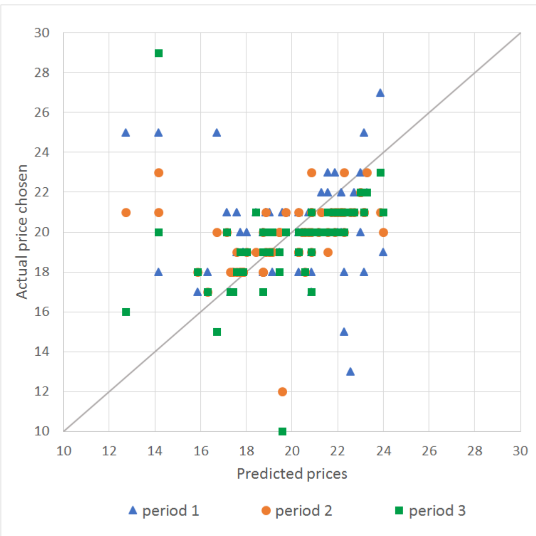
(b) n = 2, rounds 6 to 10



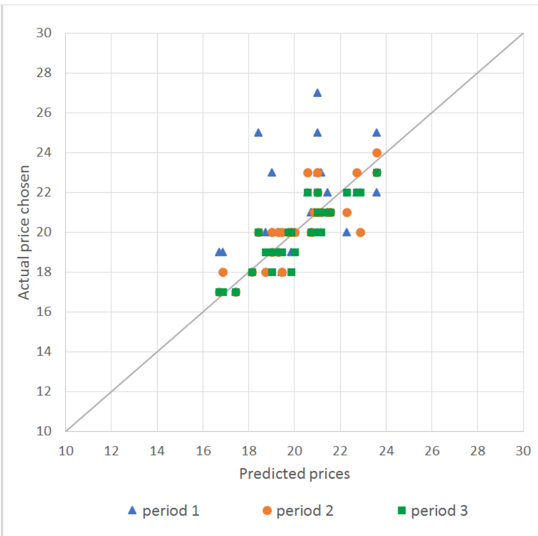
(c) n = 3, rounds 1 to 5



(d) n = 3, rounds 6 to 10



(e) n = 4, rounds 1 to 5



(f) n = 4, rounds 6 to 10

Figure 3.14: Actual prices vs. ex-post predicted prices – Market Force treatment

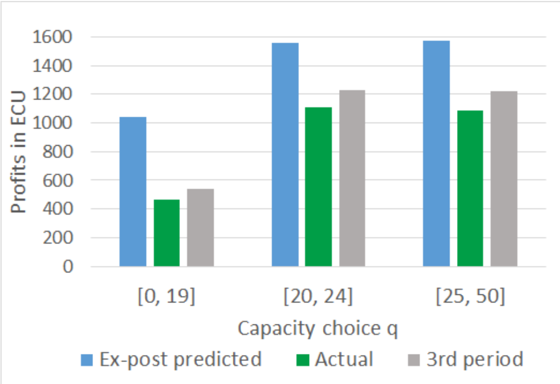
Profits

Figure 3.8b gives an overview of how profits developed over time. Similar to what we have observed before in Figure 3.8a for the Baseline treatment, ex-post predicted profits were always lower than the ex-ante predicted profits. This corresponds to the fact, that capacities were not chosen optimally as average capacities were significantly lower than ex-ante predicted for duopolies and higher for quadropolies. Further, profits based on the third period outcomes only were on average higher than the actual profits, again indicating that price adjustments within a round had a positive effect on profits. In fact, the difference between actual profits and profits extrapolated based on third period outcomes was significant for all three market sizes ($p < 0.004$).

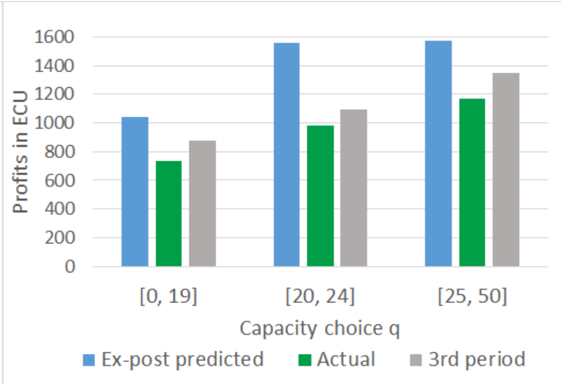
Compared to the Baseline treatment profits were higher in the first rounds and seemed to increase faster within the first five rounds. While firms were on average unprofitable in the first round, realized profits increased immediately in the second round to 50% of the ex-post predicted profits. In round five actual profits increased to approximately 70% of the ex-post predicted profits. The development in the later five rounds was comparable between both treatments. Overall, profits did not significantly differ between the two treatments for all three market sizes ($p > 0.2$).

Disaggregating the data and looking at each market size individually, profits significantly differed between the three market sizes ($p < 0.004$). Considering profits in the first half and the second half of the experiment separately shows that average profits in the second half are significantly higher. For duopolies and tripolies $p < 0.05$ and for quadropolies $p = 0.06$.

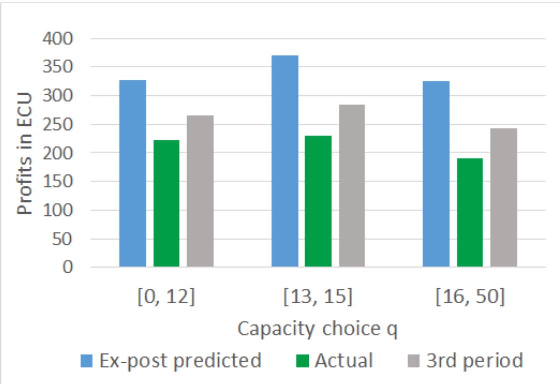
Figure 3.15 depicts the average profits given capacity choices. We again calculated the median, the lower 33% percentile as well as the upper 33% percentile of the chosen capacities for each market size. In this treatment the ex-ante equilibrium capacity were in the upper 33% percentile for duopolies and in the interval around the median for tripolies and quadropolies.



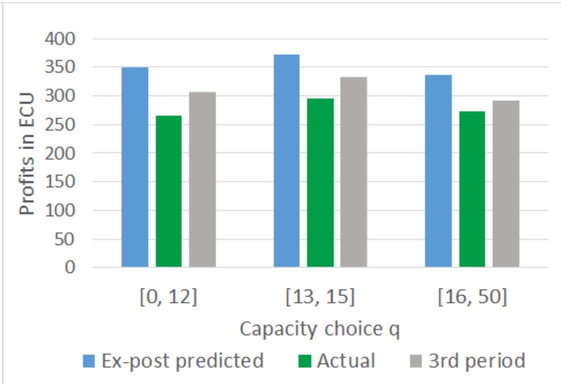
(a) n = 2, rounds 1 to 5



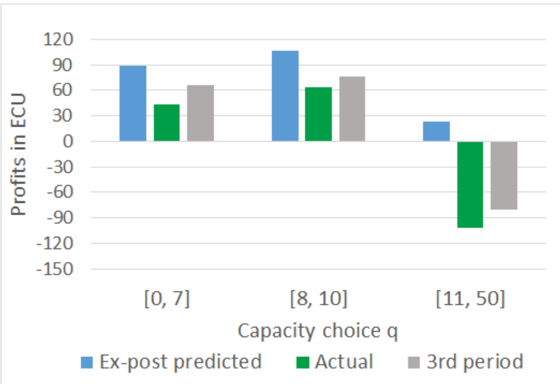
(b) n = 2, rounds 6 to 10



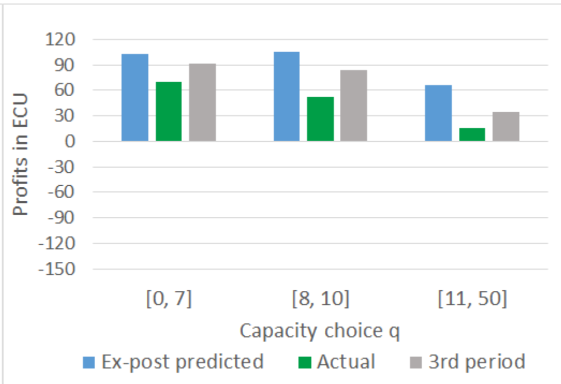
(c) n = 3, rounds 1 to 5



(d) n = 3, rounds 6 to 10



(e) n = 4, rounds 1 to 5



(f) n = 4, rounds 6 to 10

Figure 3.15: Average actual and predicted profits given capacities – Market Force treatment

In contrast to the Baseline treatment, the difference between realized profits and predicted profits were not as prominent in the first five periods for all three market sizes. As before in the Baseline treatment, Figure 3.15 conveys the impression that average profits were highest for the interval around the ex-ante equilibrium capacity. However, the difference in average profits between the interval with the ex-ante equilibrium value and the remaining two capacity intervals was only significant for quadropolies ($p = 0.006$). For duopolies the difference was not significant ($p = 0.234$) and for tripolies it was only at the boarder of significance ($p = 0.1$).

Overall, we do not find strong evidence for a significantly better market performance in the Market Force treatment compared to the Baseline treatment.

3.4.4 Side market

On the side markets where individuals were not competing against each other the average capacities and prices chosen were almost equivalent to the equilibrium values resulting in an average profit which was very close to the equilibrium profit for both treatment types (see Table 3.6).

Treatment	Capacity	Price				Profit
		Period 1	Period 2	Period 3		
Baseline	12.4	22.7	22.9	23	22.9	214.2
Market Force	12.8	22.8	22.8	22.8	22.8	228.8
Equilibrium	12.5	23	23	23	23	240

Table 3.6: Average capacity, price and profit in the side market over all ten rounds

As can be seen from Figure 3.16 and 3.17 the capacities and prices chosen were fairly constant across all rounds and not significantly different from the equilibrium values. As the data shows no difference between the side markets in the Baseline treatment and the Market Force treatment, we pooled the data of both treatments for the analysis. The realized profits were almost equal to the ex-post predicted profits – the maximum profits given capacity choices – (see Figure 3.18). Hence, once capacities were set, firms on the side market did choose the right profit maximizing prices. Looking into the data only 2.9% of prices chosen in period three were not equal to the predicted prices. Average ex-post predicted profits were only 9 ECUs lower than ex-ante predicted. Hence, firms in the side market, on average, also managed to choose the profit maximizing capacities.

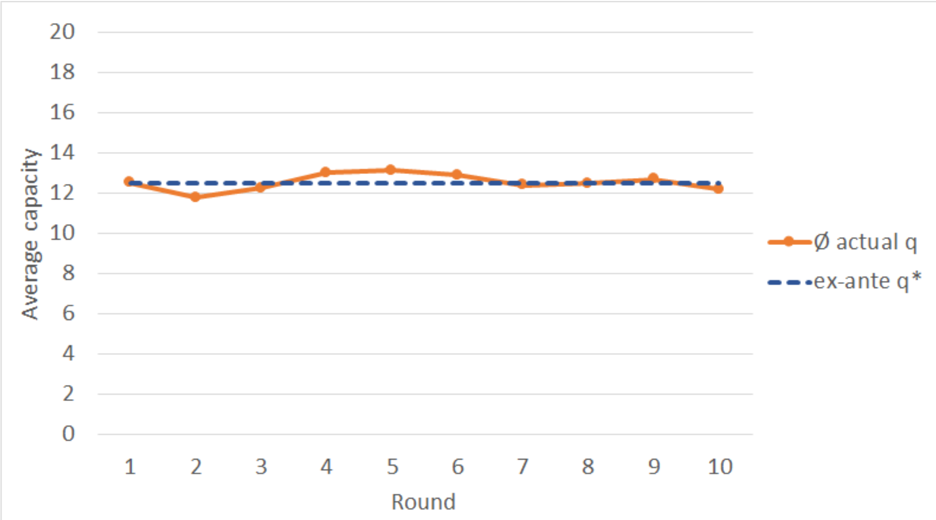


Figure 3.16: Mean capacity vs. ex-ante equilibrium capacity – side market

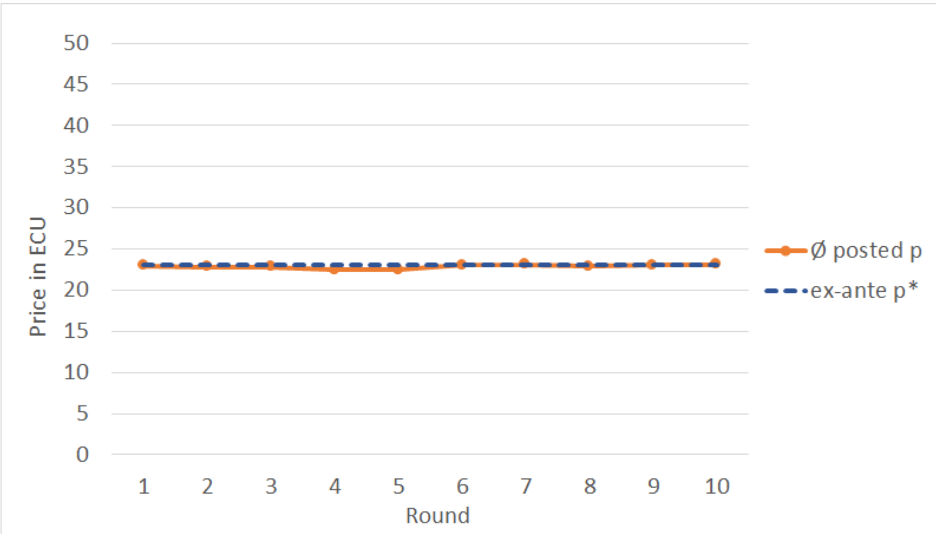


Figure 3.17: Mean posted price vs. ex-ante equilibrium price – side market

In total, we can conclude for the side market that: (i) There was no significant difference between side markets in both treatments. (ii) Firms set capacities which were very close to the equilibrium capacity throughout all rounds. (iii) Almost all firms set the profit maximizing prices and prices did not systematically change between the

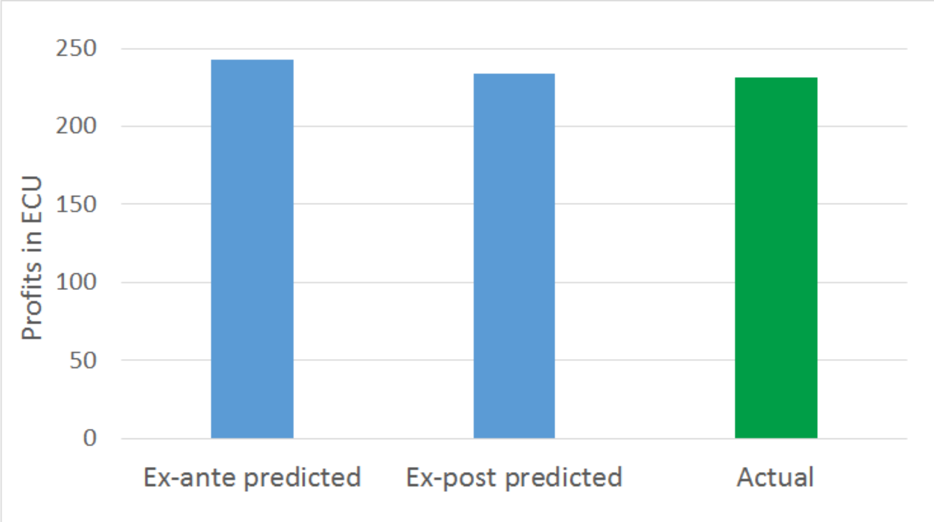


Figure 3.18: Average actual and predicted profits – side market

three periods of a round. And (iv) resulting from (ii) and (iii), average actual profits were very close to the equilibrium profits.

3.4.5 Post-tests

To investigate whether and how certain characteristics such as loss aversion, risk aversion, cognitive skills and social value orientation were related to the outcome variables entry decision, capacity and price choices as well as profits, we ran a battery of post-tests.

Figure 3.19 gives an overview of the results for the four different tasks. As expected, the distribution of subjects' choices for the loss aversion test as well as the risk aversion test has a bell shape. The mode and median is at five for the loss aversion test and at three for the risk aversion test (see Figure 3.19a and 3.19c).[23] Hence, most subjects' loss and risk aversion level were in the middle of our scale. The correlation between these two tests is positive with a correlation coefficient of 0.42.

[23]Lottery five in the loss aversion test: 50:50 chance of either losing 10 points or gaining 20 points. Lottery three in the risk aversion test: 50:50 change of either getting 44 points or 20 points.

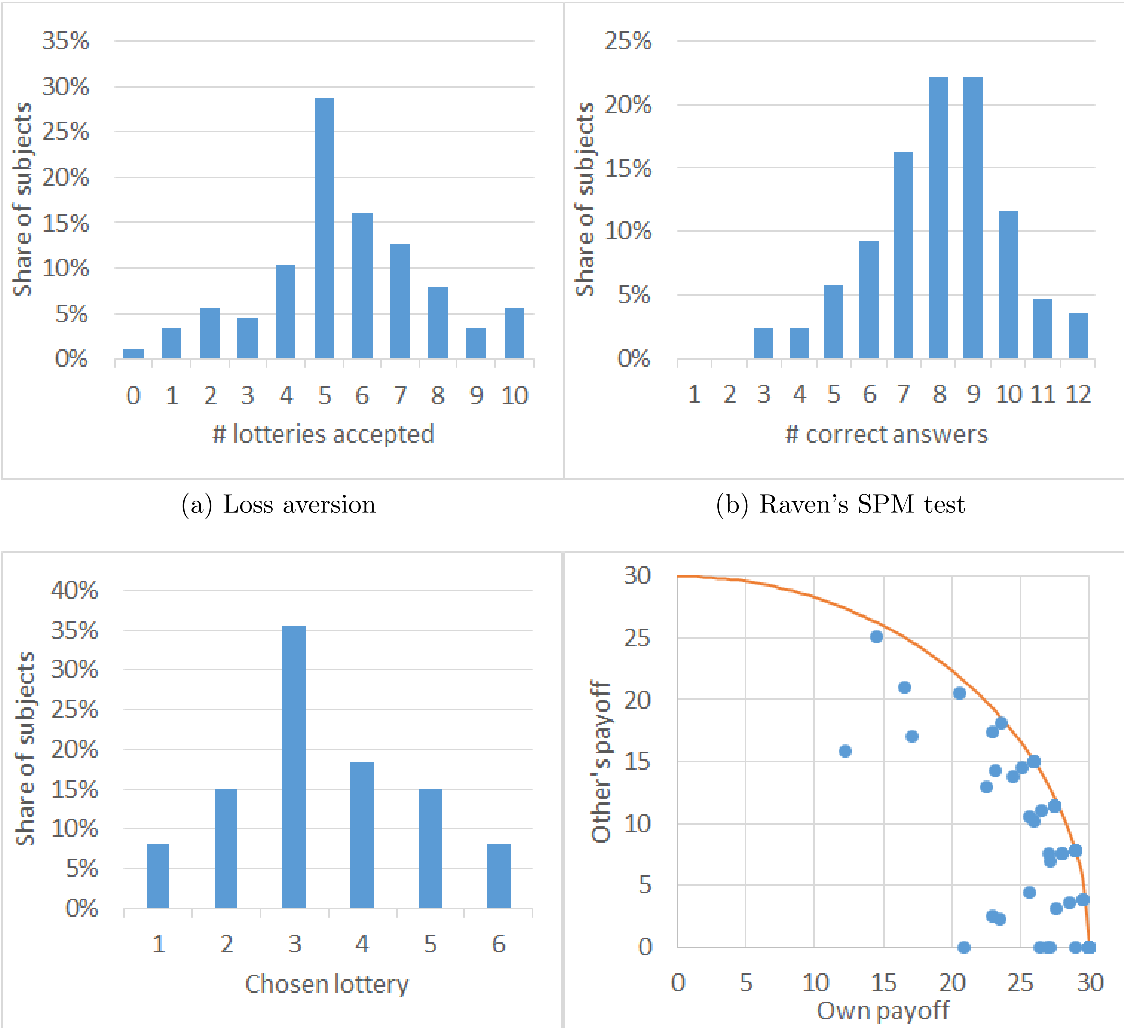
In the Raven's SPM where we tested for cognitive skills almost half of the participants managed to do either eight or nine tasks correctly. 20% of subjects even solved ten or more matrices (see Figure 3.19b).

Figure 3.19d depicts the outcomes of each subject on the the self/other allocation plane of the Liebrand's SVO test. The curve of the SVO ring is included as a reference point. As not all points lie on the boarder of the curve some choices were not pareto-efficient. There were only very few altruistic subjects who attributed more payoff to the other player than to themselves. 51% of subjects were completely "egoistic" as they always chose the option which benefited themselves the most (allocation (30, 0)). The angle of the own-other-payoff allocation vector ("Liebrand angle") which gives the social orientation was negatively correlated to the length of the vector ("Liebrand efficiency") with a coefficient of -0.31 indicating that subjects who were more pro-social tend to be less efficient. Furthermore, risk aversion is positively correlated to the "Liebrand angle" (0.38) and slightly negatively to the "Liebrand efficiency" (-0.23). This implies that subjects who were more risk averse tend to be more pro-social and slightly less efficient.[24]

For ease of interpretation of the regression results we rescaled the post-tests so that a variable value of zero corresponded to the characteristics of (i) loss neutrality, (ii) highest level of cognitive skills, (iii) risk neutrality, (iv) selfishness (focused on maximizing own payoff) and (v) efficiency maximizing.

(i) In the loss-aversion test, the lottery number nine was the loss-neutral option, offering a 50:50 chance of either losing or gaining 20 points. Hence, if a participant accepted lotteries one to nine, the variable "Loss aversion" was set equal to zero in our new scale. We set up the new scale such that the degree of loss aversion increased with increasing numbers. Accepting just lottery number one which was the option with a

[24]The correlation between all other variables were very low with coefficients between -0.04 and 0.13.



(a) Loss aversion

(b) Raven's SPM test

(c) Risk aversion

(d) Liebrand's SVO ring

Figure 3.19: Summary of post-test results

safe outcome of 28 points corresponds to a loss aversion level of eight in our new scale. For lottery number 10 "Loss aversion" was set equal to -1.

(ii) The variable "Raven's test" is equal to zero if all 12 questions were answered correctly in the Raven's SPM test. The modified scale gives minus points for each incorrectly solved puzzle. If one of the 12 tasks was answered incorrectly "Raven's test" takes the value of -1 and so on.

(iii) Lotteries five and six in our risk aversion test were both expected payoff maximizing solutions. A risk neutral individual who only cares about maximizing their expected payoffs should be indifferent between these two lotteries. However, lottery six is a mean-preserving spread of lottery five as the variance is smaller in lottery five. Hence, for lottery five “Risk aversion” = 0 and for lottery six ”Risk aversion” = -1. As with the scale for loss aversion, the higher the value the more risk averse the individual.

(iv) The “Liebrand angle” did not need to be rescaled. An angle of 0 degree corresponds to extreme egotism and an angle of 90 degrees indicates complete altruism.

(v) For the “Liebrand efficiency” metric (vector length), a length of 30 would be the most efficient outcome. Again we set the most efficient outcome to zero. Hence as with the Ravens test, the variable takes on negative values for outcomes that were less efficient.

Entry decision

We used the total number of rounds each participant operated on the main market as the dependent variable for this regression. Although we observed a strong self-selection, none of the measured characteristics had a significant effect on how often a participant chose to operate on the main market as the regression results in Table 3.7 shows.

Capacity choices

To analyze how characteristics elicited in the post tests relate to subjects' capacity choice, we used the difference between their actual capacity and the equilibrium capacity of the corresponding market size as the dependent variable. As these differences could take on positive ($+\Delta q$) as well as negative ($-\Delta q$) values, we looked at both separately.

Beside the measurements from the post-tests, we also included the number of rounds (“Round”) and the aggregated number of rounds a participant operated on the main market (“Aggregated entries”) to account for experience as an explanatory variable. Furthermore, we controlled for the market size (“Market size”) and included the squared market size (“Market size^2”) to account for non-linear effects.

	# Entry
Loss aversion	0.376
	(0.261)
Raven's test	0.248
	(0.289)
Risk aversion	−0.348
	(0.458)
Liebrand angle	−0.019
	(0.034)
Liebrand efficiency	0.057
	(0.275)
Constant	6.221***
	(1.494)
Observations	48
R^2	0.079
Adjusted R^2	-0.031

Notes: Standard deviation are in parentheses.
*p<0.1; **p<0.05; ***p<0.01

Table 3.7: Regression - Market entry decision

The regression results are presented in Table 3.8. We obtain comparable results for both treatments. When capacities were lower than the equilibrium capacity ($-\Delta q$), market size ("Market size") and experience ("Round") had a significantly positive effect on capacity choices in the sense that actual capacities were closer to the equilibrium capacity. In the Baseline treatment, participants with higher cognitive skills measured by the Raven's SPM test chose capacities that were closer to the equilibrium capacity. For negative deviations from the equilibrium this effect is stronger than for positive deviations. With the average deviation being -4.16, a coefficient of 0.7 implies that an increase in correct answers by one decreased the difference to the equilibrium by 17% on average.[25] However, in the Market Force treatment, we found no significant influence of the "Raven's test" on the accuracy of the capacity choices.

Prices choices

For price choices we also differentiated between positive ($+\Delta p$) and negative ($-\Delta p$)

[25] For positive deviations ($+\Delta q$) the mean was 3.7.

	Baseline treatment		Market Force treatment	
	$-\Delta q$	$+\Delta q$	$-\Delta q$	$+\Delta q$
	(1)	(2)	(3)	(4)
Loss aversion	0.443*	−0.032	−0.195	−0.026
	(0.251)	(0.075)	(0.178)	(0.077)
Raven's test	0.695***	−0.167**	0.314	0.051
	(0.138)	(0.075)	(0.212)	(0.082)
Risk aversion	−0.120	−0.067	0.543	0.214
	(0.387)	(0.151)	(0.375)	(0.132)
Liebrand angle	−0.046	0.012	−0.006	−0.015
	(0.043)	(0.013)	(0.026)	(0.010)
Liebrand efficiency	−0.132	0.096	0.036	0.086
	(0.189)	(0.125)	(0.279)	(0.076)
Round	0.471***	0.039	0.590***	0.095*
	(0.156)	(0.058)	(0.168)	(0.057)
Aggregated entries	−0.133	−0.045	0.064	−0.113
	(0.331)	(0.129)	(0.206)	(0.101)
Market size	12.402**	2.766	10.215***	−0.600
	(5.350)	(2.090)	(3.554)	(1.612)
Market size^2	−1.642*	−0.139	−1.067*	0.236
	(0.890)	(0.332)	(0.588)	(0.251)
Constant	−25.849***	−8.707***	−25.634***	−0.717
	(7.540)	(3.241)	(5.239)	(2.567)
Observations	79	151	113	161
R^2	0.559	0.343	0.460	0.124
Adjusted R^2	0.502	0.301	0.413	0.072

Notes: Standard deviation are in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table 3.8: Regression - Deviation of actual capacities from ex-ante equilibrium capacities

deviations from the ex-post equilibrium prices.[26] Column (1) and (3) of Table 3.9 shows that with increasing market size ("Market size") as well as experience ("Round") mean posted prices converged towards the ex-post equilibrium prices.[27]

As for capacity choices, cognitive skill had a significant positive effect on price choices. In the Baseline treatment as well as for negative price deviations in the Market Force treatment, mean prices were closer to the ex-post optimal prices when cognitive skills were higher. The effect was stronger in the Baseline treatment than in the Market Force treatment.[28]

Apart from the cognitive skills, the other characteristics taken from the post-tests had no meaningful impact on pricing. Risk aversion and loss aversion appeared to affect price choices slightly negatively: The difference between mean posted prices and ex-post predicted prices increased with a subject's level of risk or loss aversion. Risk aversion increased the difference between actual and predicted prices significantly for $-\Delta p$ in the Baseline treatment and $+\Delta p$ in the Market Force treatment. Loss aversion had a significant negative effect on the price difference for $+\Delta p$ in the Baseline treatment, in the sense that higher levels of loss aversion increased the price difference. However, it had a marginally positive effect for $-\Delta p$ in the Market Force treatment.[29]

A possible explanation for why loss and risk aversion seemed to affect price choices negatively, at least to some extend, could be that subjects who were very focused on not having any overcapacity tend to set prices that were too low compared to the predictions. On the other hand, subjects that had a more per unit focus may have

[26] We used the average price over all three periods for each participant.

[27] The effect of market size on the price difference was non-linear ("Market size^2" was significant for $-\Delta p$ in column (1) and (3)).

[28] The average deviation for $-\Delta p$ was -2.9 in the Baseline treatment and -2.6 in the Market Force treatment. For $+\Delta p$ it was 1.1 in both treatments.

[29] However, the effects are not very clear as they seemed to counteract each other. While higher levels of risk aversion increased the deviation from the ex-post predictions for $-\Delta p$ in the Baseline treatment significantly, it reduced the price difference for $+\Delta p$, though not significantly. Loss aversion had just the opposite effect in the Baseline treatment. For $+\Delta p$, increasing loss aversion increased the price difference significantly while it had a positive, though insignificant, effect on price accuracy for $-\Delta p$. In the Market Force treatment, the effect of both, loss and risk aversion, was increasing for $+\Delta p$ while for $-\Delta p$ higher loss aversion closed the gap between actual and predicted prices while higher levels of risk aversion increased the difference.

aimed for covering their unit costs with their per unit revenues which lead to higher prices.

	Baseline treatment		*Market Force treatment*	
	$-\Delta p$	$+\Delta p$	$-\Delta p$	$+\Delta p$
	(1)	(2)	(3)	(4)
Loss aversion	0.117	0.186**	−0.215*	0.049
	(0.118)	(0.070)	(0.115)	(0.081)
Raven's test	0.569***	−0.189***	0.294**	0.159*
	(0.084)	(0.060)	(0.133)	(0.086)
Risk aversion	−0.358*	−0.133	0.220	0.272*
	(0.214)	(0.127)	(0.215)	(0.143)
Liebrand angle	−0.021	0.027**	0.007	−0.013
	(0.020)	(0.011)	(0.015)	(0.014)
Liebrand efficiency	−0.173	0.191*	−0.133	−0.114
	(0.126)	(0.106)	(0.119)	(0.119)
Round	0.350***	−0.134**	0.330***	−0.106
	(0.079)	(0.053)	(0.095)	(0.068)
Aggregated entries	−0.278	0.229*	0.129	−0.043
	(0.169)	(0.134)	(0.131)	(0.114)
Market size	9.336***	−2.814	16.532***	0.944
	(2.801)	(2.285)	(2.700)	(1.282)
Market size^2	−1.170**	0.508	−2.172***	−0.117
	(0.464)	(0.350)	(0.435)	(0.206)
Constant	−18.731***	3.579	−32.637***	0.284
	(4.090)	(3.680)	(4.117)	(2.027)
Observations	166	64	193	81
R^2	0.490	0.465	0.466	0.211
Adjusted R^2	0.460	0.376	0.440	0.111

Notes: Standard deviation are in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table 3.9: Regression - Deviation of mean posted prices from ex-post predicted prices

Profits

Regressing profits of each round on the data collected in the post-tests shows that market size has a non-linear effect on profits which is intuitive as profits decrease with market size in our setting. Experience captured by the round variable has a highly significant increasing effect indicating that profits were increasing over time. As before for the capacity and price choices we observe a highly significant positive effect of cognitive skills on profits. According to the regression results higher levels of loss aversion and a more pro-social attitude decreased profits in the Baseline treatment. In the Market Force treatment the effect of these two variables is not significant.

	Profits	
	Baseline	*Market Force*
	(1)	(2)
Loss aversion	-13.064^{***}	-7.430
	(4.820)	(4.643)
Raven's test	19.988^{***}	13.184^{**}
	(3.591)	(5.285)
Risk aversion	-13.336	7.504
	(8.736)	(8.725)
Liebrand angle	-2.121^{***}	0.405
	(0.805)	(0.667)
Liebrand efficiency	-5.963	-6.750
	(5.660)	(5.342)
Round	25.069^{***}	18.258^{***}
	(3.335)	(3.880)
Aggregated entry	-19.122^{**}	5.154
	(7.525)	(5.696)
Market size	$-1,694.355^{***}$	$-1,846.499^{***}$
	(119.823)	(95.537)
Market size^2	212.424^{***}	233.849^{***}
	(19.361)	(15.309)
Constant	$3,461.471^{***}$	$3,607.203^{***}$
	(180.264)	(146.681)
Observations	230	274
R^2	0.819	0.807
Adjusted R^2	0.811	0.800

Notes: Standard deviation are in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table 3.10: Regression - Profits on the main market

3.5 Conclusion

In the economics literature laboratory experiments are often used to examine firms decisions. A major advantage of such experiments is that they allow researchers to analyze specific market settings and behavior in an isolated environment where they can control for and vary the influencing factors. However, as subjects in such experiments are often students, critics question the experiments external validity. One common line of reasoning is that in "real" markets only firms that conform to the model of the sophisticated rational firm succeed since competition and market forces eliminate those that do not. One major concern that has been voiced is that students who are confronted with a new and possibly complex market situation behave rather naively and use simple heuristics when making decisions.

To investigate the relevance of this claim to the validity of laboratory experiments we confronted our subjects with a deliberately complex multi-layered strategic market setting, in which they first chose capacity levels and then competed by setting prices in three consecutive periods in the main market. We set up two different treatments: the Baseline treatment and the Market Force treatment. In the Baseline treatment subjects were allocated at random to markets of two, three or four players. Poor performance in this treatment did not lead to exclusion from the market. For the Market Force treatment, endogenous market entry was introduced. Subjects in this treatment could actively choose to enter the market provided they had sufficient liquidity. If their balance fell below the required minimum they were forced out of the market.

Although we observed strong self-selection into markets in the Market Force treatment, in the sense that participants "specialized in the later rounds and either operated in the main market or stayed in the side market, we observed no substantial difference between the Baseline and Market Force treatments in terms of performance. In both treatments, outcomes were closest to equilibrium predictions in tripolies. In duopolies firms oftentimes did not fully exploit their market power and capacities as well as prices were too low, while in quadropolies capacities and prices tend to be too high affecting profits negatively. Regardless of market size, we observed price adjustments in the three consecutive periods of a round which affected profits positively. Overall, although the setting was deliberately complex, participants decisions converged on average to the equilibrium solution.

In both treatments we find that the market entry decision did not vary significantly with any of the characteristics (loss aversion, risk aversion, cognitive skills and social value and efficiency orientation) from the post-tests. Risk and loss aversion had a slightly negative effect on prices, increasing the gap between posted prices and the optimal prices given total market outputs (ex-post predicted prices). Experience and cognitive skills affected capacity and pricing decisions was well as profits significantly in the sense that with increasing experience and cognitive skills capacity and price choices were closer to the equilibrium predictions and profits increased.

In total, we did not find any substantial differences between the outcomes of the two treatments in our setting. It seems that although particpants in the Baseline treatment did not actively choose to enter the main market, they were able to adjust to the market situation and the decisions they were confronted with.

4 Procrastination and Deadlines in Long-term Team Projects

4.1 Introduction

When the term “procrastination” comes up in conversation, we almost immediately think of our own behavior and the last time we put something off until the last minute. Although we all know that we should *“never put off until tomorrow what we can do today”*, we still struggle to live up to this adage. However procrastination is not only an individual issue but can also be a major roadblock to teamwork.

Even in our everyday lives we encounter situations where we are confronted with inefficient delay in teams. Group projects at school or voluntary contributions to local community projects are examples we are all familiar with. Here a conflict between collective and individual interests can arise since costs are private while benefits are public. The collective interest is in achieving a common goal that would be beneficial for everyone involved in the sense that it would maximize joint payoffs. In such situations selfish team members who pursue their own interests will minimize their own investment while relying on the efforts and investments of the other team members. Such detrimental holding back on investments might have fatal consequences, since it can lead to increased delay costs or even the failure of a project. The situation is worsened when contribution decisions are made independently and there is uncertainty about the actions of the other people involved.

In the economic literature, the conflict of collective and individual interests is modeled predominantly by public good games. One special case of public good games includes a threshold that has to be met to guarantee the provision of these goods. In general, the underlying games for most common public good settings are one-shot decision games. However, there are many situations where a one-shot game even if it is repeatedly played is not realistic: Many group projects as well as cooperation among firms normally last for more than one period. Investments have to be provided continuously over a certain time period in order to complete a project or achieve a goal. So what happens if contributions have to be accumulated over time? Will teams be able to coordinate on the most efficient outcome when there is no deadline, or will they delay their work or even fail to complete a project at all? How can the efficiency of teamwork be improved?

Deadlines are a common way of counteracting procrastination. But how effective are deadlines in a team setting? The effect of a deadline will depend on different factors such as how binding it is and how it was set. For example, a deadline can be externally imposed (exogenous) or set by the individuals or firms themselves. In the case of a school project an exogenous deadline might be one set by the teacher. On the other hand, sometimes students are allowed to set their own deadlines. Some self-imposed deadlines are binding, e.g. when a group has to commit to the hand-in day for a project. Other self-imposed deadlines, such as intermediate deadlines, are non-binding and are often used to prevent everything from being postponed until the last minute. But do deadlines increase efficiency? And what exactly is the difference between exogenous deadlines and self-imposed ones?

The issue of procrastination in a team setting has been analyzed theoretically by Bonatti and Hörner (2011). They investigate collaboration over an infinite time horizon where efforts are private information and only the outcome can be observed by the agents. If agents wanted to maximize joint profits, it was best for them to exert maximal effort. However, the authors found that when agents could do things later, they did. Agents exerted too little effort too late, relying on the efforts of others. Bonatti and Hörner refer to this specific type of inefficient delay as procrastination.

Building on the work of Bonatti and Hörner, Külpmann (2015) defines an important effect seen after contributions have been made over an extended period as the encouragement effect: An individual's actions today will affect the efficiency of every players future efforts and therefore their choice of effort. In continuous contribution models such as Bonatti and Hörner's, the encouragement effect is usually negative: Putting in more effort today will "encourage" the other team members to exert less effort tomorrow, whereas exerting less effort today will force the others to increase their efforts tomorrow in order to complete the project.

In the experimental literature some well-known work has been done on the issue of coordination failure by Van Huyck et al. (1990, 1991). Following Van Huyck et al., other authors have also studied the prevalence of coordination failures and tested the robustness of the findings.[1]

In addition, voluntary contribution to threshold public goods has been analyzed experimentally in various settings: While continuous contribution (Cadsby and Maynes, 1999), a money-back guarantee (Cadsby and Maynes, 1999) and sequential move (Rau and Normann, 2011) increase total contribution, completion rate and payoffs, they do not stop attempts to free-ride. In a continuous contribution setting, the average first-round investment also has an effect on total contributions: The lower the average first-round investment, the lower the total contribution (Dannenberg et al., 2011).

So far experimental research on procrastination has mainly focused on the aspects of time-inconsistency and present bias (Bisin and Hyndman, 2014) as well as self-control problems in an individual setting. Some studies have analyzed the effect of certain methods such as costly self-imposed deadlines and intermediate deadlines to enforce self-discipline. According to Ariely and Wertenbroch (2002) subjects in the exogenous

[1] Anderson (2001) and Goeree and Holt (2005) show that lower effort costs lead to higher frequency of coordination on the payoff-dominant equilibrium even with random matching. Blume (2005) introduces communication and finds that cheap talk facilitates quick convergence towards coordination.

deadlines treatment performed best, while subjects in the no-deadline treatment performed the worst. In addition, self-imposed deadlines were often not set optimally and hence did not improve performance as much as expected. Possible reasons for what Ariely and Wertenbroch call imperfect sophistication (or "partial naiveté") might be biased self-perception and cognitive limitations in calibrating deadlines or even a deliberate mixed strategy of balancing flexibility and self-control. The results of Bisin and Hyndman (2014) are in line with that: There setting a self-imposed binding deadline did not increase completion rates either, even though sophisticated people show a high demand for commitment in the form of self-imposed deadlines. Bisin and Hyndman argue that there are various factors besides present-bias causing procrastination such as over-confidence and lack of preference. Ariely and Wertenbroch (2002) also find that under intermediate deadlines performance is higher and there is less delay. However, Burger et al. (2011) observed the exact opposite. In their experiments, exogenous intermediate deadlines did not improve the completion rate. On the contrary, they lowered it. The major difference between their experimental design and Ariely and Wertenbroch's was that not meeting intermediate deadlines resulted in failure of completion while Ariely and Wertenbroch's subjects were penalized for missing a deadline but could still participate in the experiment. Burger et al. conclude that some deadlines may in fact reduce flexibility enough to overcome any benefits they provide in terms of improving performance.

To the best of my knowledge, procrastination and the effect of deadlines have not been analyzed experimentally in a teamwork setting where contributions accumulate over time. The lack of research in this area leaves interesting possibilities that need to be investigated. My study addresses these issues by empirically investigating individual investment behavior in long-term team projects. I use a threshold public good setting with a continuous contribution that requires multilateral contributions over an extended period for completion. This setting is then varied in two dimensions: type of deadline and team size. I seek to analyze whether deadlines improve efficiency and whether there is a difference between self-imposed deadlines and exogenous ones.

According to the literature, deadlines might help reduce procrastination or at least encourage teams to make a final sprint. However, they might also have a negative effect on subsequent levels of effort if the first-period investment is very low. I also want to see how team size affects the results. Conventional research indicates that the larger the team, the easier it is for an individual to free-ride. Therefore increasing team size should lower individual effort provision. Hence, the larger the team, the longer it should take to finish a project.

I find clear evidence of inefficient procrastination by teams in the absence of deadlines. The tendency to free-ride and delay the project increases with team size. In fact, large teams in this study needed almost twice as long as small ones to complete projects. Project completion was also negatively affected by team size. While introducing a deadline helped to mitigate the problem of procrastination, it had a negative effect on project completion, particularly when the deadline was exogenous rather than self-imposed. The strong negative effect of team size remained even when deadlines were set.

The structure of this chapter is as follows. Section 4.2 explains the study's general design. Section 4.3 provides a short theoretical prediction, and Section 4.4 gives a detailed description of the experiments that were conducted. I present the results of the experiments and my conclusions in Sections 4.5 and 4.6 respectively.

4.2 Experimental Design

4.2.1 General Setting

Consider a team of n players working together on a project. To complete the project, the aggregated investment must be at least $H > 0$ with $H \in \mathbb{R}_+$ ("effort threshold"). It takes at least $k > 1$ periods to finish the project. Each player has an initial wealth of $W_i \in \mathbb{R}_+$. In each period, every player decides how much to invest in the project.

Individual investments in period t are denoted by $x_{i,t} \in [0, b]$, where $b \in R_+$.[2] Whatever players decide to invest in the project will be subtracted from their current wealth immediately and is non-refundable. Each player can invest at most W_i in total. Once the team's total investment is greater than or equal to the threshold, the project is completed. As a reward for completing the project, each member receives $R \in \mathbb{R}_+$. However, the net reward decreases over time. When the team needs more than k periods (the minimum number of periods) to complete the project, an amount of $d > 0$ is subtracted for each additional period ("delay cost"). Define the aggregated total team investment until period t as $Y^t = \sum_{i=1}^{n} \sum_{k=1}^{t} x_{i,k}$. Then, payoffs are calculated as

$$EU(x_{i,t}, x_{-i,t}) = \begin{cases} W - \sum_{t=1}^{T} x_{i,t} + R - d(T-k) & \text{for } Y^T \geq H \\ W - \sum_{t=1}^{\infty} x_{i,t} & \text{else} \end{cases}$$

where time is denoted by t and T is the period in which the project is completed.

Now consider a setting where a deadline $\hat{T}$ is introduced. If the project is not completed before the deadline the team will not receive the reward R. All investments are non refundable. Then payoffs are

$$EU(x_{i,t}, x_{-i,t}) = \begin{cases} W - \sum_{t=1}^{T} x_{i,t} + R - d(T-k) & \text{for } Y^T \geq H \text{ and } T \leq \hat{T} \\ W - \sum_{t=1}^{\hat{T}} x_{i,t} & \text{else} \end{cases}$$

where time is denoted by t and $T \leq \hat{T}$ is the period in which the project is completed which is either $\hat{T}$ or less then $\hat{T}$ if the project is completed earlier.

4.2.2 Experimental Setting

The setting in this study is varied in two different dimension: team size and deadlines. To analyze the effect of team size, I distinguished between $n = 1$, $n = 3$ and $n = 7$.

[2] The introduction of an upper limit on investment b can be justified by the fact that efforts are naturally limited in real life. For example, a day only has 24 hours, and when it comes to monetary contributions, people are usually budget constrained.

The behavior in an individual setting ($n = 1$) is used as a benchmark. I chose $n = 3$ as a team size for small teams, firstly because it is frequently chosen in the literature as a good proxy for a small group, and secondly, because $n = 2$ is a special case in which a player can always directly infer his or her teammate's individual investment. To create a noticeable difference between small and large teams, I wanted to at least double the team size for large teams. As an odd number is required for determining the maximum number of periods in one of the treatments with a deadline as will be explained in detail below, I chose $n = 7$ for large teams.

To analyze the effect of deadlines, three different treatments are introduced.
(i) No deadline (ND): Since this is a lab experiment it was not possible to implement a treatment with an infinite deadline. But in order to simulate a situation in which subjects felt like they were working on a long-term project without the pressure of a deadline the maximum number of periods had to be sufficiently long. Therefore, the maximum number of periods was set at 30 periods in this treatment.
(ii) Self-imposed deadline (SID): Team members made suggestions for the maximum number of periods and a deadline was defined through a mechanism based on the suggestions made.
(iii) Exogenous deadline (ED): Relatively short deadlines of between five and ten periods were set exogenously and communicated to the teams . The way decisions were made did not change between treatments. The only thing that changed was that teams had to finish the project before the given deadline $5 \leq \hat{T} \leq 30$. Otherwise they do not receive any reward.

The individual investment per period was $x_{i,t} \in [0, 10]$. Every subject was given the same initial endowment of 75 ECUs (Experimental Currency Units). Each team member was allowed to invest at most 75 ECUs over the course of a single project. The completion threshold for small teams was defined as $H = 150$ and $H = 350$ for large teams. Given these thresholds, teams of both sizes needed at least five periods to complete a project. Once a project was completed, each team member received a reward of 90 ECUs. However, a delay cost of 2 ECUs was subtracted from the reward for every period over five.[3] If subjects maximized their joint profits as a team, investing 10 ECUs in each period and finishing the project in five periods, they achieved the optimal and most efficient outcome.

The payoff function for each individual is then given by

$$EU(x_{i,t}, x_{-i,t}) = \begin{cases} 175 - \sum_{t=1}^{T} x_{i,t} - 2T & \text{for } Y^T \geq H \text{ and } T \leq 30 \\ 75 - \sum_{t=1}^{30} x_{i,t} & \text{else} \end{cases}$$

Periods are denoted by t, and T is the last period of a round which is either 30 or less then 30 if the project was completed earlier.

	No Deadline	**Deadline**	
		Self-imposed	**Exogenous**
n = 3	30	5 - 30	5 - 10
n = 7	30	5 - 30	5 - 10

Table 4.1: Maximum number of periods

4.3 Theoretical Prediction

The joint payoff-maximizing outcome, with everyone investing 10 ECUs in each period and completing the project within five periods is not sustainable as an equilibrium, as each team member has an incentive to stop their investment if the other team members continue to contribute.

Focusing only on symmetric equilibria, $x_{i,t} = 0 \; \forall \; i$ and t is one possible equilibrium, as the best response to everyone else not investing is also not to invest. The project will not be completed and everyone will have a payoff of 75 ECUs.

[3]The delay costs were set such that teams would not give up on completing the project too early.

One other symmetric equilibrium involves all participants playing a trigger strategy in which the total investment in period t is equal to $\hat{X}^t$. If someone deviates and the total investment in a periods is $X^t < \hat{X}^t$ everyone will stop their investment and the project is not completed. In order for the trigger strategy to be credible the payoff from continuing the investment and completing the project by oneself has to be smaller than by sticking to the trigger strategy while taking into account the time constraint ($T \leq 30$) and the budget restriction ($W_i = 75$).

As before $Y^t = \sum_{i=1}^{n} \sum_{k=1}^{t} x_{i,k}$ is the aggregated total team investment until period t. For $n = 3$ and $n = 7$ every team member invests $x_{i,t} > 0$ in each period such that

$$\hat{Y}^t \leq \begin{cases} 25(2n-3)\frac{n}{(n-1)} & \text{for } t \leq \frac{27.5n-30}{(n-1)} \text{ and } T \leq 30 \\ 50n + 10t - 300 & \text{for } t \geq \frac{27.5n-30}{(n-1)} \text{ and } T \leq 30 \end{cases} \tag{4.1}$$

holds.

Assuming that individuals will chose to invest whenever they are indifferent between investing and not investing, the individual per-period investment in this symmetric equilibrium can be defined as

$$x_{i,t} = \begin{cases} 25(2n-3)\frac{1}{(n-1)} & \text{for } Y^{t-1} \geq \hat{Y}^{t-1} \text{ and } t \leq \frac{27.5n-30}{(n-1)} \\ \frac{1}{n}[50n + 10t - 300] & \text{for } Y^{t-1} \geq \hat{Y}^{t-1} \text{ and } \frac{27.5n-30}{(n-1)} < t \leq 30 \\ 0 & \text{else} \end{cases} \tag{4.2}$$

In this equilibrium the project is completed in period 30, which is the last period in our treatments without a stricter deadline. It is not possible to complete the project earlier as once the aggregated total team investment is too high, players have an incentive to deviate from the equilibrium strategy and start to free-ride since the threat to immediately stop investing once someone defects is not credible anymore.

After completing the project in 30 periods, every team member receives a payoff of $EU(x_{i,t}, x_{-i,t}) = 125 - \sum_{t=1}^{30} x_{i,t} = 125 - 50 = 75$. The payoff is equal to the initial endowment and hence also equal to the payoff when $x_{i,t} = 0 \; \forall \; i$ and t. This is because the variable values in this setting were constructed such that the net reward is equal to the delay costs whenever the maximum of 30 periods is needed to complete the project.

As condition 4.1 has to be fulfilled, there is no other symmetric equilibrium in which the project is completed before the very last period of a round.[4]

4.4 Experimental Procedures

4.4.1 General procedures

The experiment consists of two parts. In part one each participant works alone for one round to complete one project. The maximum number of periods for this round is 30. In part two participants are assigned to teams. This part of the experiment consists of either twelve rounds (self-imposed deadline) or nine rounds (exogenous deadline), with each round constituting an entire game either with the maximum number of periods being 30 or the type of deadline specified in brackets. Table 4.2 shows, each subject took part in three different treatments during a session: single-player, teams without deadline and teams with either self-imposed deadline or an exogenous deadline.

Endogenous deadline treatment		**Exogenous deadline treatment**	
Part 1	**Single-player**	**Part 1**	**Single-player**
1 round	without a deadline	1 round	without a deadline
Part 2	**Teams**	**Part 2**	**Teams**
3 rounds	wihout a deadline	3 rounds	wihout a deadline
9 rounds	with self-imposed deadline	6 rounds	with exogenous deadline

Table 4.2: Session sequence depending on deadline variation

I conducted the experimental sessions in the Experimental Economics Laboratory at the University of Mannheim (mLab).[5] Across fourteen sessions, a total of 276

[4]For details see Appendix C.1.

[5]The experiment was programmed and conducted using the software z-Tree (Fischbacher, 2007) and subjects were recruited through ORSEE (Greiner, 2004).

undergraduate and postgraduate students participated. All students were recruited from the general student population of the University of Mannheim.

Participants were compensated based on their ECU earnings in one randomly selected round.[6] The exchange rate was 6 ECUs = 1 Euro. The average payoff per subject was around 18 Euros.

4.4.2 Treatments

Single-player (without deadline)

At the beginning of each session, the instructions were read aloud and explained to all participants. Participants were not allowed to talk or communicate during the sessions and were unable to influence each other than through their actions. Then, all participants played one round of the game on their own with a threshold of 50 ECUs. This not only enabled participants to familiarize themselves with the game and the interface but was also a chance to analyze participants behavior when they are working on their own and to observe whether inefficient procrastination occurred even when there is no scope for free-riding.

After the single-player round, participants were randomly placed into groups of three or seven. Random matching was employed at the beginning of each new round to allocate participants to groups. Participants took part in either a small team ($n = 3$) or a large team ($n = 7$) treatment, not both. The small-team sessions involved 18 participants who were separated into two matching groups to guarantee the existence of independent observations from two completely unrelated samples. For the experiments with large teams, each session consisted of one matching group with 14 participants.

[6] At the end of the session, one of the participants was asked to draw a ball with a number on it. The number picked represented the round that would be payoff relevant.

Teams without deadline

For rounds one to three of part two, the maximum number of periods was 30 per round. In each period, all team members made their investment decisions independently. At the end of each period, participants were informed of their teams total investment, the total amount that they had invested thus far and how much their team still needed to invest to reach the threshold. Once a teams total investment (summed up over all previous periods and the current period) reached or exceeded the threshold, the current round ended as soon as the team completed its project. If the project was not completed within 30 periods, the round ended and no rewards were given.

Teams with self-imposed deadline

In rounds four to 12 of the self-imposed deadline treatment teams could decide on the maximum number of periods. At the beginning of each round, each team member was asked to suggest a maximum number of periods (at least five and up to 30). As soon as all team members had submitted a suggestion, the suggestions were ranked from lowest to highest. The median value was then imposed as the relevant maximum number of periods for that round. For transparency, all participants received feedback on the suggestions of their fellow team members and the resulting deadline for that round. The deadlines in this treatment can therefore be seen as the self-imposed result of a voting process that all team members were able to influence to a certain degree.

Teams with exogenous deadline

In rounds four to nine of the exogenous deadline treatment, teams were informed of the maximum number of periods at the beginning of each round. The deadlines were chosen based on the previously obtained results of the self-imposed deadline treatment. In the self-imposed deadline treatment, a very clear and fast convergence towards the jointly efficient deadline of five was observed. With much lower frequency and mostly in earlier rounds, deadlines of six, eight and ten periods were also played. The sequence of exogenous deadlines chosen for this teamwork section was thus $5, 6, 10, 5, 8, 5$.

Throughout the experiment, participants were provided with a calculator on their screen. There they could enter their average investment per period starting from the period they were in and their belief about all others average investment per person per period. Then they were told what the last period and their expected payoff would be. They were allowed to use the calculator as often as they wanted. At the beginning of each period participants received feedback on the total team investment of the last period, the total team investment up to the current period and how much was still missing to reach the threshold. They were also informed about how much they invested in the last period and how much their aggregated investment was until the current period. Participants were able to see their current account balance throughout the entire round. At the end of each round, they were informed whether the project had been completed and about their payoffs for the round.[7]

4.5 Results

4.5.1 One-shot single-player

Result 1:

The completion rate in the single-player treatment was 100%, and very little procrastination was observed.

Figure 4.1 shows the cumulative share of individuals ($n = 1$) and teams ($n = 3$ and 7) who completed their projects within the deadline. As one can see, all subjects completed their projects when working alone. On average, they needed six periods to do this. 70% of participants finished their projects in five periods, which is

[7] Appendix C.2 provides screenshots of the games.

payoff-maximizing regardless of risk preferences. 30% of subjects needed more than five periods. One possible explanation could be confusion. Some subjects might not have realized in this very first round that finishing the project as fast as possible was payoff-maximizing. Another reason may have been a preference to experiment. Some participants may have used the first round as an opportunity to get acquainted with the game, accepting a small decrease in payoff in the bargain. Although some delay was observed, it was not too severe. 97% of participants finished within the first ten periods.

The average payoff in the single-player part was 113 ECUs, which is very close to the maximum possible payoff of 115 ECUs.[8] Using the ratio between an individual's actual profit and the maximum possible payoff, we can see how efficient individuals were on average. I define this ratio as the degree of efficiency $DoE = \frac{\text{actual payoffs}}{max \text{ payoffs}}$. In case of teams the DoE was defined as the ratio between a team's actual joint profit and the maximum possible joint profit, $DoE = \frac{\sum_{i=1}^{n} \text{actual payoffs}}{max \sum_{i=1}^{n} \text{payoffs}}$. Achieving the maximum possible profit leads to a DoE of 1. For the single-player treatment the DoE was 0.98, which is extremely high, indicating that subjects were very efficient in the single-player round.

4.5.2 Teams without a deadline

Result 2:

(i) There is clear evidence for procrastination by teams when there is no deadline. Completion rates were affected negatively by working in teams compared with working alone.

(ii) Team size had a strong negative effect on completion rates, completion time and efficiency.

[8] The maximum possible payoff can be achieved if every team member invests 10 ECUs in each period and the project is then completed in period five.

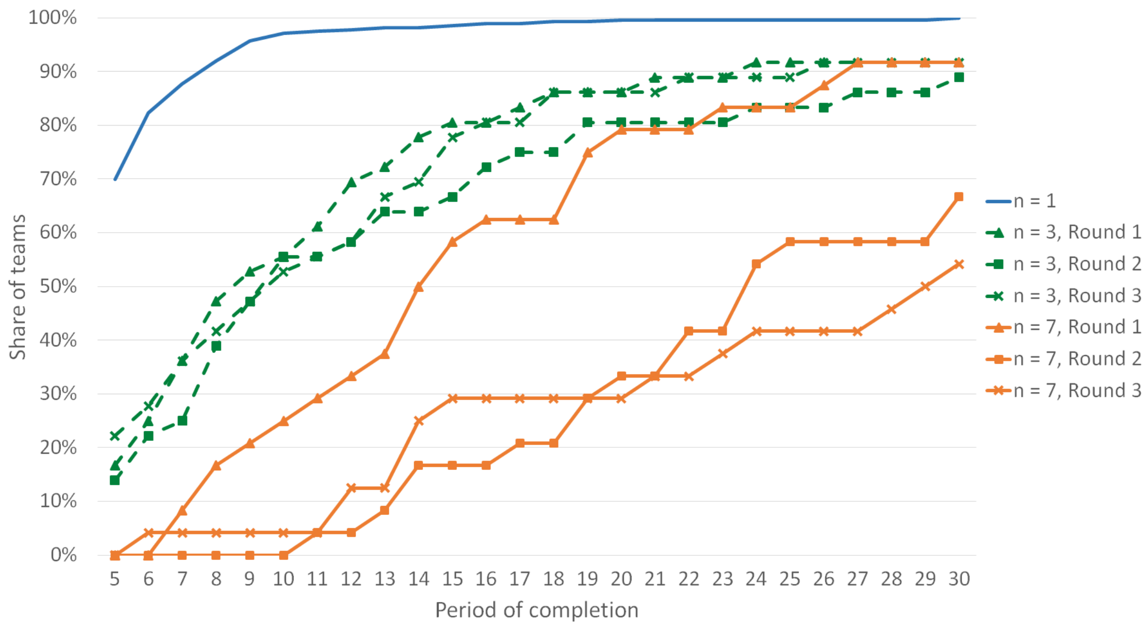


Figure 4.1: Cumulative distribution of completion time for the "single-player" and the "teams without deadline" treatments

Let us consider completion rates first. In contrast to the single-player treatment, not every team managed to complete projects within 30 periods. Aggregated over all three rounds 98 out of 108 three-person teams completed their projects, for a completion rate of approximately 92%. For seven-person teams the completion rate was only 71% (51 out of 72 teams). Disaggregating the data and comparing the completion rates for each round separately, Figure 4.1 and Table 4.3 show that for small teams the completion rate was constant over all three rounds. By contrast, there was a big difference between the data in the first round and the latter two rounds for seven-person teams (see Figure 4.1 and Table 4.3). Comparing the completion rate in round one to the average combined completion rate for rounds two and three, the difference was statistically very significant ($p = 0.004$, Wilcoxon signed-rank test). The difference between round two and round three was not statistically significant ($p = 0.299$).

Pooling the data for rounds two and three, we can see that the mean completion rates for small teams was significantly higher than for large teams as the p-value of a Mann- Whitney U test was 0.008. Hence, team size had a strongly negative effect on completion rates.

Given the substantial difference in the behavior of seven-person teams between round one and the latter two rounds, I will pool the data for rounds two and three and consider the data from round one separately in my analysis from now on. Further, between-subject comparisons were tested using the Mann-Whitney U test and within-subject comparisons with the Wilcoxon signed-rank test if not stated otherwise.

	Single-player ($N = 276$)	3-person teams ($N = 12$)			7-person teams ($N = 12$)		
Round	1	1	2	3	1	2	3
Completion rate in %	100	91.7	88.9	91.7	91.7	66.7	54.2
∅ Completion time	5.9	11.7	13.6	12.2	16.1	23.7	23.8
Degree of efficiency (DoE)	0.98	0.87	0.84	0.87	0.78	0.59	0.59
DoE (only completed)	0.98	0.91	0.88	0.90	0.83	0.73	0.76

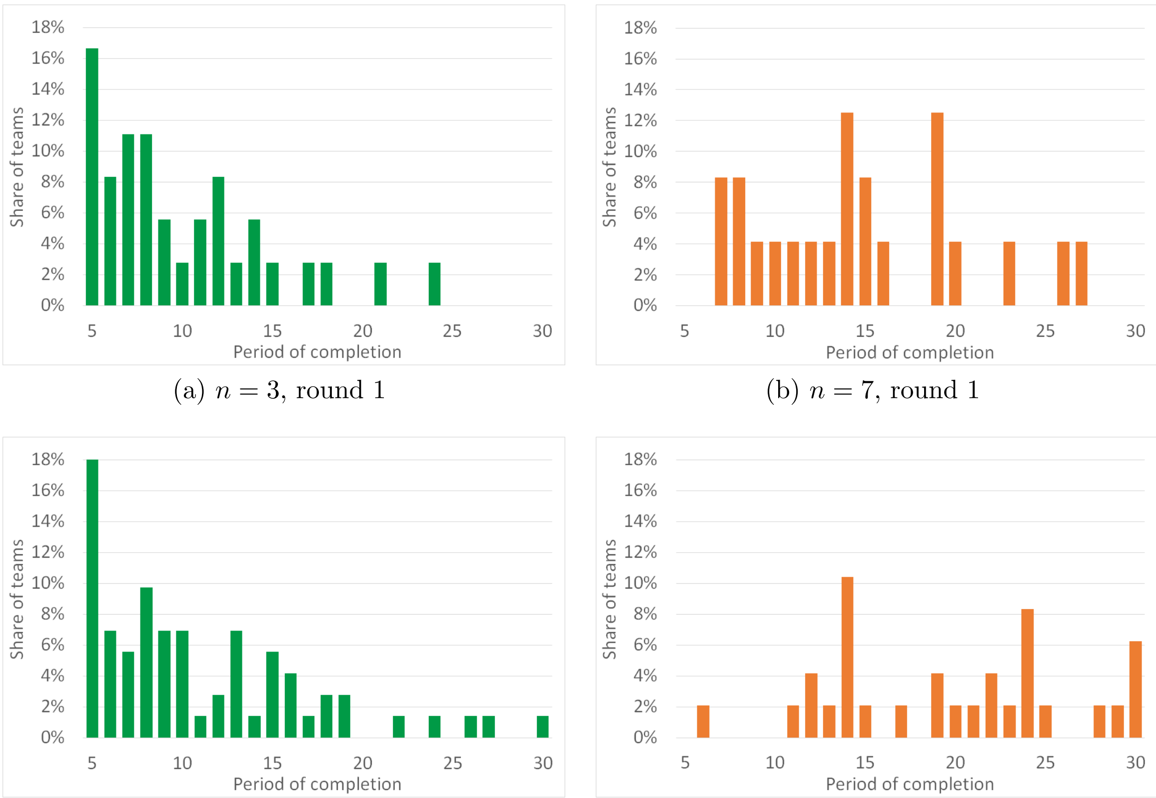
Notes: (i) N = number of independent observations. (ii) As the DoE includes teams that did not complete their projects, I calculated a DoE (only completed) for teams that did.

Table 4.3: Summary of completion rate, completion time and efficiency

Now to examine completion time. The average number of periods small teams needed for project completion over three rounds was 10.7, or more than twice as many as the five-period minimum. This shows that even in small teams there were signs of coordination failure and procrastination. For large teams the average was 17.6 periods. For rounds two and three the average completion time for seven-person teams was 19.7 periods. Comparing this to the average second- and third-round completion time of the three-person teams, the difference is highly significant ($p < 0.001$).

The distribution of completion times by team size is depicted in more detail in figure 4.2. The distributions for three-person teams peak at five periods, indicating that most teams finished their projects in five periods. The median is nine periods in the first round and ten periods in the second and third rounds. Figure 4.2a and 4.2c show that the distribution remained almost the same across all three rounds. The share of three-person teams who completed projects in five periods increased slightly, and some teams needed more than 25 periods, which did not occur in round one. But these shifts in the distribution are not significant ($p = 0.248$). For large teams ($n = 7$), completion time was more evenly distributed over the 30 periods, with small peaks at periods

14 and 19 in the first round and periods 14 and 24 in rounds two and three. Not a single large team managed to finish a project within the first five periods. Looking at Figure 4.2b and 4.2d, we see that delays were much more pronounced in rounds two and three. In the first round 25% of all teams completed their projects within the first ten periods while only one team managed to do so in rounds two and three. In the later two rounds 31% of teams needed more than 25 periods, which is about twice as many as in the first round (16%). This shift towards longer completion times is marginally significant, with a p-value of 0.056. Hence, we can conclude that a larger team size fostered procrastination here, with delays increasing over time.



(a) $n = 3$, round 1

(b) $n = 7$, round 1

(c) $n = 3$, rounds 2 and 3

(d) $n = 7$, rounds 2 and 3

Figure 4.2: Distribution of completion time depending on team size

That small teams were more efficient overall than large ones is also indicated by the degree of efficiency (DoE) in Table 4.3. Although the three-person groups were not as efficient as individuals working on their own, the average DoE was still very high. On average, teams achieved 86% of the maximum joint payoff. The first round DoE for the seven-person teams was also rather high. In rounds two and three the

DoE dropped by about 20 percentage points and was significantly lower than for small teams ($p = 0.005$), which corresponds to the significantly lower completion rate and the longer completion time. As the *DoE* reports the average efficiency of all teams, it also includes those that failed to complete projects. If we take only those teams into account that completed their projects, the difference in efficiency between small teams and large teams is not as large but still significant ($p = 0.018$). Similarly, the difference in efficiency between the first round and the second and third rounds also decreases for large teams if one focuses solely on cases of project completion. This indicates that project non-completion was a major driver of the observed loss in efficiency.

Result 3:
Teamwork in large teams is characterized by strong free-riding behavior, which increases over time leading to more and longer delays and lower completion rates.

Marked free-riding behavior was a major reason for the low completion rates for large teams in this study, especially in the latter two rounds. Table 4.4 shows that the average investment level per team member was much lower in the seven-person teams than in the teams with three members. Looking just at teams that completed their projects, a member of a seven-person team invested only about half as much as someone in a three-person team in rounds two and three.

	n=3		n=7	
Round	1	2, 3	1	2, 3
completed	6.1	5.9	4.0	2.9
uncompleted	0.8	0.7	1.3	1.1
∅ in total	5.7	5.4	3.7	2.2

Table 4.4: Average per period investment per team member

Looking at individual investment in seven-person teams, a large portion of participants chose to free-ride, investing only 0 or 1 ECU per period.[9] Figure 4.3 shows the share of participants who chose to invest 1 ECU or less in each period of each round. In the first period of the first round 15% of subjects chose to free-ride. This share increased dramatically during the round to 59% in period ten. At the start of rounds two and three around 35% of subjects are already free-riders, a significantly higher share than in the first round ($p = 0.002$).

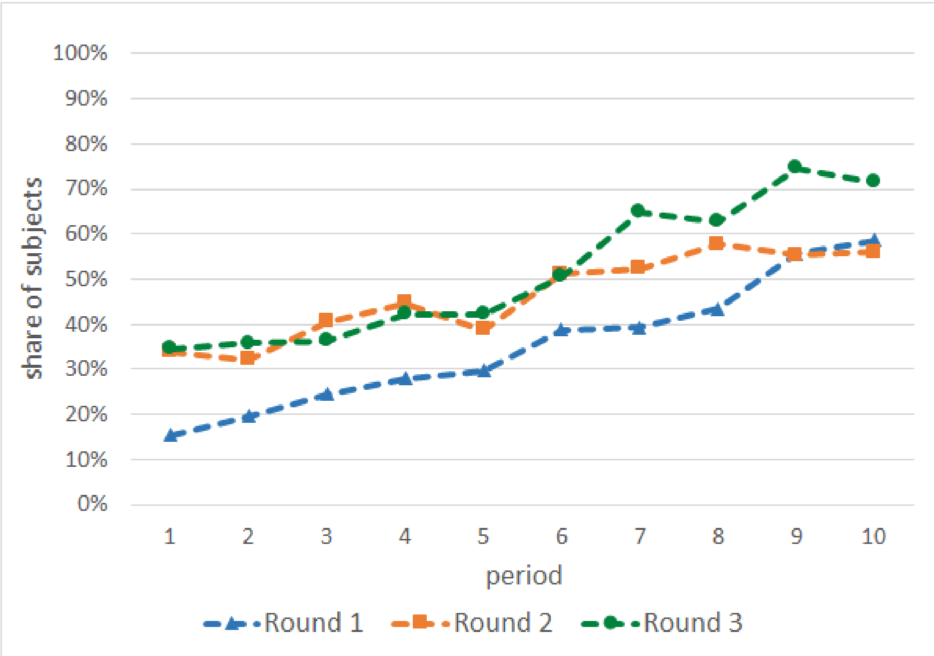


Figure 4.3: n = 7, per period share of subjects who free-ride

As can be seen from Figure 4.3 the share of free-riders increased considerably not only from round to round but also from period to period within the same round. To test whether the observed increase in the share of free-riders over the first ten periods of each round was statistically significant, I first calculated the regression coefficient of a simple linear regression with the share of free riders as the dependent variable and periods as the independent variable for each of the 12 independent observations. This regression coefficient tells us how much the share of free-riders increased or decreased on average

[9]Investing 1 ECU or less per period is considered severe free-riding here for reasons: Firstly, if every team member invested only 1 ECU per period, the team would be unable to complete its project. Secondly, investing only 1 ECU is similar to giving to charity to ease ones mind. Only a very small fraction of participants chose to invest just 1 ECU. Therefore, the summary statistics are not much different than if only per-period investments of 0 ECU counted as severe free-riding.

from period to period. If there is no systematical difference in the number of free-riders, the regression coefficients should not be significantly different from zero. Hence, the null hypothesis is that changes are equal to zero. Then I tested the coefficients against the null hypothesis using a Wilcoxon signed-rank test. Applying this method to round one and the pooled data for rounds two and three shows that the increase in the share of free-riding subjects is highly significant ($p = 0.003$ for round one and $p = 0.005$ for rounds two and three).

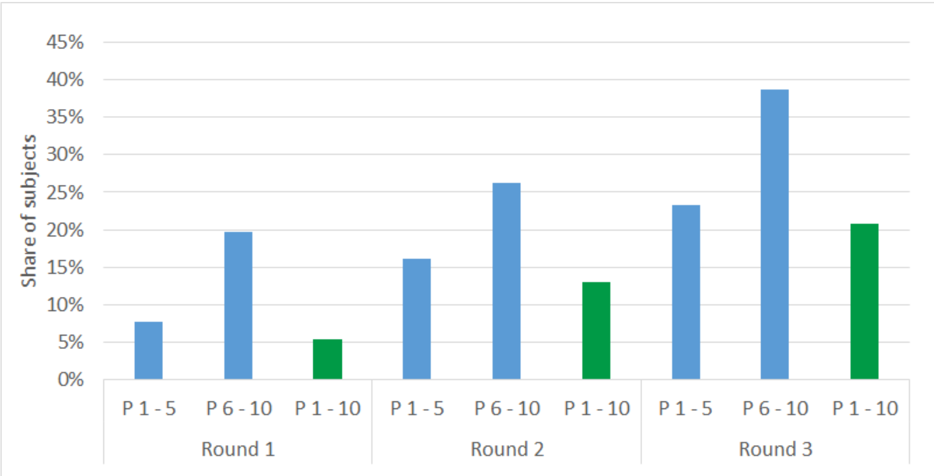


Figure 4.4: n = 7, share of free-riders throughout periods 1−5, 6−10 and 1−10

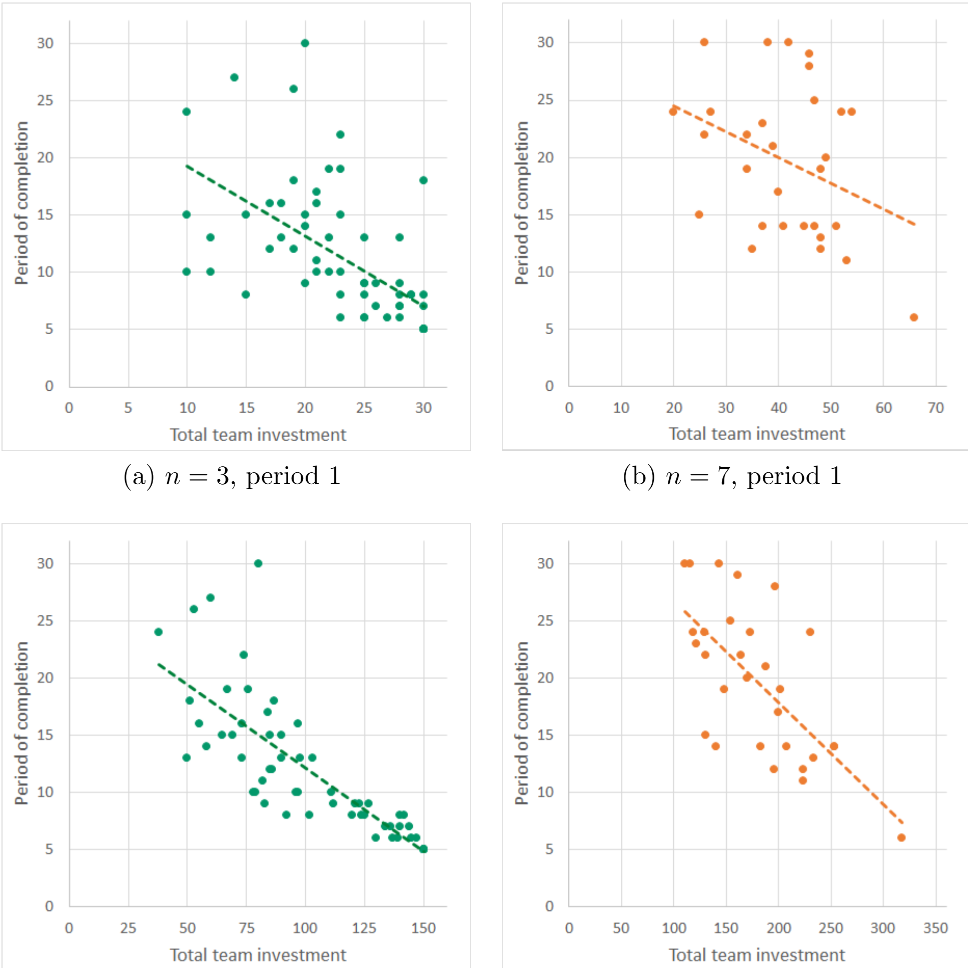
Figure 4.3 shows the share of subjects who systematically chose to free-ride throughout the entire first ten periods of a round. It also breaks down the share of systematic free-riders for periods one to five and periods six to ten. As with the findings in figure 4.4 for the total share of free-riders, the share of subjects who systematically chose to free-ride increased substantially from round to round and within each round. While only 5% of subjects chose to free-ride over the first ten periods of round one, this share increased to 21% in round three. Between periods one to five and periods six to ten, the number of systematic free-riders increased by 10 to 16 percentage points.

One possible explanation is that those who were willing to invest more became discouraged by the rather low average investment during the initial periods of round one and hence started reducing their own investments, leading to even more severe delays. Frustrated by their experience in the first round, participants invested much less right from the start in the second and third rounds. This led to even more severe delays and lower completion rates in the two latter rounds. On average, only 11% of participants did not free-ride at all within the first ten periods of a round. This share was fairly constant over all three rounds of this treatment.

Result 4:
The investment level in the first period had no significant effect on the completion rate. However, when extending the focus to the first five periods the effect becomes significant.

The average individual first-period investment for the three-person teams that completed their projects was 7.8 ECUs while for teams that did not complete their projects in time it was only 4.1 ECUs. The corresponding values for seven-person teams were 6.6 ECUs and 5 ECUs respectively. For small teams the difference between first-period investments of those teams that succeeded in completing a project and those that did not was statistically not significant ($p = 0.106$). However, when the focus is extended to the first five periods the difference in total investment on teams that did and those that did not complete their projects does become marginally significant ($p = 0.059$). In the case of large teams this difference in average first-period investment was not statistically significant ($p = 0.343$). Once again when the focus is extended to the first five periods the difference becomes significant ($p = 0.024$).

Taking a closer look at how first-period investment affects investment behavior in subsequent periods, we see two different effects. In some teams observing full investment by all teammates in the first period seems to have had a signaling effect and induced cooperative behavior. This was mainly observed for three-person teams. For some other teams, in particular seven-person teams, however, a very high first-period investment level was followed by a rapid decrease in total investments over the next few periods. In those teams, some members seem to have taken a very high first-period team investment as an opportunity to free-ride and rely on the investments of their teammates. Considering the first five periods instead of only the very first period accounts for these free-riders.



(a) $n = 3$, period 1

(b) $n = 7$, period 1

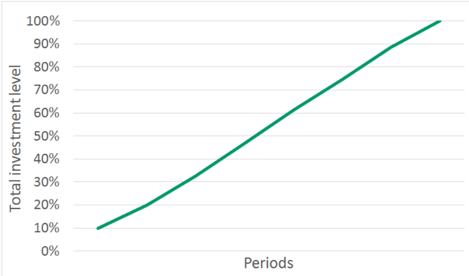
(c) $n = 3$, period 1 to 5

(d) $n = 7$, period 1 to 5

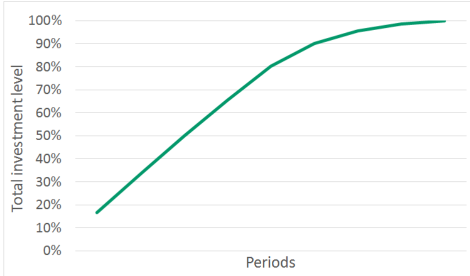
Figure 4.5: Effect of first-period and first-five-period investment on completion times

Figure 4.5a and 4.5b depict the correlation between teams' first-period investment and the time they needed for project completion.[10] As the trendline in those graphs indicates, the higher the first-period team investment, the less time teams need to complete a project. Figure 4.5c and 4.5d show the correlation between investments in the first five periods and project completion time. Compared to Figure 4.5a and 4.5b the data points are less scattered and we can clearly see the correlation between the investment levels of the first five periods and completion time.

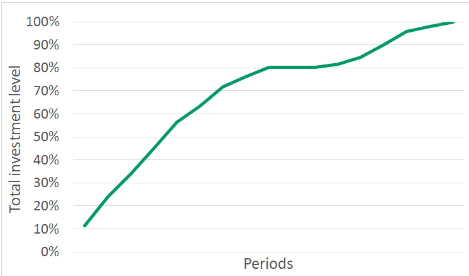
[10]For this analysis I excluded all teams that did not complete their projects in time. I also only looked at data from rounds two and three for better comparability as round one is somewhat special for seven-person teams.



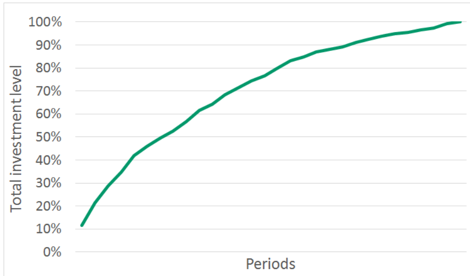
(a) Type 1 - short completion time



(b) Type 2 - median completion time



(c) Type 3 - median completion time



(d) Type 4 - long completion time

Figure 4.6: Four typical investment paths for aggregated team investments

The data also show that most teams' investments decreased over time regardless of the first-period investment level. In general, for a completion time above five periods four types of team investment paths were observed, as illustrated in Figure 4.6:

(a) Type 1 was observed for some small teams that completed their projects in six to nine periods. These teams usually did not start out with full investments (everyone investing 10 ECUs) in the first period, but then either increased their investments a bit or maintained fairly constant investment levels throughout the remaining rounds.

(b) Type 2 is typical of teams who have a completion time around the median value. As the figure shows, investments were fairly constant over the first few periods, then gradually decreased. This was the most frequently observed investment path for both small and large teams.

(c) Type 3 is a rather special case that did not occur very often. Here there seemed to be a revival phase or final sprint. After total investment had dropped to 0 ECU or a value close to it, one or two team members suddenly spurred investments again by making a rather high contribution.

(d) Type 4 is typical of teams that needed more than 20 periods to complete a project.

Their total investments gradually decreased over time, and in the last few rounds different team members alternately invested very low amounts.

4.5.3 Self-imposed deadline treatment

Result 5:

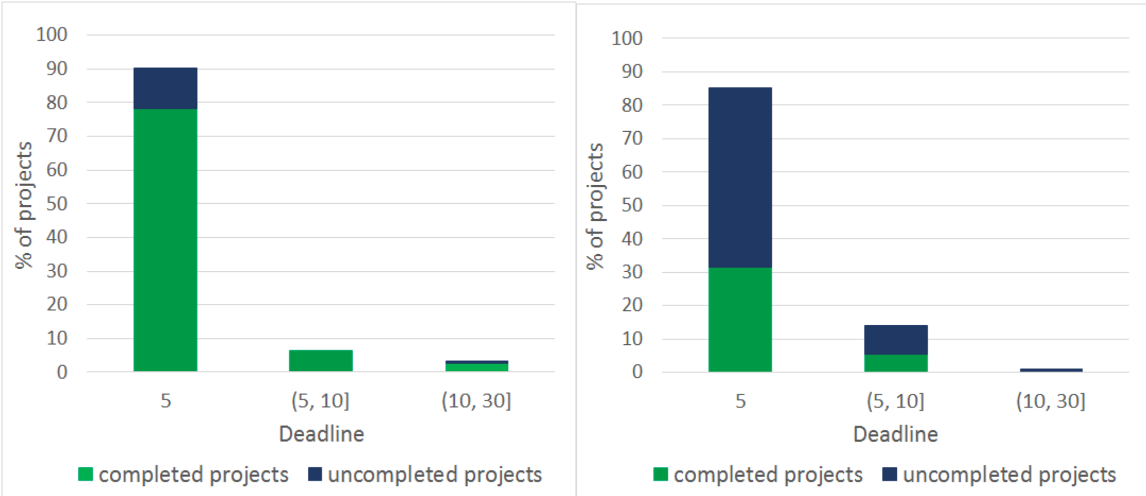
Regardless of team size, suggestions for the maximum number of periods very quickly converged to five, as did the deadlines actually implemented, indicating that participants were aware that completing a project in five periods was the efficient outcome.

In this treatment, where participants were able to influence the maximum number of periods they played, the suggestions for a deadline converged quickly to five. For three-person teams 61% of team members suggested a deadline of five periods in the first round. This share increased over subsequent rounds to 92% in round nine. For seven-person teams, 59% of participants suggested five as the maximum number of periods in the first round. Just as for small teams, this share then increased rapidly, to over 80%. In total, five was suggested as the maximum number of periods 80% of the time when $n = 3$ and 76% of the time when $n = 7$.

Similarly, the deadlines implemented by teams of both sizes converged to five. Almost all three-member teams played a five-period deadline from round two onwards. For $n = 7$, the deadline was five periods for over 80% of teams starting with round three. Aggregated over all nine rounds, 90% of small teams and 85% of large teams played a deadline of five periods (see Figure 4.7).

Result 6:

For small teams the completion rate was similar to that in the treatment without a deadline and introducing a self-imposed deadline increased efficiency as the average completion time was much shorter. However, large teams systematically failed to com-



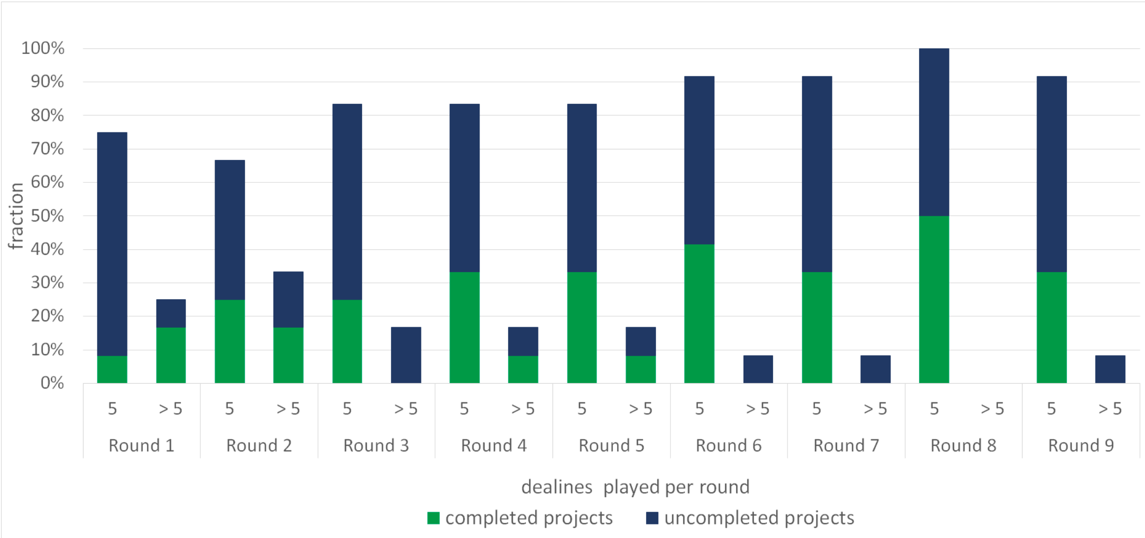
(a) $n = 3$. (b) $n = 7$.

Figure 4.7: Self-imposed deadlines and completion across all nine rounds

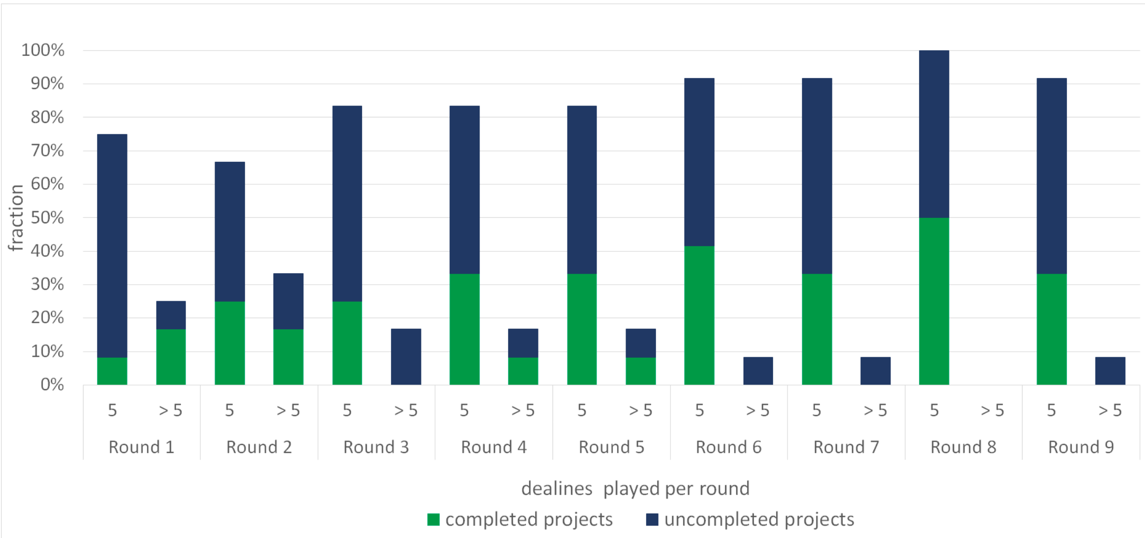
plete their projects within the self-imposed deadlines.

Although teams of both sizes almost immediately chose deadlines of five periods, team size still had a significantly negative effect on completion rates: While for $n = 3$ the average completion rate over all rounds was 87%, it was only 37% for $n = 7$. This difference was significant ($p = 0.030$). As we can see in Figure 4.7 the average completion rates did not differ much between projects with a deadline of five periods and those with a longer deadline. For small teams the average completion rate was 87% under a five-period deadline and 93.3% under longer deadlines. For large teams the average completion rate was 37% no matter what deadline was selected.

Looking more closely at how completion rates changed over time, we see no significant improvement from one round to the next (see Figure 4.8). As shown in Figure 4.8a the completion rates for small teams were rather stable throughout all nine rounds. Focusing solely on the completion rates when the deadline was five periods, we see no significant improvement. To test whether the completion rate for a deadline of five periods changed over the nine rounds, I again used the slope of a linear regression line for each independent observation to get the average per-period change, as explained above in Section 5.2. A Wilcoxon signed-rank test gives a p-value of 1, indicating that completion rates did not change over time. For large teams the completion rates were much more volatile, ranging from 0% to 67%. Although the average completion rate in rounds five to nine (42%) was higher than in rounds one to four (30%), this increase was not significant ($p = 0.198$).



(a) $n = 3$.



(b) $n = 7$.

Figure 4.8: Self-imposed deadline and completion rate over all nine rounds

Comparing this treatment with a self-imposed deadline to the treatment without a deadline, the average completion rates were quite similar for small teams (87% versus 91%). However, the average completion time was much lower when subjects had the opportunity to choose a deadline. In the treatment with a self-imposed deadline small teams completed their projects in 5.1 periods on average, twice as fast as in the treatment without a deadline (10.8 periods). Hence, introducing a self-imposed deadline increased efficiency for small teams. For teams that completed the project the DoE was almost 1, as most of them were able to coordinate on completing the project in five periods. This was an improvement compared to the treatment without a deadline, where the average DoE for completed projects was 0.9 ($p = 0.059$).[11]

For seven-person teams the average completion rate dropped by more than 20 percentage points once a self-imposed deadline was introduced compared with rounds two and three of the treatment without a deadline. This drop was not significant, however, as the p-value was 0.4. The overall DoE for seven-person teams was 0.71, which was higher than the DoE for rounds two and three in the treatment without a deadline (0.59). But this difference was not significant either ($p = 0.402$). By contrast, the DoE of teams that completed their projects was significantly higher (0.99) than in the treatment without a deadline ($p = 0.059$), as these teams completed their project in a much shorter time period. In the self-imposed treatment the average completion time was 5.5 periods, while in the last two rounds of the treatment without a deadline it was 19.7 periods.

Result 7:
There were four main reasons why participants withheld investments when the deadline was 5: (i) They did not want to play 5 as a deadline. (ii) They were expecting other team members not to invest. (iii) They were deliberately sabotaging the game. (iv) They were simply confused and made a mistake.

But why was the completion rate of large teams so low when five was implemented as a self-imposed deadline? One might expect completion rates to improve when teams select their own deadline. One possible explanation might be that when participants

[11] The DoE of all teams, including teams that did not complete their projects, was also higher than in the no deadline treatment (0.93 vs 0.86), though the difference was not statistically significant ($p = 0.178$).

suggested a deadline larger than five periods but were outvoted − as the median of all suggestions was imposed as the relevant deadline of the round they felt dissatisfied about being overruled or simply did not feel committed to the deadline. A closer look at the individual data reveals that 41% (34 out of 84) of the participants were overruled in the sense that they suggested a deadline larger than five but then had to play a deadline of five in at least in one of the nine rounds. Of these 34 participants 19 (56%) did not invest 10 ECUs when the deadline was five. This behavior also affected the investment decisions of the other team members negatively. In 76% of the uncompleted projects team members who suggested a deadline of five refrained from investing once they observed some other team members suggesting a deadline above five periods. Although some participants who were outvoted did cooperate once a deadline of five was selected, observing a suggestion larger than five was still taken as a negative signal and induced some very risk-averse players to withhold their investments.

What is rather puzzling is that not all teams that unanimously elected five as a deadline managed to complete their projects. In fact, only 65% of the projects were completed under these circumstances. Taking a closer look at the data on an individual level, we observe that twelve participants always suggested five as a deadline but then invested less than 10 ECUs in at least one round in which a deadline of five periods was implemented. Seven of these participants did this in more than one round. One might assume that these participants were among those who did not complete their projects within five periods when working alone in the single-player treatment. However, this was the case for only four of the twelve participants. The remaining eight participants all finished the single-player treatment in five periods.

These twelve players can be categorized into three groups. (i) Deliberate saboteurs: Four players seemed to deliberately sabotage the game as they always suggested a five-period deadline but then invested 0 ECU. It is not clear, why they behaved like this. (ii) Extremely risk-averse players: Five players fully invested in early rounds with a five-period deadline but the projects were not completed as someone else did not cooperate. In response these players stopped their investments in subsequent rounds. (iii) Confused players: The remaining three players were probably confused as they only deviated once and in all other rounds where five was the deadline they cooperated and invested fully.

For three-person teams, we also observed that not every project where every team member suggested five as a deadline was completed. There the completion rate was 83%, which was much higher than for seven-person teams. In total, there were five subjects who did not invest 10 ECUs when they and their teammates unanimously suggested a five-period deadline. Two of them did not complete the single-player treatment in five periods. But unlike the seven-person treatments there was only one subject who could be categorized as a "deliberate saboteur", and that only for the first four rounds. From round five onwards s/he invested 10 ECUs. One subject falls into the extremely risk-averse category, while the remaining three players were "confused".[12]

Result 8:
Team dynamics strongly affected project success. Subjects learned from experience and adapted their investments accordingly. However, they did not adjust their choice of deadlines.

It is evident from the data how important team dynamics were for the success of a project. In one session with seven-person teams two players deliberately sabotaged the game and two players who wanted a deadline longer than five periods reacted to being overruled by investing 0 ECU. This led to a downward spiral in investments. Once participants experienced defection in their teams they were frustrated and stopped cooperating in the next round because they expected some of their teammates not to cooperate. It turns out that four of the five extremely risk-averse players mentioned

[12] Two of the "confused" players only invested less than 10 ECUs in one round. Only one player seemed to have difficulties in choosing his/her investments. His/her investments ranged between 1 ECU and 9 ECUs for a deadline of five.

above were in this session too. Hence, not a single project was completed in this session.[13]

However, the opposite was also observed. Once subjects experienced cooperation in the sense that everyone consistently invested 10 ECUs per period in a round with a five-period deadline, they continued to cooperate in all subsequent rounds with a five-period deadline. Thus, although teams were reshuffled after each round, participants still seem to have based their decisions on what they experienced in previous rounds and adjusted their investments accordingly. This behavior was observed in 50% of the large-team treatments and in 80% of the small-team treatments.

One very surprising observation is that subjects kept suggesting five as the maximum number of periods, even after their teams were repeatedly unsuccessful in completing projects within the minimum number of periods. One would assume that people would learn from their experience and make suggestions greater than five after they had experienced failure a few times in a row. In total, 70% of the participants in large teams encountered failure in two consecutive rounds with a deadline of five and 90% of these participants nevertheless kept suggesting five as a deadline.

4.5.4 Exogenous deadline treatment

Result 9:

Completion rates were lower in the treatment with exogenous deadlines than in the treatment with self-imposed deadlines. As before, the effect of team size on completion rates was negative.

In the exogenous treatment, teams were given a binding deadline.[14] As before, the completion rate was much higher in small groups regardless of the deadline, as

[13]Excluding the data from this one session leads to an overall completion rate of 44%.

[14]For the six rounds played in this treatment the deadlines were 5, 6, 10, 5, 8 and 5.

Figure 4.9 and Table 4.5 show. Over all six rounds, the average completion rate for $n = 3$ was 69%; for $n = 7$ it was only 33%. However, irrespective of the externally imposed deadline, the difference between these two team sizes was not as large as in the treatment with self-imposed deadlines and only marginally significant ($p = 0.077$).

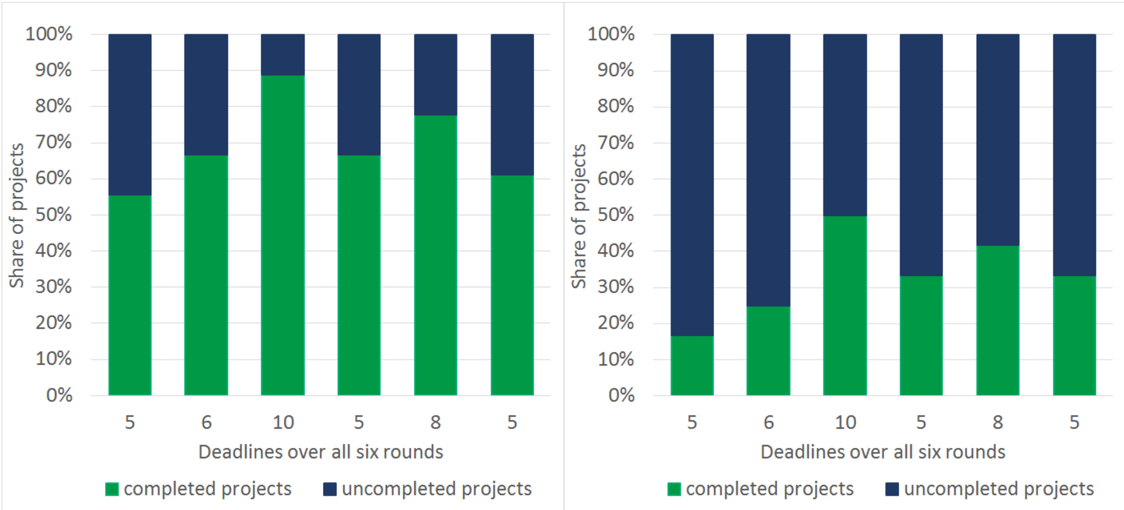
Looking at the different deadlines separately in Figure 4.9, it seems the more time teams had to complete a project, the higher the completion rate. For small teams the average completion rate was 61% when the deadline was five (see Table 4.5).[15] The completion rate peaked at 88% when the deadline was ten. This difference between the completion rate when the deadline was ten and when it was five was significant ($p = 0.034$). In contrast to small teams, having more time did not significantly improve the completion rates of large teams. There the average completion rate was 28% over the three rounds where the deadline was five. Although the completion rate with a deadline of ten was 50%, the difference was not statistically significant ($p = 0.269$).

Focusing on the three rounds where five was imposed as the relevant deadline, we saw no improvement in completion rates over time either for large or for small teams. As the deadline of five was always played in three rounds (rounds one, four and six), I used the same regression coefficient method as before to test for possible improvements in the completion rate. By calculating the regression coefficient with the number of completed projects as the dependent variable and testing it against the null hypothesis that the difference between the rounds is zero, I obtain a p-value of 0.890 for small teams and 0.586 for large teams. This suggests that the average completion rate did not systematically increase between the three rounds with five as a deadline.

Compared with the setting where teams selected their own deadline, the completion rates were noticeably but statistically not very significantly lower for small teams ($p = 0.064$).[16] For large teams the completion rates of both treatments – self-imposed deadline (37%) and exogenous deadline (33%)– are not significantly different ($p = 0.686$). Table 4.5 shows that for a deadline of five the completion rates were much

[15] 61% is the average over all three rounds where five was imposed as a deadline.

[16] In the SID treatment the average completion rate for small teams was 87%.



(a) $n = 3$. (b) $n = 7$.

Figure 4.9: Completion rates under exogenous deadlines

	Deadline	ED completion rate	ED # projects	SID completion rate	SID # projects
	5	61%	54	87%	141
$n = 3$	[6, 10]	78%	54	100%	10
	> 10	–	–	80%	5
	5	28%	36	37%	92
$n = 7$	[6, 10]	39%	36	40%	15
	>10	–	–	0%	1

Notes: # projects = total number of projects.

Table 4.5: Comparison of completion rates under self-imposed and exogenous deadlines

lower in the exogenous deadline treatment for both team sizes. However this difference was only marginally significant for small teams ($p = 0.065$) and not significant at all for large teams ($p = 0.678$). Due to the low number of observations for certain deadlines, it is somewhat difficult to compare completion rates separately for the two treatments for each deadline length. Nevertheless, if we look at all the deadlines between six and ten periods, we see that for three-person teams the completion rate was significantly higher in the self-imposed deadline treatment ($p = 0.016$) whereas for seven-person teams the difference between the two treatments is not significant (39% versus 40%).

As for efficiency, the average *DoE* for small teams was 0.83 and for large teams it was 0.78. These values are quite similar to those in the team treatment with a self-imposed deadline. If only teams that completed their projects are taken into account, the DoE increases to 0.97 for small teams and 0.98 for large ones. As these participants were restricted by deadlines there was no room for procrastination. Hence, the teams that managed to complete their projects did not encounter large delay costs. This also applied to teams that were unable to complete their projects. Because rounds ended as soon as the deadline was reached, projects could not be dragged on. This prevented players from "wasting assets" by continuing to invest in uncompletable projects.

Overall it seems that for shorter deadlines, completion rates were higher when teams coordinated on the deadline themselves. This might be because subjects who have a say in setting a deadline will feel more committed to it. In addition, they might be more aware of the optimal number of periods and how much to contribute. Furthermore, when participants in a team see the other team members arguing for a particular deadline, they can be more certain that their teammates will cooperate and fully invest in completing the project. This sense of assurance, though no guarantee for success, is absent when the deadline is externally imposed.

4.6 Conclusion

In this experimental study I analyze whether inefficient procrastination occurs in long-term team projects when no strict deadlines are given. To investigate the effects of team size on investment behavior and completion rates I distinguish between single-players ($n = 1$), small teams ($n = 3$) and large teams ($n = 7$). In addition, the efficiency of deadlines was examined while differentiating between self-imposed and exogenous deadlines.

The setting involves participants making effort decisions over several periods in order to complete a project with a contribution threshold. I find that in the absence of deadlines, participants are rather efficient when working alone. In the single-player treatments all subjects managed to complete their projects and very little procrastination was observed. Once participants were assigned to teams however, inefficient procrastination did occur, and completion rates were negatively affected. Delays were much longer and completion rates significantly lower in large teams than in the other two treatments. While single players completed projects in 6 periods on average, small teams needed almost twice as long and large teams needed twice as long as small teams. A strong tendency to free-ride was observable in large teams − a tendency that increased from round to round and over time within individual rounds. This affected completion rates negatively, especially in the later rounds of the team treatment without a deadline.

In the self-imposed deadline treatment, teams were allowed to select a deadline using a mechanism where each team member made a suggestion for the maximum number of periods played and the median was then taken as the deadline. In this treatment, suggestions quickly converged to five, the minimum number of periods teams needed to complete a project. In fact, five was chosen as a deadline for over 80% of the projects. Almost none of the self-imposed deadlines was above ten periods, showing that the majority of participants realized what option would be efficient to play. This strong convergence towards five periods as a deadline was independent of team size. However, while introducing a deadline can mitigate procrastination, it can have a negative effect on completion rates, especially when the deadline is very tight. For small teams the completion rate did not change when a deadline was introduced, but for large teams the average completion rate was lower than in the treatment without a deadline. Large teams are particularly prone to failure when deadlines are short as the non-cooperative behavior of a single team member can lead to non-completion of a project. It was rather surprising to see that even after they experienced failure in previous rounds subjects kept suggesting five periods as a deadline instead of adjusting their suggestions.

Comparing the two different deadline types, completion rates were higher when deadlines were self-imposed than when they were exogenous. This difference, however, was only marginally significant for three-person teams and not statistically significant for seven-person teams.

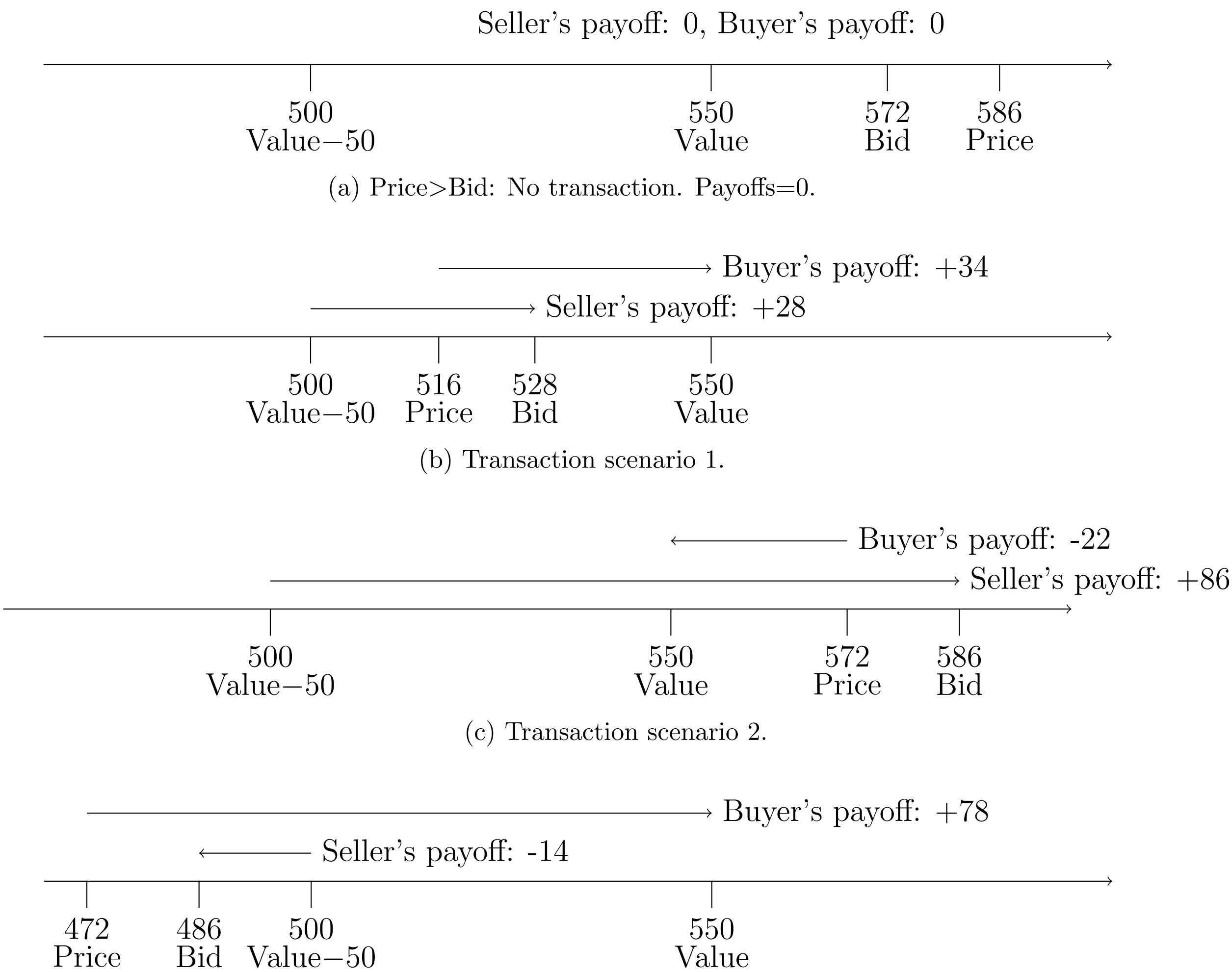
To sum up, inefficient procrastination was observed in the team treatments. Introducing deadlines can help with procrastination, but it can negatively affect project completion when deadlines are too tight. Aiming for the most efficient outcome, e.g. by setting the deadline to dive periods in this experiment, only improved the outcome when all team members were willing to contribute fully. For teams that were able to complete their projects within the given deadline, efficiency increased considerably as they encountered much lower delay costs. Overall, team size had a strong negative effect on completion time and completion rates regardless whether a deadline was introduced or not. One modification of the self-imposed deadline treatment worth investigating further would be to assign subjects to teams according to their suggestions for a specific deadline. Participants who suggested the same maximum number of periods could be matched together, which might help to increase completion rates and improve efficiency. For example, if only subjects who suggested a 5-period deadline were assigned to a group, there would be no dissatisfaction among participants over being "outvoted", increasing the probability of success.

Introducing a deadline is only one of many ways to prevent procrastination. Penalizing delay more severely by increasing the delay costs, introducing more transparency in effort choices or even allowing for communication are, among others, also possible ways to improve the efficiency in long-term team projects. For future research it would be interesting to investigate these methods and how they affect the investment dynamics of teams.

Appendices

A Appendix Chapter 2

A.1 Examples of Possible Outcomes



(a) Price>Bid: No transaction. Payoffs=0.

(b) Transaction scenario 1.

(c) Transaction scenario 2.

(d) Transaction scenario 3.

Figure A.1: Transaction scenarios.

A.2 Simulation of Publication Bias

The benchmark values of the difference between the kth-highest evidence and the true value v in the first line of panel (a) in table 2.4 do not take into account that $v \in [200, 1000]$. In order to analyze the effect of such boundaries on the inference, we simulate the distribution of the true value given the realization of a certain kth highest evidence with the help of a set of 10000 draws of 10 evidences for each true value between 200 and 1000.

For each realization of kth highest evidence between 200 and 500, we calculate the difference between the true means in this simulation and the benchmark value. If the deviation is different from 0 the boundaries for the true values are still at play, causing a skewed distribution of the true means underlying the observed kth evidence. The results of these simulations are depicted in figure A.2.

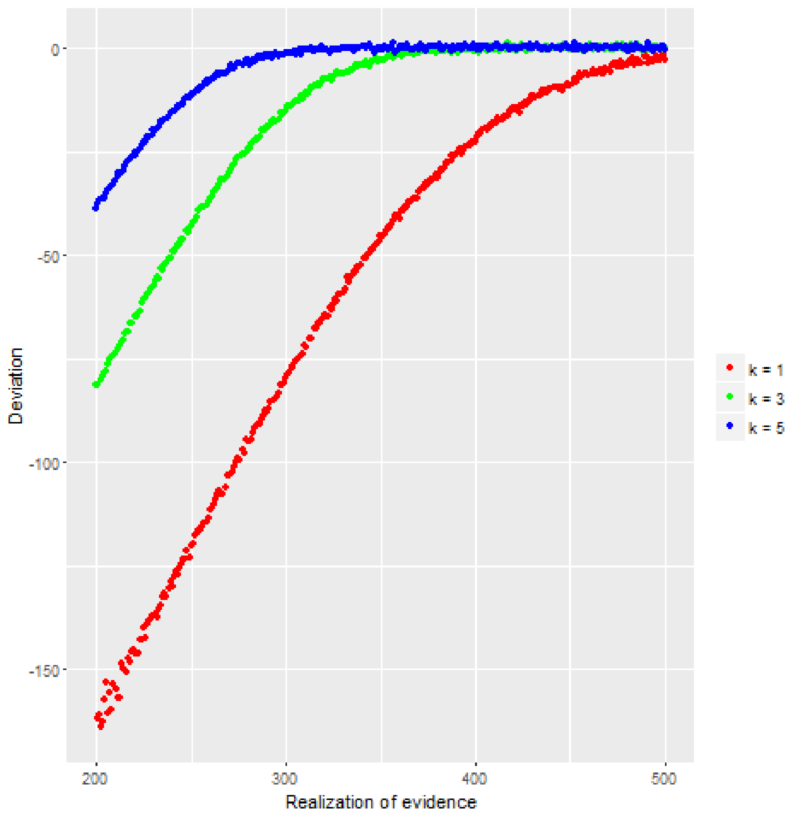


Figure A.2: Deviation from benchmark value for realized levels of the kth evidence close to the lower bound of v, 200.

A.3 Screenshots

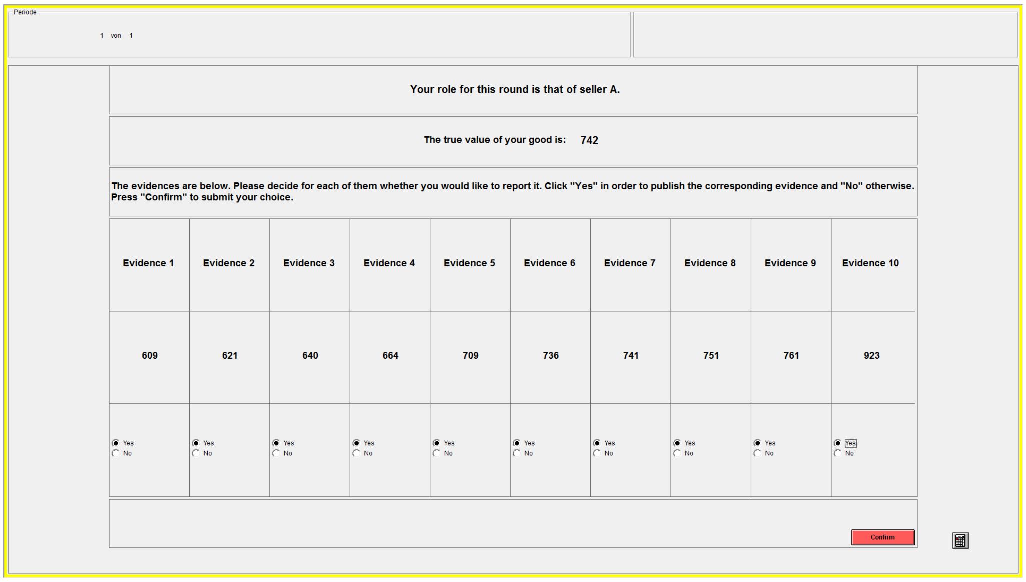


Figure A.3: Sellers' screen in C10

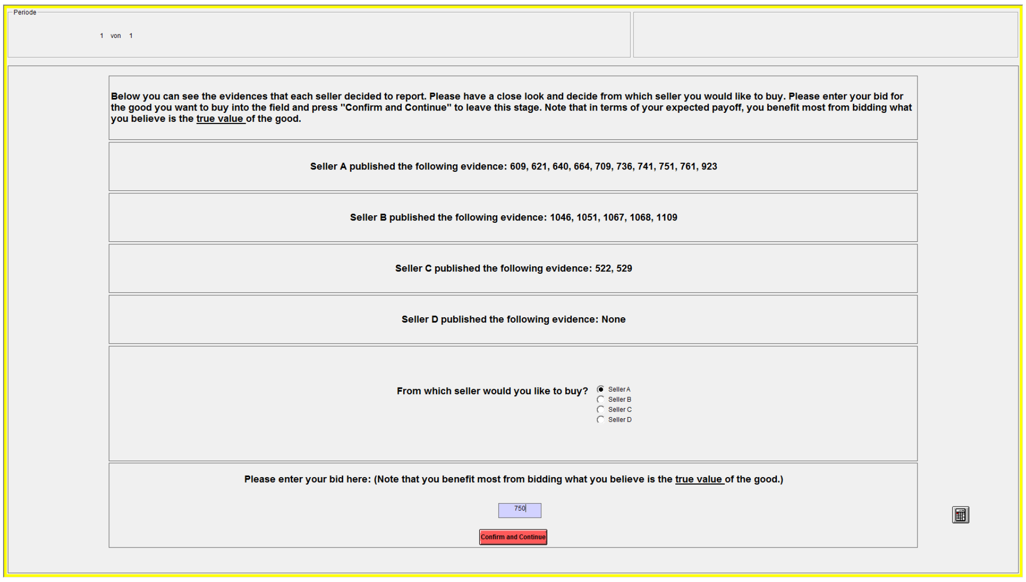


Figure A.4: Buyers' screen in C10

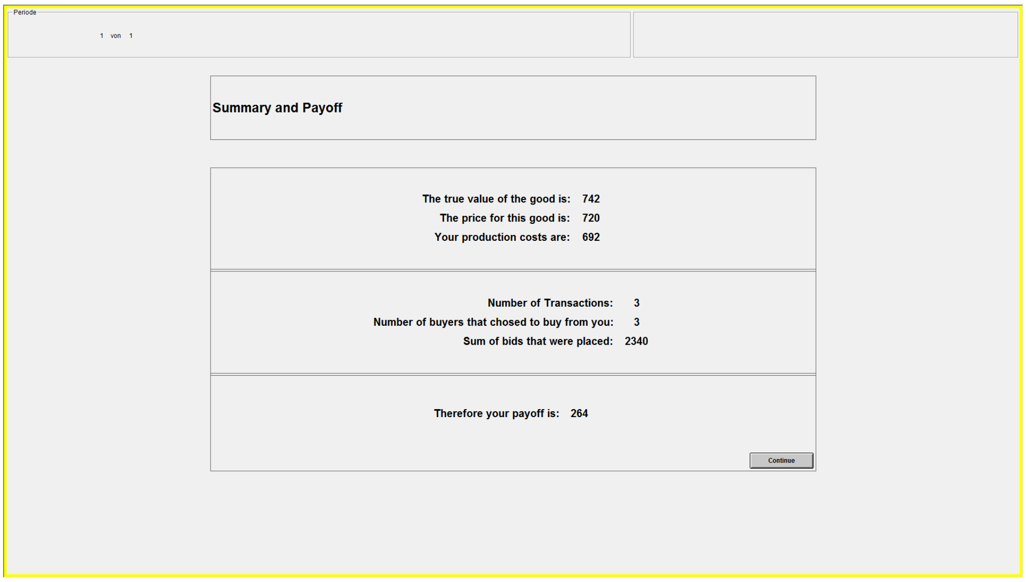


Figure A.5: Sellers' feedback screen in C10

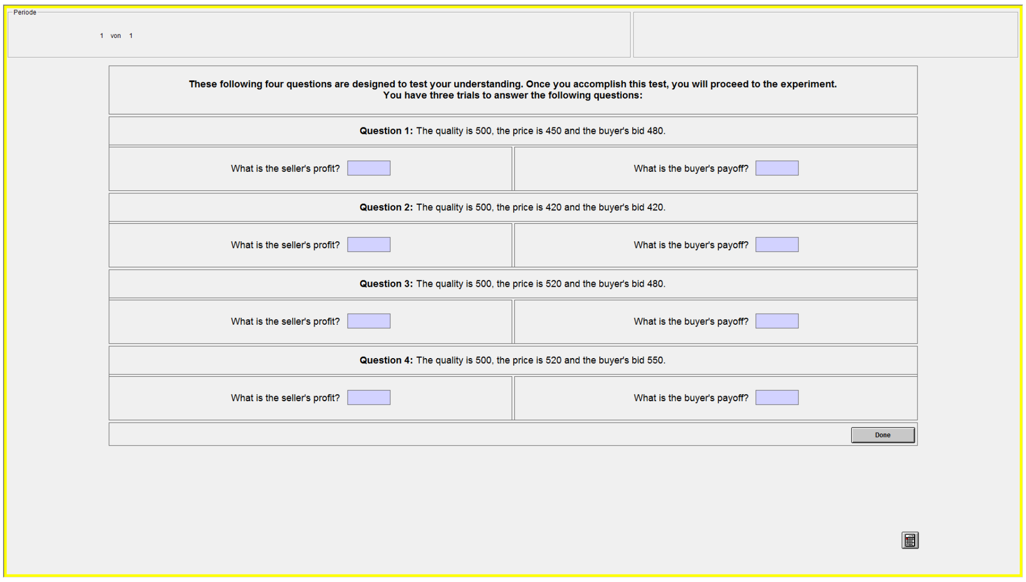


Figure A.6: Screen of the understanding test

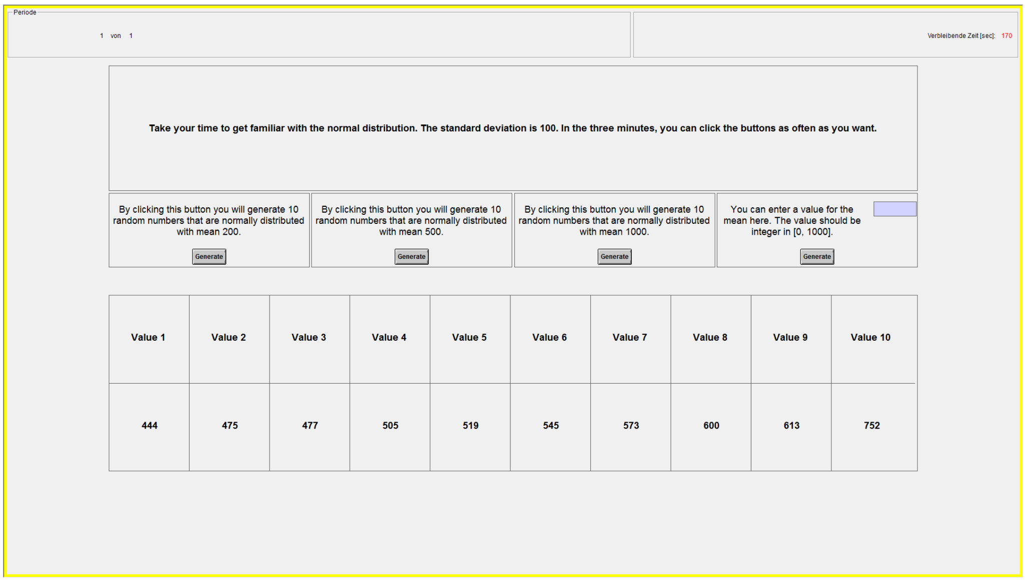


Figure A.7: Tool to generate 10 draws from a normal distribution $N(v, 100)$, $v \in [0, 1000]$. Here: $v = 500$

A.4 Instructions for C10 - M10

You are about to participate in an experiment in a market setting. You may earn a considerable amount of money. The amount you earn will depend on your decisions and the decisions of others, so please follow the instructions carefully. All that you earn is yours to keep, and will be paid to you in private, in cash, at the end of today's session.

During the experiment your payoffs are denominated in points. Your point earnings will be converted to cash at the end of today's session at an exchange rate of 60 points = 1 Euro. No matter what your payoffs are in the experiment, you will be paid at least 2 Euro.

It is important to us that you remain silent and do not look at other people's screens. If you have any questions or need assistance of any kind, please raise your hand, and an experimenter will come to you. If you talk, exclaim out loud, etc., you will be kindly asked to leave.

The experiment consists of two parts (**Part I**, **Part II**) which are independent of each other. Each of these parts, in turn, consists of up to 12 rounds.

Part I

In this part of the experiment, four Sellers and four Buyers form a market. The Seller knows the Value (in points) of his/her good but cannot report it to the Buyers. The Seller has requested 10 external test institutes to officially evaluate the Value of the good and can indeed report these 10 official "Evidences" about the Value to the Buyers. The external Evidences are informative about the good's Value but noisy. In particular, they follow a normal (Gauss+ian) distribution around the Value with a constant standard deviation of 100. The standard deviation measures the dispersion of the Evidences, how far away they are from the Value. Figure A.8 shows that Evidences are more likely to be close to the Value than far away. You will be able to get familiar with this distribution later.

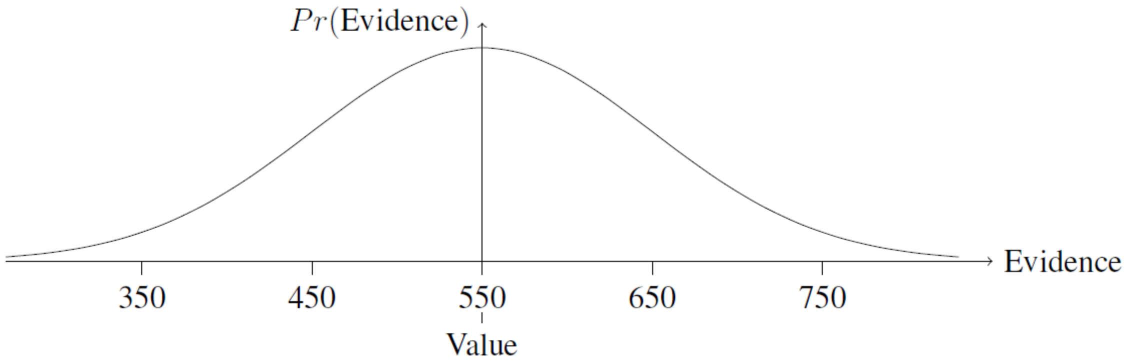


Figure A.8: Normal probability distribution with a Value of 550 and a standard deviation of 100, indicating the probability of Evidences.

At the beginning of a round, each Seller is informed about the true Value of her/his good. The true Value lies between 200 and 1000 points, each Value level in this interval is equally likely. As we said, the Seller cannot report the true Value to the Buyers, but she/he can choose for each of the 10 Evidences whether to report it to the Buyers or not. The Seller cannot change or manipulate the Evidences in any way.

After the Sellers' choice, the Buyers will see the reports of all 4 Sellers. Sellers as well will be informed about the reports of the other Sellers. Before the Buyers place their bid, they have to chose from which Seller to buy. For the chosen Seller's good, the Buyer places a Bid. The Bid has to be an integer value between 0 and 1200.

When does the transaction take place? The computer generates randomly a Price of the Seller's good. Neither Seller nor Buyer will be informed about the Price when they take their decisions. The Price takes integer values between the Value minus 200 and the Value plus 200, with each Price level being equally likely (figure A.9). The transaction takes place whenever the Buyer's Bid is greater than or equal to the Price.

So, in our example, a Bid of

1. 349 will never lead to a transaction, since the Price is certainly above.

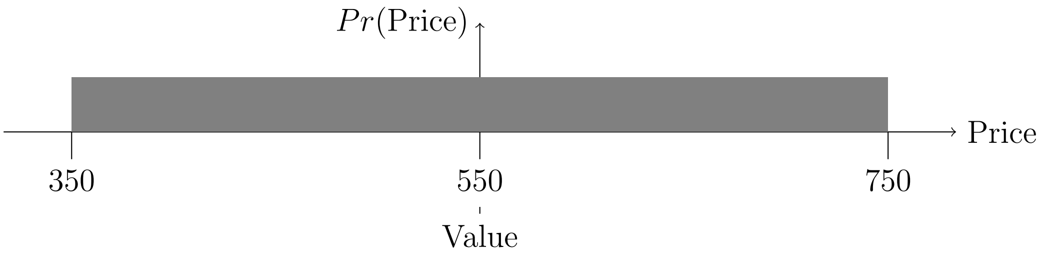


Figure A.9: Price distribution around the Value of 550.

2. 750 will always lead to a transaction since the Price is certainly below or the same.

3. 550 corresponding to the Value will lead to a transaction with 50% chance.

How are the Seller's and the Buyer's payoff determined? First, if no transaction takes place, both Seller and Buyer get a payoff of 0 points. If a transaction takes place, as we said depending on the Buyer's Bid and the random Price, the Seller produces the good at the cost of the Value minus 50. The Seller sells one item of their good to each Buyer whose Bid exceeds the Price. A Seller might sell to none, one, two, three, or four Buyers. The Seller's payoff per transaction at the end of the round will be the Bid placed by the Buyer minus the cost:

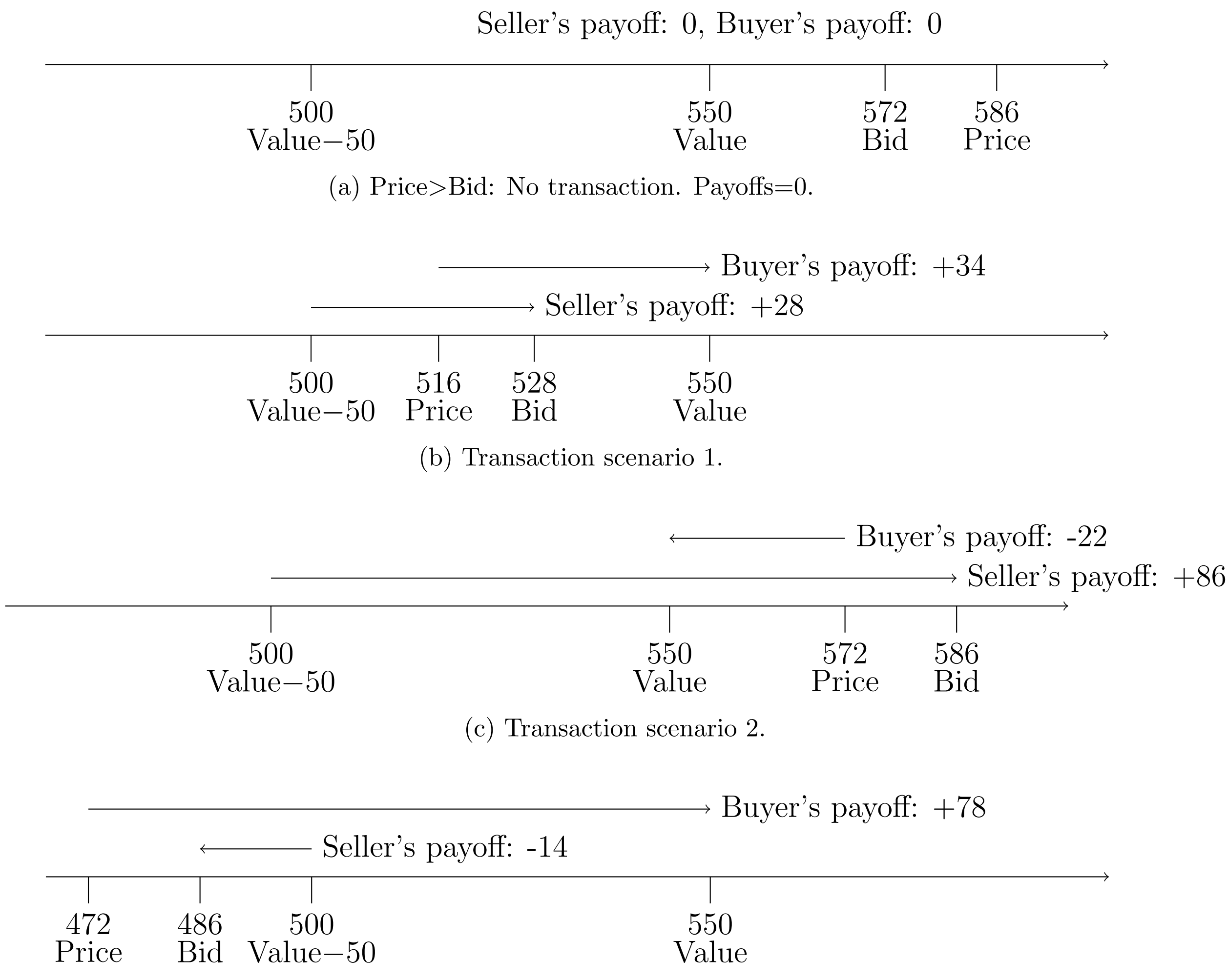
$$\mathbf{Payoff_{Seller} = Bid_{Buyer} - (Value - 50).}$$

The Buyer's payoff is the true Value of the good minus the Price:

$$\mathbf{Payoff_{Buyer} = Value - Price.}$$

Notice that a Bid equal to the Value will ensure the Buyer to never make losses. If the Price was higher than the Bid=Value, the transaction would not take place. Recall from 3. that under a Bid=Value, the transaction does not take place half the times. Further, note that in terms of expected payoff, the Buyer benefits most from bidding what he believes the Value of the item is. The bids are limited to be between 0 and 1200. Figures A.10a to A.10d present various scenarios.

Part I of this experiment consists of 2 practice rounds and 2 blocks of 5 experiment rounds. At the beginning of each round, you will be informed about the randomly



(a) Price>Bid: No transaction. Payoffs=0.

(b) Transaction scenario 1.

(c) Transaction scenario 2.

(d) Transaction scenario 3.

Figure A.10: Transaction scenarios.

chosen role (Seller or Buyer) that you take. You keep this role in the first block of 5 rounds, and take on the other role in the second block. You will keep the same role within a block, but you will face randomly chosen market counterparts. Throughout, four Sellers and four Buyers will form a market.

In order to participate in the experiment, you will go through a brief understanding test. Here and throughout the experiment, you can access a calculator via a button in the right bottom corner of your screen. Once everybody accomplishes this test, you can get more familiar with the normal distribution with standard deviation of 100. For that purpose, you will have three minutes to simulate as often as you want the process of generating 10 Evidences for different Values. The experiment will start with the 2 practice rounds that are not paid. Finally, you proceed to the paid rounds.

Are there any questions? If not, please turn to your screens and follow carefully the instructions there.

Part II

You are about to start Part II of the experiment, which consists of no practice rounds and 2 blocks of 5 experiment rounds. Like before, in each block of 5 rounds you will take the same role, and you will face randomly chosen market counterparts.

In this part, one market will consist of one Seller and one Buyer. Just like before, the Seller first chooses the Evidences s/he wants to report. The Buyer will observe only the Evidences that the Seller chose to report. S/He then places a bid for the good. The Buyer's and the Seller's payoffs are determined in the same fashion as before.

A.5 Instructions for M15 - C15

You are about to participate in an experiment in the economics of decision making in a market setting. You may earn a considerable amount of money. The amount you earn will depend on your decisions and the decisions of others, so please follow the instructions carefully. All that you earn is yours to keep, and will be paid to you in private, in cash, at the end of today's session.

During the experiment your payoffs are denominated in points. Your point earnings will be converted to cash at the end of today's session at an exchange rate of 60 points = 1 Euro. No matter what your payoffs are in the experiment, you will be paid at least 2 Euro.

It is important to us that you remain silent and do not look at other people's screens. If you have any questions or need assistance of any kind, please raise your hand, and an experimenter will come to you. If you talk, exclaim out loud, etc., you will be kindly asked to leave.

The experiment consists of two parts (**Part I**, **Part II**) which are independent of each other. Each of these parts, in turn, consists of up to 12 rounds.

Part I

In this part of the experiment, one Seller and one Buyer form a market. The Seller knows the Value (in points) of his/her good but cannot report it to the Buyer. The Seller has requested 10 external test institutes to officially evaluate the Value of the good and can indeed report these 10 official "Evidences" about the Value to the buyer. Additionally, the Seller has the possibility to ask 5 more test institutes to evaluate the Value of his/her good at a package price of 15 points. These 5 additional Evidences can also be reported to the Buyer. The external Evidences are informative about the good's Value but noisy. In particular, they follow a normal (Gaussian) distribution around the Value with a constant standard deviation of 100. The standard deviation measures the dispersion of the Evidences, how far away they are from the Value. Figure A.11 shows that Evidences are more likely to be close to the Value than far away. You will be able to get familiar with this distribution later.

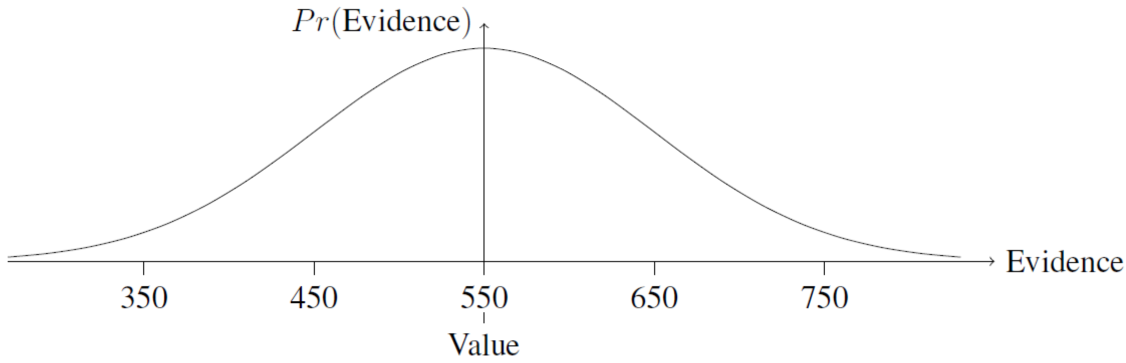


Figure A.11: Normal probability distribution with a Value of 550 and a standard deviation of 100, indicating the probability of Evidences.

At the beginning of a round, the Seller is informed about the true Value of her/his good. The true Value lies between 200 and 1000 points, each Value level in this interval is equally likely. After observing the initial 10 Evidences, the Seller has the opportunity to get 5 additional Evidences for a price of 15 points. As we said, the Seller cannot report the true Value to the Buyer, but s/he can choose for each of the 10 (or 15) Evidences whether to report it to the Buyer or not. The Seller cannot change or manipulate the Evidences in any way.

After the Seller's choice, the Buyer observes only the Evidences that the Seller chose to report. S/He then places a Bid for the good. The Bid has to be an integer value between 0 and 1200.

When does the transaction take place? The computer generates randomly a Price of the Seller's good. Neither Seller nor Buyer will be informed about the Price when they take their decisions. The Price takes integer values between the Value minus 200 and the Value plus 200, with each Price level being equally likely (figure A.12). The transaction takes place whenever the Buyer's Bid is greater than or equal to the Price.

So, in our example, a Bid of

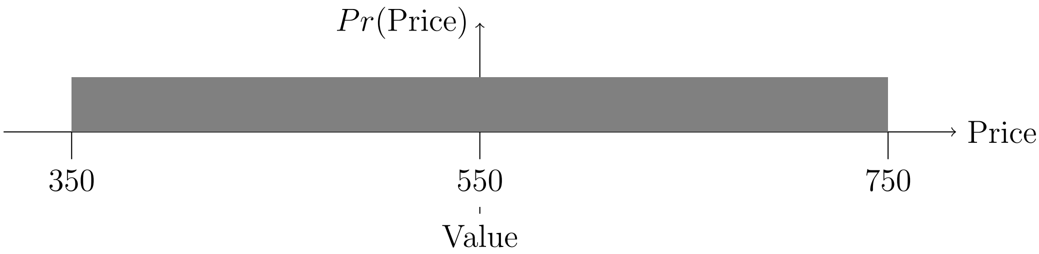


Figure A.12: Price distribution around the Value of 550.

1. 349 will never lead to a transaction, since the Price is certainly above.
2. 750 will always lead to a transaction since the Price is certainly below or the same.
3. 550 corresponding to the Value will lead to a transaction with 50% chance.

How are the Seller's and the Buyer's payoff determined? First, if no transaction takes place, the Buyer gets a payoff of 0 points. In case the Seller didn't purchase additional Evidences her/his payoff is 0 points as well. Otherwise her/his payoff is -15 points. If the transaction takes place, as we said depending on the Buyer's Bid and the random Price, the Seller produces the good at the cost of the Value minus 50. The Seller's payoff at the end of the round will be the Bid placed by the Buyer minus the cost (production cost and possibly cost from purchasing 5 additional Evidences):

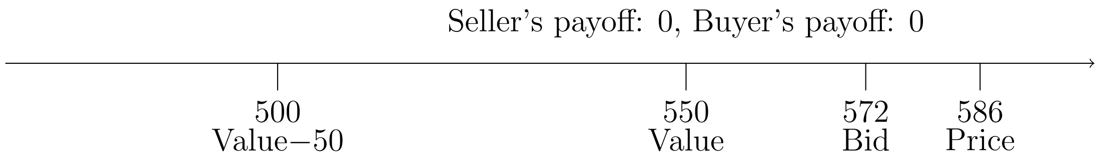
$$\mathbf{Payoff_{Seller}} = \begin{cases} \mathbf{Bid_{Buyer} - (Value - 50)} & \text{without purchase of additional Evidences} \\ \mathbf{Bid_{Buyer} - (Value - 50) - 15} & \text{with purchase of additional Evidences} \end{cases}$$

The Buyer's payoff is the true Value of the good minus the Price:

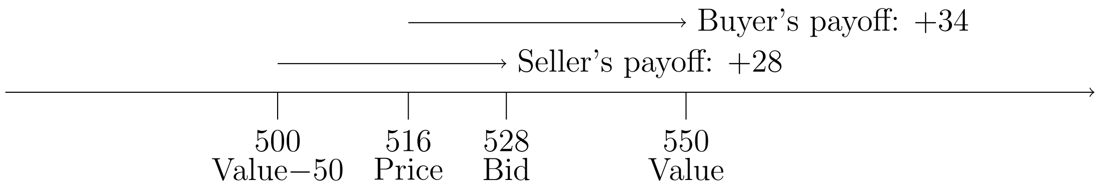
$$\mathbf{Payoff_{Buyer} = Value - Price.}$$

Notice that a Bid equal to the Value will ensure the Buyer to never make losses. If the Price was higher than the Bid=Value, the transaction would not take place. Recall from 3. that under a Bid=Value, the transaction does not take place half the times.

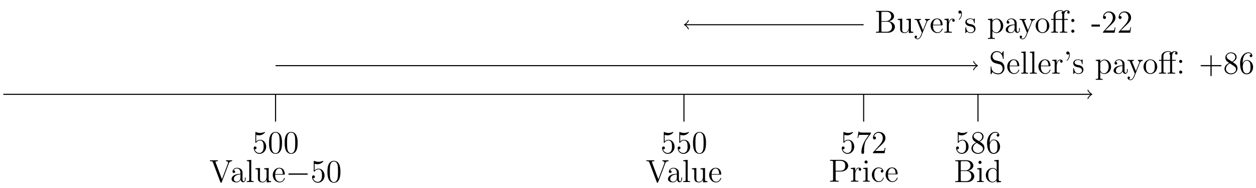
Further, note that in terms of expected payoff, the Buyer benefits most from bidding what he believes the Value of the item is. The bids are limited to be between 0 and 1200. Figures A.13a to A.13d present various scenarios. In these scenarios the purchasing of additional Evidences is not considered.



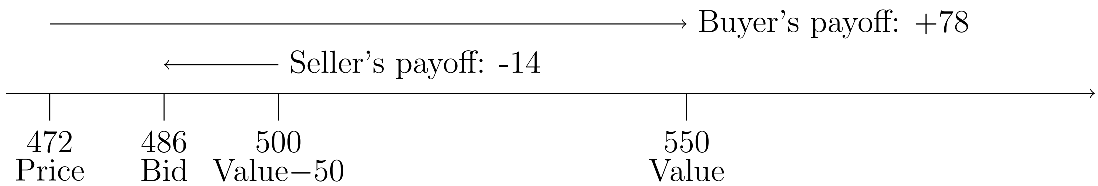
(a) Price>Bid: No transaction. Payoffs=0.



(b) Transaction scenario 1.



(c) Transaction scenario 2.



(d) Transaction scenario 3.

Figure A.13: Transaction scenarios.

Part I of this experiment consists of 2 practice rounds and 2 blocks of 5 experiment rounds. At the beginning of each round, you will be informed about the randomly chosen role (Seller or Buyer) that you take. You keep this role in the first block of 5 rounds, and take on the other role in the second block. You will keep the same role within a block, but you will face randomly chosen market counterparts. Throughout, one seller and one buyer will form a market.

In order to participate in the experiment, you will go through a brief understanding test. Here and throughout the experiment, you can access a calculator via a button in the right bottom corner of your screen. Once everybody accomplishes this test, you can get more familiar with the normal distribution with standard deviation of 100. For that purpose, you will have three minutes to simulate as often as you want the process of generating 10 Evidences for different Values. The experiment will start with the 2 practice rounds that are not paid. Finally, you proceed to the paid rounds.

Are there any questions? If not, please turn to your screens and follow carefully the instructions there.

Part II

You are about to start Part II of the experiment, which consists of no practice rounds and 2 blocks of 5 experiment rounds. Like before, in each block of 5 rounds you will take the same role, and you will face randomly chosen market counterparts.

In this part, one market will consist of 4 Sellers and 4 Buyers. Just like before, the Sellers first decide whether they want to purchase 5 additional Evidences. Then they choose the Evidences they want to report. The Buyers will see the reports of all 4 Sellers. Sellers as well will be informed about the reports of the other Sellers. Before the Buyers place their bid, they have to chose from which Seller to buy. For the chosen Seller's good, the Buyer places a bid and the Buyer's payoff is determined in the same fashion as before. The payoff of the Sellers is as well determined in the same fashion as before. The Seller sells one item of their good to each Buyer whose Bid exceeds the Price. A Seller might sell to none, one, two, three, or four Buyers.

B Appendix Chapter 3

B.1 Screenshots

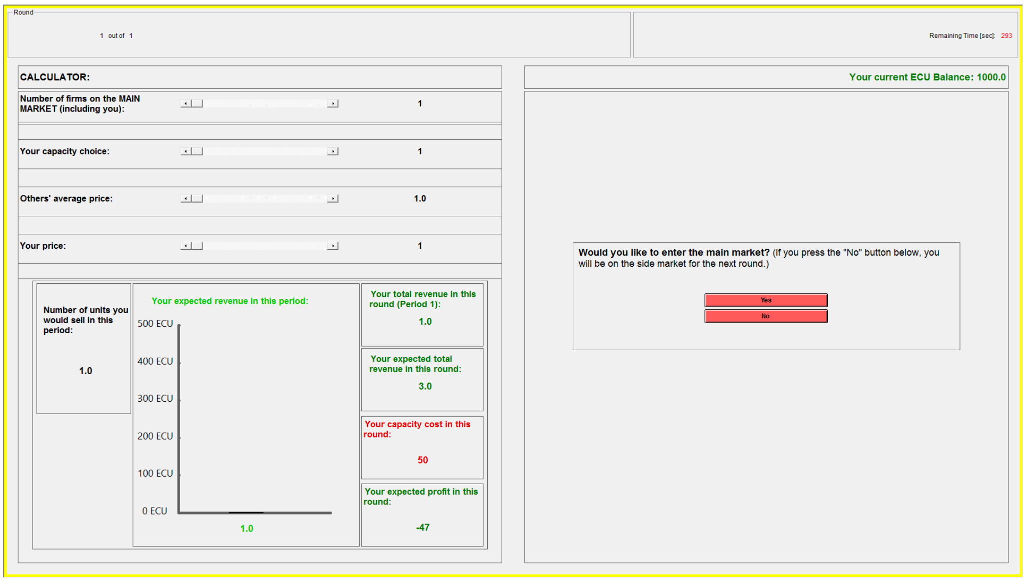


Figure B.1: Screen in entry stage in Market Force treatment

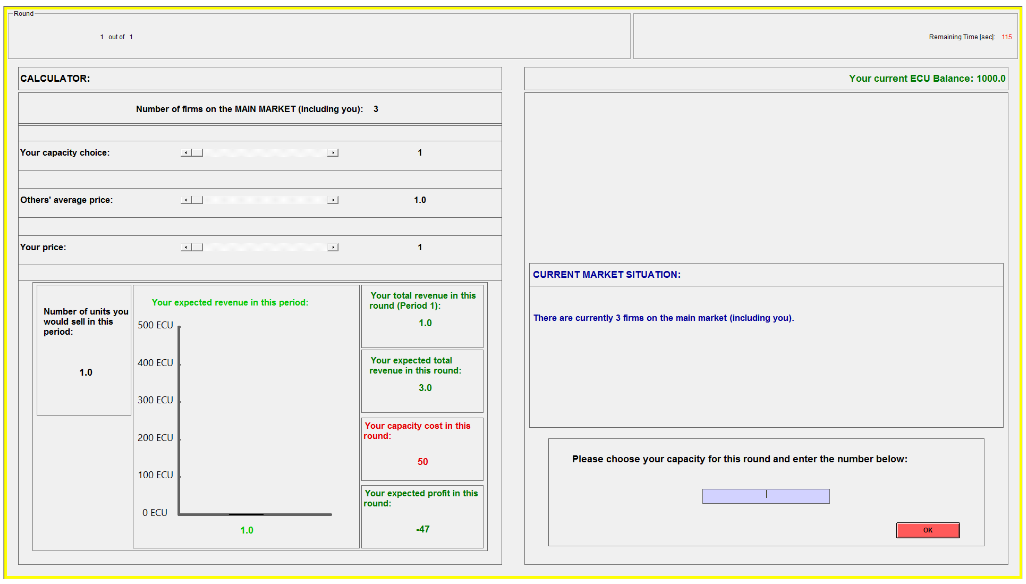


Figure B.2: Screen in capacity choice stage – main market

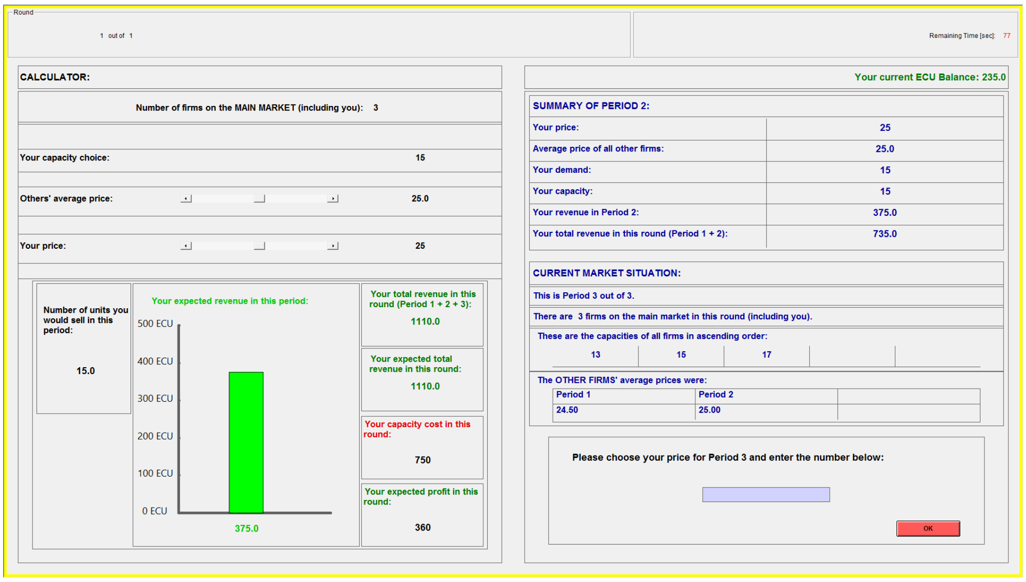


Figure B.3: Screen in the third period of the pricing stage – main market

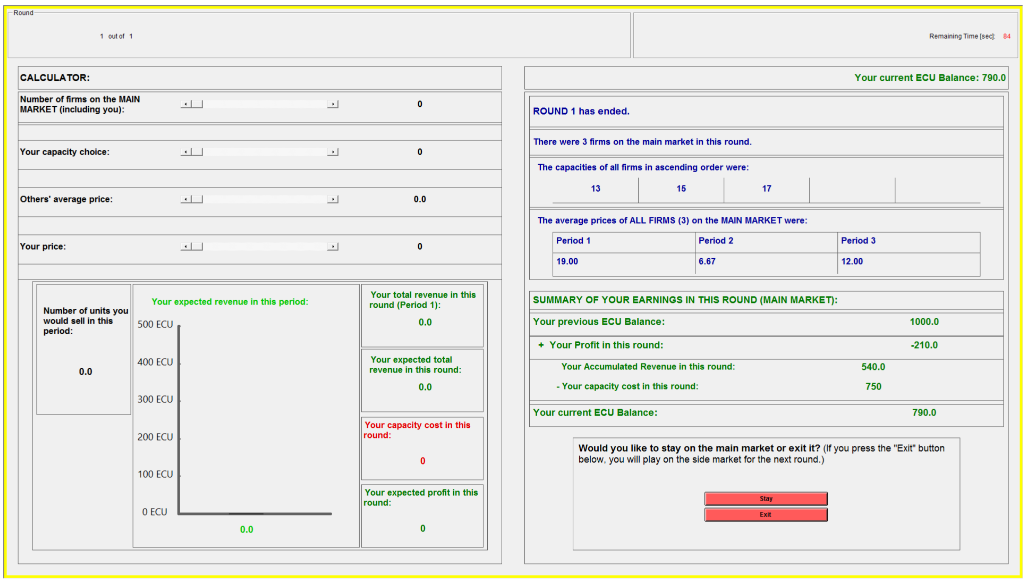


Figure B.4: Screen at the end of a round in the Market Force treatment

B.2 Instructions for Market Force treatment

Welcome and thank you for participating. In this experiment you will have the opportunity to earn a certain number of ECUs (experimental currency units). At the end of todays session we will convert these ECUs to an amount of money, using an exchange rate of 175 ECUs = 1 Euro. We will pay you in cash and in private. How much money you earn will depend on your decisions so please follow these instructions carefully.

Please note the following before we begin: It is very important to us that all participants are focussed exclusively on their own decision making. Please do not talk to the other people and also do not communicate with them by any other means. If you have a question during the experiment, just raise your hand and we will come to you.

At the beginning of the experiment the computer will randomly allocate the participants in this session to **groups of 5**. The experiment consists of **10 rounds**, and in these rounds you will interact with the other four people in your group but not with the members of other groups. You will not learn who has been allocated to your group.

What is this about?

In todays experiment you assume the role of a firm. Your task will be to decide how much to produce and at what price you want to offer your product. Furthermore, in each round you will choose whether you want to operate on the **main market** or on the **side market**. The other members of your group must take the same decisions for themselves. Your starting capital at the beginning is ECU 1,000. Depending on the decisions that you and others make you may earn additional ECUs. Losses are possible too, however.

The organization of a round

At the beginning of round 1 you first decide whether you wish to go on the main market or on the side market for this round. So that you are able to evaluate the consequences of your decision we will provide you with a calculator via the software that allows you to run through possible scenarios and how they affect your ECU earnings in advance. We will describe this calculator in more detail further below.

If you choose the main market you will compete with those in your group who have also selected the main market for the current round. If you choose the side market you will not compete with other firms in the current round (that is, there is a separate side market for each firm). You and the other four members of your group make the main market/side market decision simultaneously and individually. Therefore, you will only learn how many other firms have chosen the main market after you have submitted your own decision. Up to 5 firms in total (all members of a group) can simultaneously operate on the main market.

The main market

As soon as everyone has made their decision you will be informed about the number of firms on the main market. The round then consists of two phases. In phase 1 each firm decides on its **production capacity** for the current round, and in phase 2 the firms on the main market compete in 3 consecutive periods by choosing **prices** for their goods. After the third period the round ends.

A firms production capacity determines how man units it can maximally produce in each period of phase 2. At the beginning of phase 2 the production capacities of all firms involved will be announced.

In each of the three periods of phase 2 each firm chooses a **price** (in ECUs) for its product. At the end of each period the computer then calculates how many units each firm sells. For this a particular formula is used, which is of course the same that underlies the already mentioned calculator that allows you to try out how different prices affect your earnings.

Basically, there are the following three principles:

Principle 1: The higher your price the fewer units you will sell: For each ECU you charge more you sell two units less. The number of units you sell can even fall to zero. Principle 2: The higher the prices of your competitors the more units you will sell: For each ECU that the other firms average price increases you sell one additional unit. Principle 3: Per period you cannot sell more units than you have production capacity. Please note that at this stage you cannot change the production capacity you have chosen in phase 1. Each capacity unit costs ECU 50. You bear these costs only once, in phase 1. In phase 2 (in the 3 periods) there are no costs.

Here are two examples. The numbers are completely arbitrary and are just for illustration. **Example 1:** A firm chooses a production capacity of 5 units. This means that it can sell up to 5 units in each of the 3 periods. Its costs in phase 1 are 5 ECU 50 = ECU 250. In period 1 it chooses a price of ECU 10 and sells 5 units (although it would have been able to sell more units at this price if it had chosen a higher capacity level in phase 1). In period 2 it chooses a price of ECU 20 and again sells exactly 5 units due to its constrained capacity. In period 3 it chooses a price of ECU 30 and this time it sells 4 units.

The firms account at the end of the round:

Account balance at the start:			ECU	1,000
Cost for 5 units of capacity:	5 × ECU –50	= ECU –250		
Revenue in period 1:	5 × ECU 10	= ECU 50		
Revenue in period 2:	5 × ECU 20	= ECU 100		
Revenue in period 3:	4 × ECU 30	= ECU 120		
Profit in this round:		**= ECU 20**		
Updated account balance:			ECU	1,020

Example 2: A firm chooses a production capacity of 30 units. This means that it can sell up to 30 units in each of the three periods. Its costs in phase 1 are 30 ECU 50 = ECU 1,500. In period 1 it chooses a price of ECU 25 and sells 25 units. In period 2 it chooses a price of ECU 40 and sells 0 units. In period 3 it chooses a price of ECU 15

and sells 20 units.

The firms account at the end of the round:

Account balance at the start:			ECU	1,000
Cost for 30 units of capacity:	30 × ECU –50	= ECU –1,500		
Revenue in period 1:	25 × ECU 25	= ECU 625		
Revenue in period 2:	0 × ECU 40	= ECU 0		
Revenue in period 3:	20 × ECU 15	= ECU 300		
Profit in this round:		**= ECU –575**		
Updated account balance:			ECU	425

Investing in capacity is necessary in order for you to be able to obtain a profit. Note also that you have the possibility to invest an amount that exceeds your current account balance.

At the end of the round the firms that are present on the main market individually decide for themselves whether they wish to remain on the main market for the subsequent round or wish to exit the main market. Firms that decide to exit will operate on the side market during the subsequent round. There is one special rule: All firms whose account balance falls below ECU 300 at the end of the round must leave the main market.

At the same time the firms that are not present on the main market individually decide for themselves whether they want to enter the main market. Entering the main market requires an account balance of at least ECU 300, however. Furthermore, entry is also restricted in an-other respect: The main market is allowed to grow by only one firm from round to round. An example: In the round that has just been completed there were 2 firms present on the main market. One of these firms now decides to leave the main market, and the other one chooses to stay on the main market. Meanwhile, all firms that were not previously on the main market now want to enter the main market. However, because the main market is allowed to grow by only one firm overall (to then 3 firms), only two of the firms interested in the main market are admitted.

The selection of these two firms is done at random. The place of the one firm that had previously operated on the main market is secure.

The side market

As already explained above the side market and the main market differ in that there is no interaction between firms on the side market. Thus, each firm operates in its own individual side market. Another difference between the main and the side market is that the demand level on the side market fluctuates randomly from round to round while the demand situation on the main market remains in principle the same (although the demand for your product is affected by the number of firms on the main market and by their actions).

In all other respects the side market and the main market are practically identical. Here, too, you determine first (in phase 1) your production capacity and then (in phase 2) choose prices in three consecutive periods. At the end of the round the computer shows you a statement with your expenditures for ca-pacity and your revenue from the three periods. Moreover, you get a summary of what has happened on the main market in the current round (number of firms, chosen capacity levels, average prices). Finally, you will be asked whether you want to enter the main market or stay on the side market.

The calculator

As already mentioned above, our software provides you with a calculator which you can use to run through possible scenarios. The calculator for the main market has four sliders that allow you to set (a) the number of firms on the main market, (b) your production capacity, (c) your price and (d) the average price of the other firms on the main market. The calculator shows you your expected costs and revenues (the costs and revenues that would emerge in the scenario you selected by choosing values for a, b, c and d).

Figure B.5: The four sliders (initial settings)

During the round the calculator becomes somewhat simpler: As soon as the actual number of firms on the main market is known, the first slider disappears and the calculator uses only the actual number of firms, and as soon as you have chosen your production capacity the second slider also disappears and the calculator uses only your actual capacity.

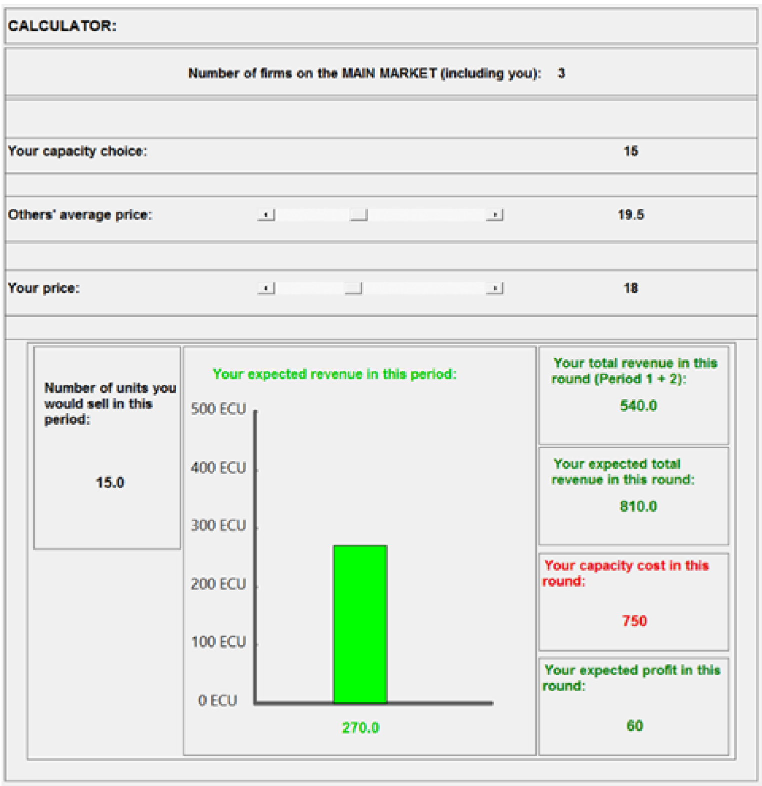


Figure B.6: The sliders again (during phase 2) and the output for particular prices

Notes: The example here displays a case where the firm is currently choosing its price for period 2. As shown the actual number of firms on the main market is 3 and the firm has chosen a capacity of 15 units for this round. The calculator is set to show what would happen if the firms own price is ECU 18 and the others average price turns out to be ECU 19.5. The result: The firm would sell 15 units and obtain a revenue of ECU 270 (= 15 ECU 18). On the right hand side you see the currently expected total revenue from periods 1 and 2 (if the ECU 270 really materialize in period 2), the projected total revenue from all 3 periods (if the ECU 270 materialize in both period

2 and then once again in period 3), the capacity cost of ECU 750 (15 units ECU 50) and finally the expected profit (expected revenue from all 3 periods minus the ECU 750).

We provide you with a very similar calculator on the side market when you operate on the side market. However, at the very beginning of the session and in between the rounds (i.e. whenever you have to decide whether you wish to enter the main market) the computer will, for simplicity, display only the calculator for the main market. To give you some orientation: On the side market you can expect a profit of between approximately ECU 160 and approximately ECU 330 per round, if you choose your capacity and your prices optimally. Your potential profit on the main market depends on the individual decisions of all firms that are involved.

Final remarks

On the main market prices up to ECU 50 and individual production capacity levels of up to 50 units are allowed. Only whole numbers are permissible. On the side market prices can also be up to ECU 50 but the maximum possible capacity is only 20 units because demand levels are generally lower than on the main market.

You have to submit your decisions within certain time limits. In the first 5 rounds these time limits are more generous than in the last 5 rounds. At the beginning of the very first round, you have a one-off opportunity to spend 5 minutes on trying out the calculator intensively and to take a first decision between main market and side market. For rounds 2 to 5 you have 2 minutes for selecting main or side market. You must choose your production capacity in a round within 2 minutes and your price in a period within 90 seconds.

After round 5 you will be reminded that you are now entering the second half of the experiment. The second half also consists of 5 rounds and is identical to the first half in every respect. The only difference is that from now on you only have 90 seconds to choose between main market and side market and that you only have 60 seconds to choose your price. The time limit for your capacity choice remains at 2 minutes.

When the time is up and you have not submitted a decision, the computer will in each case automatically choose an option for you, and to do this it will submit the value that you, at that precise moment, had selected for the relevant variable in your calculator (production capacity or price). If you fail to submit the main market/side market decision in time, the computer will choose the side market for you. The applicable countdown is always shown in the top right corner of your screen.

If you have any questions, now or during the experiment, please raise your arm and we will come to you.

C Appendix Chapter 4

C.1 Theoretical considerations

Consider a team of n players working together on a project. The individual investment in period t is denoted by $x_{i,t}$ $\forall$ $i = 1, 2, ..., n$ and the total individual investment until period t, including investments in period t, is denoted by $\sum_{k=1}^{t} x_{i,k} = y_i^t$. Hence, the total aggregated team investment at the end of period t is given by $\sum_{i=1}^{n} y_i^t = Y^t$.

In order to complete a project the teams' total investment has to meet the threshold of $H = 50n$. In each period, team members have to decide how much to invest. The amount that is still missing to reach the threshold is denoted by $Z_t = H - Y^{t-1}$ in period t.[1]

Assume that the trigger strategy is such that the total investment in period t is equal to $\hat{X}^t$. If someone deviates and the total investment in a periods is $X^t < \hat{X}^t$ everyone will stop their investment and the project is not completed.

Assume further that until period $\tilde{t} - 1$ everyone has invested such that $X^t \geq \hat{X}^t$ but in period $\tilde{t}$ there is a deviation and $X^{\tilde{t}} < \hat{X}^t$. Hence, the total investment at the end of period $\tilde{t}$ is $Y^{\tilde{t}} = (t - 1)X^t + X^{\tilde{t}}$. According to the trigger strategy all players should stop their investments in the following periods $t \geq \tilde{t} + 1$. Hence, $x_{i,t} = 0$ for i and $\forall$ $t > \tilde{t}$ and the project is not completed.

This trigger strategy is only credible if none of the players have an incentive to continue investing and completing the project on his/her own. In period $\tilde{t}+1$ the team is still missing $Z_{\tilde{t}+1} = H - Y^{\tilde{t}}$ to reach the threshold of H. Consider a player i that did not deviate in the previous periods. In order for this player to have an incentive to continue to invest to complete the project given that all other players stick to the trigger strategy and stop their investments, three conditions have to be fulfilled:

(i) $T \leq 30$ (Time restriction)

(ii) $Z_{\tilde{t}+1} \leq 75 - y_i^{\tilde{t}}$ (Budget restriction)

[1]Note: When players make their investment decisions in period t the aggregated team investment is Y^{t-1}.

(iii) Payoff of competing the project has to be larger or equal to sticking to the trigger strategy $75 + 90 - 2(T - 5) - y_i^T \leq 75 - y_i^{\tilde{t}}$

For (i): As the maximum investment per period is 10 ECUs, i would need at least

$$T - \tilde{t} = \frac{H - Y^{\tilde{t}}}{10} = \frac{50n - Y^{\tilde{t}}}{10}$$

additional periods to complete the project. Taking into account that the maximum number of periods is only 30:

$$\frac{50n - Y^{\tilde{t}}}{10} + \tilde{t} \leq 30$$

$$Y^{\tilde{t}} \geq 50n + 10\tilde{t} - 300$$

For (ii): Considering the case of symmetric investment, the budget restriction $Z_{\tilde{t}+1} \leq 75 - y_i^{\tilde{t}}$ can be written as

$$75 - \frac{Y^{\tilde{t}}}{n} \geq 50n - Y^{\tilde{t}}$$

$$Y^{\tilde{t}} \geq 25(2n - 3)\frac{n}{(n-1)}$$

For (iii):

$$75 + 90 - 2(T - 5) - Y^T \geq 75 - \frac{Y^{\tilde{t}}}{n}$$

Substituting $T = \frac{H - Y^{\tilde{t}}}{10} + \tilde{t}$ and solving for $Y^{\tilde{t}}$ gives the following condition:

$$Y^{\tilde{t}} \geq \frac{5}{3}t - \frac{250}{3} + 50n$$

When does which condition apply?

(i) > (ii) for

$$50n + 10t - 300 \geq 25(2n - 3)\frac{n}{(n-1)}$$

$$t \geq \frac{27.5n - 30}{(n-1)}$$

(iii) > (ii) for

$$\frac{5}{3}t - \frac{250}{3} + 50n \geq 25(2n-3)\frac{n}{(n-1)}$$

$$t \geq \frac{35n - 50}{(n-1)}$$

(i) > (iii) for

$$50n + 10t - 300 \geq \frac{5}{3}t - \frac{250}{3} + 50n$$

$$t \geq 26$$

For $n = 3$ and $n = 7$ condition (ii) is stricter than condition (iii) for $t < 28$ and for $t \geq 26$ (i) is stricter than (iii). Hence for $T \leq 30$ the trigger strategy to be credible the total team investment in period t has to fulfill 4.1

$$Y^t \leq \begin{cases} 25(2n-3)\frac{n}{(n-1)} & \text{for } t \leq \frac{27.5n-30}{(n-1)} \text{for } T \leq 30 \\ 50n + 10t - 300 & \text{for } t \geq \frac{27.5n-30}{(n-1)} \text{for } T \leq 30 \end{cases} \tag{C.1}$$

C.2 Screenshots

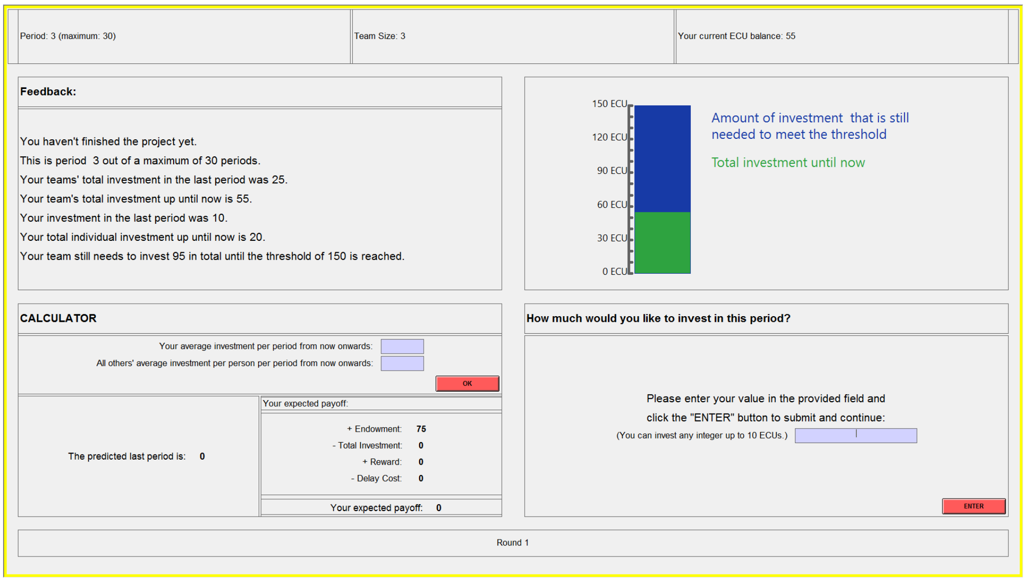


Figure C.1: Screen in treatment without deadline (three-person teams)

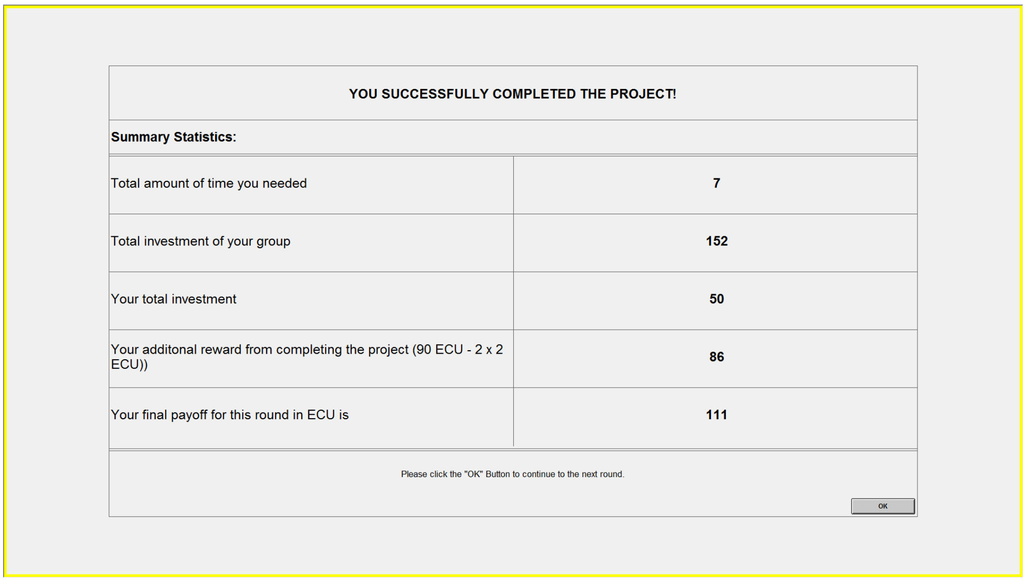


Figure C.2: Feedback at the end of each round

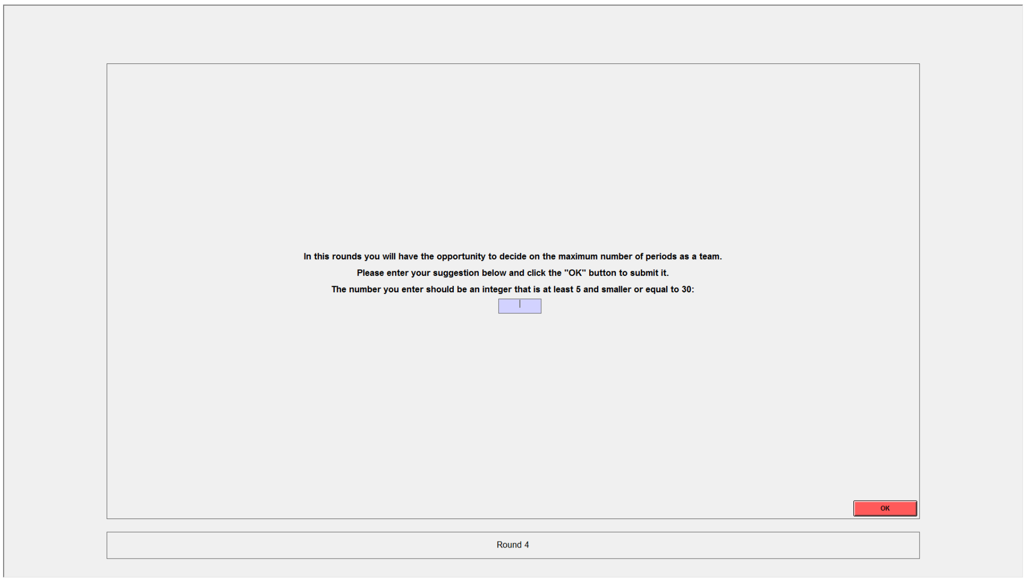


Figure C.3: Screen to enter suggestion for a deadline

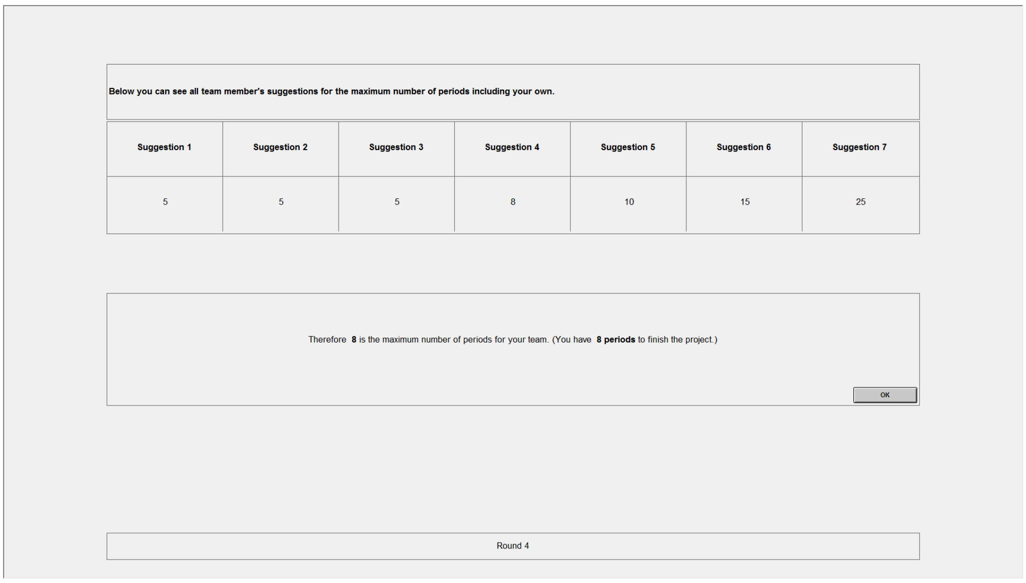


Figure C.4: Feedback on suggestions made by all team members and deadline

C.3 Instructions – Endogenous deadline, three-person teams

Instructions General rules

Welcome! You are about to take part in an experiment in the economics of decision making. You will be paid in private and in cash right after the end of the experiment. The amount you earn will depend on your decisions, so please follow the instructions carefully. During the experiment your payoffs are denominated in ECUs (Experimental Currency Units). Your ECU earnings will be converted to cash at the end of todays session at an exchange rate of 30 ECUs = 1 Euro. Thus, the more ECUs you earn the more cash you will take home today.

The experiment will consist of two parts. In the first part you will be working alone while in the second part you will be assigned to teams. Each part will consist of a certain number of rounds, as will be explained in detail in a few moments. At the end of the session one of these rounds will be selected at random. Your cash earnings from the experiment will be calculated based on your ECU earnings in this randomly selected round.

Note: It is important that you do not talk to any of the other participants until the experiment is over. If you have a question at any time, raise your hand and someone will come to your desk to answer it.

Instructions for Part 1

This part will consist of just a single round. This round, in turn, will consist of several periods (up to a maximum of 30, as explained below).

At the beginning of the round you receive an initial endowment of 300 ECUs. In each period of the round you will decide how many ECUs to invest into a project. You can invest any integer number between 0 and 10 ECUs in each period. Whatever you decide to invest into the project will be subtracted from your initial endowment and is non-refundable. Your total investment will be summed up over the periods. As soon as your total investment is 50 ECUs or greater the project is completed and the round ends. As a reward for the completion of the project you will receive an additional 300 ECUs.

Since you can invest no more than 10 ECUs in a single period and since the threshold for the completion of the project is 50 ECUs, you will need at least 5 periods to complete the project. If you do not complete the project within 5 periods an amount of 15 ECUs will be subtracted from the 300 ECU reward for each additional period you need to complete the project. If you do not complete the project within 30 periods the round ends and you will not receive the reward.

Thus, your earnings in this round will be calculated as:
Initial endowment Your total investment + Reward if you complete the project
The reward for completing the project is calculated as:
300 ECUs 15 ECUs for each period after period 5

You are provided with a calculator in the bottom left box of your screen. There you can enter your average investment per period starting from the current period your in and the computer will tell you what your last period and your expected payoff will be. You can use the calculator as often as you want.

Instruction for Part 2

This part consists of twelve rounds in total. At the beginning of each round you will be matched with two other participants, randomly selected from the people in this room. The three of you form a team. As before, your task is to decide how much you want to invest into a project but this time all members of your team invest into the same project. The way you take decisions will be the same in all rounds but the identity of the people you are teamed up with will change from round to round.

Each round consists of several periods. In the first three rounds the maximum number of periods is 30. In the remaining nine rounds you will have the opportunity to decide on the maximum number of periods as a team. This will be explained in detail below.

Just like in part 1 you receive an initial endowment of 300 ECUs and you can invest any integer number between 0 and 10 ECUs in each period. All team members make their investment decisions independently but at the end of each period you will be informed about your teams total investment, the total amount that you have invested thus far and how much your team still needs to invest to reach the threshold of 150 ECUs. As soon as your teams total investment (summed up over the periods) does reach or exceed the threshold the project is complete and the round ends. As a reward for the completion of the project each team member will receive an additional 300 ECUs.

Since you can invest no more than 10 ECUs per person in a single period and since the threshold for the completion of the project is 150 ECUs, you will need at least 5 periods to complete the project. If your team does not complete the project within 5 periods an amount of 15 ECUs will be subtracted from the 300 ECU reward for each additional period your team needs to complete the project. If your team does not complete the project within the maximum number of periods the round ends and you will not receive the reward.
Thus, your individual earnings in this round will be calculated as:
Initial endowment Your individual total investment + Reward if your team completes the project
The reward for completing the project is calculated as:
300 ECUs 15 ECUs for each period after period 5

In rounds 4-12 teams will use the following procedure to determine the maximum number of periods for completing the project. At the beginning of the round each team member will be asked to suggest a maximum number of periods (up to 30). As soon as all the team members have submitted a suggestion the computer will rank these suggestions from lowest to highest and will impose the middle one as the relevant maximum number of periods for that round.

As before, you are provided with a calculator in the bottom left box of your screen. There you can enter
- your average investment per period starting from the current period
- what you believe all others' average investment per person per period is going to be
Then the computer will tell you what the last period and your expected payoff will be in this case. You can use the calculator as often as you want.

C.4 Instructions – Exogenous deadline, seven-person teams

Instructions General rules

Welcome! You are about to take part in an experiment in the economics of decision making. You will be paid in private and in cash right after the end of the experiment. The amount you earn will depend on your decisions, so please follow the instructions carefully. During the experiment your payoffs are denominated in ECUs (Experimental Currency Units). Your ECU earnings will be converted to cash at the end of todays session at an exchange rate of 30 ECUs = 1 Euro. Thus, the more ECUs you earn the more cash you will take home today.

The experiment will consist of two parts. In the first part you will be working alone while in the second part you will be assigned to teams. Each part will consist of a certain number of rounds, as will be explained in detail in a few moments. At the end of the session one of these rounds will be selected at random. Your cash earnings from the experiment will be calculated based on your ECU earnings in this randomly selected round.

Note: It is important that you do not talk to any of the other participants until the experiment is over. If you have a question at any time, raise your hand and someone will come to your desk to answer it.

Instructions for Part 1

This part will consist of just a single round. This round, in turn, will consist of several periods (up to a maximum of 30, as explained below).

At the beginning of the round you receive an initial endowment of 75 ECUs. In each period of the round you will decide how many ECUs to invest into a project. You can invest any integer number between 0 and 10 ECUs in each period. Whatever you decide to invest into the project will be subtracted from your initial endowment and is non-refundable. Your total investment will be summed up over the periods. As soon as your total investment is 50 ECUs or greater the project is completed and the round ends. As a reward for the completion of the project you will receive an additional 90 ECUs.

Since you can invest no more than 10 ECUs in a single period and since the threshold for the completion of the project is 50 ECUs, you will need at least 5 periods to complete the project. If you do not complete the project within 5 periods an amount of 2 ECUs will be subtracted from the 90 ECU reward for each additional period you need to complete the project. If you do not complete the project within 30 periods the round ends and you will not receive the reward.

Thus, your earnings in this round will be calculated as:
Initial endowment Your total investment + Reward if you complete the project
The reward for completing the project is calculated as:
90 ECUs 2 ECUs for each period after period 5

You are provided with a calculator in the bottom left box of your screen. There you can enter your average investment per period starting from the current period youre in and the computer will tell you what your last period and your expected payoff will be. You can use the calculator as often as you want.

Instruction for Part 2

This part consists of nine rounds in total. At the beginning of each round you will be matched with six other participants, randomly selected from the people in this room. The seven of you form a team. As before, your task is to decide how much you want to invest into a project but this time all members of your team invest into the same project. The way you take decisions will be the same in all rounds but the identity of the people you are teamed up with will change from round to round.

Each round consists of several periods. In the first three rounds the maximum number of periods is 30. In the remaining six rounds the maximum number of rounds will change from round to round. You will be informed about the current maximum number of periods at the beginning of each round.

Just like in Part 1 you receive an initial endowment of 75 ECUs and you can invest any integer number between 0 and 10 ECUs in each period. All team members make their investment decisions independently but at the end of each period you will be informed about your teams total investment, the total amount that you have invested thus far and how much your team still needs to invest to reach the threshold of 350 ECUs. As soon as your teams total investment (summed up over the periods) does reach or exceed the threshold the project is complete and the round ends. As a reward for the completion of the project each team member will receive an additional 90 ECUs.

Since you can invest no more than 10 ECUs per person in a single period and since the threshold for the completion of the project is 350 ECUs, you will need at least 5 periods to complete the project. If your team does not complete the project within 5 periods an amount of 2 ECUs will be subtracted from the 90 ECU reward for each additional period your team needs to complete the project. If your team does not complete the project within the maximum number of periods the round ends and you will not receive the reward.
Thus, your individual earnings in this round will be calculated as:
Initial endowment Your individual total investment + Reward if your team completes the project
The reward for completing the project is calculated as:
90 ECUs 2 ECUs for each period after period 5

As before, you are provided with a calculator in the bottom left box of your screen. There you can enter
- your average investment per period starting from the current period
- what you believe all others' average investment per person per period is going to be
Then the computer will tell you what the last period and your expected payoff will be in this case. You can use the calculator as often as you want.

Bibliography

Aghion, Philippe, Patrick Bolton, Christopher Harris, and Bruno Jullien, "Optimal learning by experimentation," *The Review of Economic Studies*, 1991, *58* (4), 621–654.

Akerlof, George A, "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, 1970, *84* (3), 488–500.

Anderson, Lisa R., "Payoff Effects in Information Cascade Experiments," *Economic Inquiry*, October 2001, *39* (4), 609–15.

Ariely, Dan and Klaus Wertenbroch, "Procrastination, Deadlines, and Performance: Self-Control by Precommitment," *Psychological Science*, 2002, *13* (3), 219–224. PMID: 12009041.

Becker, Gordon M, Morris H DeGroot, and Jacob Marschak, "Measuring utility by a single-response sequential method," *Behavioral science*, 1964, *9* (3), 226–232.

Benndorf, Volker, Dorothea Kübler, and Hans-Theo Normann, "Privacy concerns, voluntary disclosure of information, and unraveling: An experiment," *European Economic Review*, 2015, *75* (C), 43–59.

Bisin, Alberto and Kyle Hyndman, "Present-Bias, Procrastination and Deadlines in a Field Experiment," Working Paper 19874, National Bureau of Economic Research January 2014.

Blume, Andreas, "A Learning-Efficiency Explanation of Structure in Language," *Theory and Decision*, November 2005, *57* (3), 265–285.

Bonatti, Alessandro and Johannes Hörner, "Collaborating," *American Economic Review*, April 2011, *101* (2), 632–63.

Brown, Alexander L, Colin F Camerer, and Dan Lovallo, "To review or not to review? Limited strategic thinking at the movie box office," *American Economic Journal: Microeconomics*, 2012, *4* (2), 1–26.

_ , _ , and _ , "Estimating structural models of equilibrium and cognitive hierarchy thinking in the field: The case of withheld movie critic reviews," *Management Science*, 2013, *59* (3), 733–747.

Burdea, Valeria, Maria Montero, and Martin Sefton, "Communication situations with partially verifiable information: an experimental approach," Technical Report, Mimeo 2016.

Burger, Nicholas, Gary Charness, and John Lynham, "Field and online experiments on self-control," *Journal of Economic Behavior & Organization*, 2011, *77* (3), 393–404.

Burks, Jeffrey J, Christine Cuny, Joseph Gerakos, and Joao Granja, "Competition and voluntary disclosure: Evidence from deregulation in the banking industry," Technical Report 12-29 2015.

Cadsby, Charles and Elizabeth Maynes, "Voluntary provision of threshold public goods with continuous contributions: experimental evidence," *Journal of Public Economics*, 1999, *71* (1), 53–73.

Cai, Hongbin and Joseph Tao-Yi Wang, "Overcommunication in strategic information transmission games," *Games and Economic Behavior*, 2006, *56* (1), 7–36.

Chamberlin, Edward H., "An Experimental Imperfect Market," *Journal of Political Economy*, 1948, *56* (2), 95–108.

Crawford, Vincent P. and Joel Sobel, "Strategic Information Transmission," *Econometrica*, November 1982, *50* (6), 1431–51.

Dannenberg, Astrid, Andreas Lschel, Gabriele Paolacci, Christiane Reif, and Alessandro Tavoni, "Coordination under threshold uncertainty in a public goods game," GRI Working Papers 64, Grantham Research Institute on Climate Change and the Environment November 2011.

Dranove, David and Ginger Zhe Jin, "Quality disclosure and certification: Theory and practice," *Journal of Economic Literature*, 2010, *48* (4), 935–963.

Dziuda, Wioletta, "Strategic argumentation," *Journal of Economic Theory*, 2011, *146* (4), 1362–1397.

Eckel, Catherine C. and Philip J. Grossman, "Sex differences and statistical stereotyping in attitudes toward financial risk," *Evolution and Human Behavior*, 2002, *23* (4), 281 – 295.

Eckel, Catherine, Philip Grossman, and Angela Milano, "Is More Information Always Better? An Experimental Study of Charitable Giving and Hurrican Katrina," *Southern Economic Journal*, 2007, *74* (2), 388–411.

Felgenhauer, Mike and Elisabeth Schulte, "Strategic private experimentation," *American Economic Journal: Microeconomics*, 2014, *6* (4), 74–105.

Fischbacher, Urs, "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, June 2007, *10* (2), 171–178.

Forsythe, Robert, R Mark Isaac, and Thomas R Palfrey, "Theories and tests of "blind bidding" in sealed-bid auctions," *The Rand Journal of Economics*, 1989, pp. 214–238.

Fouraker, Lawrence E. and Sidney Siegel, *Bargaining and group decision making: Experiments in bilateral monopoly*, Vol. 134, New York: McGraw-Hill, 1960.

_ **and** _ , "Bargaining Behavior," *Journal of Political Economy*, 1963.

Frchette, Guillaume R., "Laboratory Experiments: Professionals Versus Students," in "Handbook of Experimental Economic Methodology," Oxford University Press, 2015, chapter 17.

Friedman, James W., "Individual Behaviour in Oligopolistic Markets: An Experimental Study," *Yale Economic Essays*, 1963, *3*, 359417.

_ , "An Experimental Study of Cooperative Duopoly," *Econometrica*, 1967, *35* (3/4), 379–397.

Friedman, Milton, "The Methodology of Positive Economics," in Milton Friedman, ed., *Essays in Positive Economics*, University of Chicago Press, 1953, pp. 3–43.

Glazer, Jacob and Ariel Rubinstein, "Debates and Decisions: On a Rationale of Argumentation Rules," *Games and Economic Behavior*, August 2001, *36* (2), 158–173.

_ **and** _ , "On Optimal Rules of Persuasion," *Econometrica*, 2004, *72* (6), 1715–1736.

_ **and** _ , "A study in the pragmatics of persuasion: a game theoretical approach," *Theoretical Economics*, 2006, *1* (4), 395–410.

Goeree, Jacob and Charles Holt, "An experimental study of costly coordination," *Games and Economic Behavior*, 2005, *51* (2), 349–364.

Greiner, Ben, "The online recruitment system orsee 2.0-a guide for the organization of experiments in economics," *University of Cologne, Working paper series in economics*, 2004, *10* (23), 63–104.

Griffin, John D. and Alice H. Lichtenstein, "Dietary Cholesterol and Plasma Lipoprotein Profiles: Randomized-Controlled Trials," *Current Nutrition Reports*, 2013, *2* (4), 274–282.

Grossman, Sanford J., "The Informational Role of Warranties and Private Disclosure about Product Quality," *The Journal of Law & Economics*, 1981, *24* (3), 461–483.

_ **and O. D. Hart**, "Disclosure Laws and Takeover Bids," *The Journal of Finance*, 1980, *35* (2), 323–334.

Guala, Francesco, "On the scope of experiments in economics: comments on Siakantaris," *Cambridge Journal of Economics*, 2002, *26* (2), 261–267.

Hagenbach, Jeanne and Eduardo Perez-Richet, "Communication with Evidence in the Lab," Technical Report, Mimeo 2015.

_ , Frédéric Koessler, and Eduardo Perez-Richet, "Certifiable Pre-Play Communication: Full Disclosure," *Econometrica*, 2014, *82* (3), 1093–1131.

Holt, Charles A., *Industrial Organization: A Survey of Laboratory Research*, In The Handbook of Experimental Economics, 1995.

Huck, Steffen and Jidong Zhou, "Consumer behavioural biases in competition: A survey," Technical Report 2011.

Huyck, John B. Van, Raymond C. Battalio, and Richard O. Beil, "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure," *American Economic Review*, March 1990, *80* (1), 234–248.

_ , _ , and _ , "Strategic Uncertainty, Equilibrium Selection, and Coordination Failure in Average Opinion Games," *The Quarterly Journal of Economics*, 1991, *106* (3), 885–910.

Jin, Ginger Zhe, "Competition and disclosure incentives: an empirical study of HMOs," *RAND Journal of economics*, 2005, pp. 93–112.

_ , Michael Luca, and Daniel Martin, "Is No News (Perceived as) Bad News? An Experimental Investigation of Information Disclosure," Working Paper 21099, National Bureau of Economic Research August 2016.

Johnson, Eric, Stephan Meier, and Olivier Toubia, "Money Left on the Kitchen Table: Exploring sluggish martgage refinancing using administrative data, surveys, and field experiments," Technical Report, Columbia Business School 2015.

Kaas, Klaus and Heidrun Ruprecht, "Are the Vickrey auction and the BDM-mechanism really incentive compatible? Empirical results and optimal bidding strategies in the case of uncertain willingness-to-pay," *Schmalenbach Business Reports*, 2006, *58*, 37–55.

Kamenica, Emir and Matthew Gentzkow, "Bayesian Persuasion," *American Economic Review*, 2011, *101* (6), 2590–2615.

Kreps, David M. and Jose A. Scheinkman, "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes," *The Bell Journal of Economics*, 1983, *14* (2), 326–337.

Külpmann, Philipp, "Procrastination and projects," Center for Mathematical Economics Working Papers 544, Center for Mathematical Economics, Bielefeld University July 2015.

Lesser, Lenard I, Cara B Ebbeling, Merrill Goozner, David Wypij, and David S Ludwig, "Relationship between funding source and conclusion among nutrition-related scientific articles," *PLoS Med*, 2007, *4* (1), e5.

Liebrand, Wim B. G., "The effect of social motives, communication and group size on behaviour in an N-person multi-stage mixed-motive game," *European Journal of Social Psychology*, 1984, *14* (3), 239–264.

McDougall, John A., "The Egg Industry: Exposing a Source of Food Poisoning," https://www.drmcdougall.com/misc/2016nl/jan/eggindustry.htm 2016. Accessed: 2018-05-20.

Milgrom, Paul and John Roberts, "Relying on the Information of Interested Parties," *The RAND Journal of Economics*, 1986, *17* (1), pp. 18–32.

Milgrom, Paul R., "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, Autumn 1981, *12* (2), 380–391.

Potters, Jan and Sigrid Suetens, "Oligopoly experiments in the current millennium," *Journal of Economic Surveys*, 2013, *27* (3), 439–460.

Rau, Holger, "The disposition effect and loss aversion: Do gender differences matter?," *Economics Letters*, 2014, *123* (1), 33–36.

Rau, Holger Andreas and Hans-Theo Normann, "Step-Level Public Goods: Experimental Evidence," Annual Conference 2011 (Frankfurt, Main): The Order of the World Economy - Lessons from the Crisis 48710, Verein fr Socialpolitik / German Economic Association 2011.

Raven, John C., "Mental tests used in genetic studies: The performance of related individuals on tests mainly educative and mainly reproductive," Master's thesis, University of London 1936.

—, *Manual for ravens progressive matrices and vocabulary scales. Oxford Psychologists Press* The Psychological Corporation 1981.

Rothschild, Michael, "A two-armed bandit theory of market pricing," *Journal of Economic Theory*, October 1974, *9* (2), 185–202.

Sauermann, Heinz and Reinhard Selten, "Ein Oligopolexperiment," *Zeitschrift fr die gesamte Staatswissenschaft / Journal of Institutional and Theoretical Economics*, 1959, *115* (3), 427–471.

Shin, Hyun Song, "News management and the value of firms," *The RAND Journal of Economics*, 1994, pp. 58–71.

Siakantaris, Nikos, "Experimental Economics under the Microscope," *Cambridge Journal of Economics*, 2000, *24* (3), 267–81.

Smith, Vernon L., "An Experimental Study of Competitive Market Behavior," *Journal of Political Economy*, 1962, *70* (2), 111–137.

— , "Effect of Market Organization on Competitive Equilibrium," *The Quarterly Journal of Economics*, 1964, *78* (2), 182–201.

Viscusi, W Kip, "A note on "lemons" markets with quality certification," *The Bell Journal of Economics*, 1978, pp. 277–279.

Wang, Joseph Tao-Yi, Michael Spezio, and Colin F Camerer, "Pinocchio's pupil: using eyetracking and pupil dilation to understand truth telling and deception in sender-receiver games," *The American Economic Review*, 2010, *100* (3), 984–1007.

Erklärung

Ich versichere hiermit, dass ich die Dissertation selbstständig und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Mannheim, den 28. Mai 2018

Curriculum Vitae

Sihong Zhang

Since 09/2014	PhD Studies in Economics Center for Doctoral Studies in Economics (CDSE) Graduate School of Economic and Social Sciences (GESS) University of Mannheim, Germany
09/2012 - 08/2014	Master of Economic Research (M.Sc.) University of Mannheim, Germany
10/2009 - 08/2012	Bachelor of Science in Economics (B.Sc.) Ludwig-Maximilians University Munich, Germany
07/2009	Abitur Berthold-Gymnasium, Freiburg